

Н. З. ШОР

**МЕТОДЫ НЕДИФФЕРЕНЦИРУЕМОЙ
ОПТИМИЗАЦИИ И СЛОЖНЫЕ
ЭКСТРЕМАЛЬНЫЕ ЗАДАЧИ**



КИШИНЭУ • ЭВРИКА • 2008

НАЦИОНАЛЬНАЯ АКАДЕМИЯ НАУК УКРАИНЫ
ИНСТИТУТ КИБЕРНЕТИКИ ИМ. В. М. ГЛУШКОВА

МИНИСТЕРСТВО ПРОСВЕЩЕНИЯ И МОЛОДЕЖИ
РЕСПУБЛИКИ МОЛДОВА

АКАДЕМИЯ ТРАНСПОРТА,
ИНФОРМАТИКИ И КОММУНИКАЦИЙ

Н. З. ШОР

МЕТОДЫ НЕДИФФЕРЕНЦИРУЕМОЙ
ОПТИМИЗАЦИИ И СЛОЖНЫЕ
ЭКСТРЕМАЛЬНЫЕ ЗАДАЧИ

Сборник избранных трудов

КИШИНЭУ * ЭВРИКА * 2008

CZU 519.85

III 79

Descrierea CIP a Camerei Nationale a Cartii

Шор, Н. З.

Методы недифференцируемой оптимизации и сложные экстремальные задачи: Сб. избранных трудов / Н. З. Шор; редкол.: И. В. Сергиенко, . . . ; нац. Акад. Наук Украины, Ин-т кибернетики им. В. М. Глушкова, Акад. Транспорта, информатики и коммуникаций. – К.: Эврика, 2008. – 270 с.

ISBN 978-9975-941-19-8

150 ex.

CZU 519.85

Сборник содержит работы академика Н.З.Шора, основателя направления недифференцируемой оптимизации в теории математического программирования. В сборник включены его основные работы по методам недифференцируемой оптимизации и их применению в задачах большой размерности, квадратичного, булевого, матричного и стохастического программирования.

Книга рассчитана на специалистов в области математического программирования и его приложений, а также студентов и аспирантов соответствующих специальностей.

Редакционная коллегия:

академик И. В. СЕРГИЕНКО,

доктор техн. наук Д. И. СОЛОМОН,

канд. физ.-мат. наук П. И. СТЕЦЮК,

канд. физ.-мат. наук Н. Г. ЖУРБЕНКО

канд. физ.-мат. наук Т. А. БАРДАДЫМ

M-208-96

ISBN 978-9975-941-19-8

©Издательство „Эврика“, 2008

©Шор Н. З., 2008



АКАДЕМИК Н. З. ШОР (1937 – 2006)

"Теория всего класса алгоритмов с растяжением пространства ещё далека от совершенства. Нам кажется достаточно реалистической целью построение такого алгоритма, который по своей практической эффективности не уступал бы r-алгоритму и был столь же хорошо обоснован, как метод эллипсоидов".

Шор

Шор Наум Зуселевич – основоположник направления недифференцируемой оптимизации в математическом программировании, профессор, академик НАН Украины. Родился 1 января 1937 г. в г. Киеве. После окончания в 1958 г. Киевского национального университета имени Тараса Шевченко работал в Институте кибернетики им. В. М. Глушкова НАН Украины (с 1983 г. – заведующий лабораторией, с 1990 г. – заведующий отделом методов решения сложных задач оптимизации). Он автор 10 монографий и более 200 статей по математическому программированию, вычислительной математике и теории графов. Лауреат Государственных премий УССР (1973 г.), СССР (1981 г.), Украины (1993 г., 2000 г.), премии им. В. М. Глушкова НАН Украины (1987 г.), премии им. В. С. Михалевича НАН Украины (1997 г.).

Научное наследие Н. З. Шора – выдающийся вклад в развитие теории и методов оптимизации. Его работы положили начало новому направлению математического программирования – численным методам негладкой оптимизации, и существенно повлияли на развитие методов дискретной, матричной, полиномиальной и стохастической оптимизации.

Далеко не все работы Н. З. Шора отражены в его монографиях. Очень много осталось рассредоточенным по многочисленным статьям, которые ныне стали библиографической редкостью. Книга содержит избранные труды Н. З. Шора, связанные с решением структурированных задач выпуклого программирования большой размерности и задач матричной оптимизации. Были отобраны те статьи, которые либо открывали новые этапы в его творчестве, либо давали обзорное представление о разработанном направлении.

Сборник состоит из двух частей. В первой части помещены статьи о методах недифференцируемой оптимизации и схемах декомпозиции по переменным и ограничениям, статьи о двухэтапных стохастических задачах, где схема декомпозиции по переменным играет центральную роль. Во второй части собраны статьи, связанные с применением методов негладкой оптимизации для решения задач квадратичного, булевого и матричного программирования.

Вся профессиональная жизнь Н. З. Шора прошла в Институте кибернетики имени В.М. Глушкова. Здесь¹ он начал работать после окончания Киевского национального университета имени Тараса Шевченко по приглашению Виктора Михайловича Глушкова, здесь он прошел все ступени от инженера до руководителя отдела методов решения слож-

¹В 1958 он начал трудовую деятельность в Вычислительном центре Академии наук Украины, преобразованном в 1962 году в Институт кибернетики.

ных задач оптимизации. Это были времена интенсивного развития прикладной математики в Советском Союзе, а Институт кибернетики стал одной из ведущих организаций в области информатики, вычислительной техники, численных методов оптимизации [1] – [3].

Разработка методов недифференцируемой оптимизации стала в 60-х годах одной из первоочередных задач. Именно эти методы обеспечили возможность решения сложных практических задач оптимизации на базе вычислительной техники того времени. Методы недифференцируемой оптимизации позволяли гибко использовать различные схемы декомпозиции (по переменным, ограничениям, ресурсам и т. п.), учитывающие специфику задач большой размерности, эффективно получать двойственные оценки в задачах дискретного, непрерывно-дискретного программирования и для некоторых классов многоэкстремальных задач. Появляется возможность использовать негладкие функции штрафа, позволяющие при конечных значениях штрафных параметров получать задачу безусловной минимизации, эквивалентную первоначальной задаче выпуклого программирования. Техничко-экономические характеристики оптимизируемых объектов часто хорошо аппроксимируются кусочно-гладкими функциями от неизвестных параметров, что также порождает задачи оптимизации с негладкими функциями.

Значительную часть творческого наследия Н. З. Шора занимает разработка и исследование методов недифференцируемой оптимизации. Они явились основой его кандидатской диссертации "О структуре алгоритмов численного решения задач оптимального планирования и проектирования", 1964 г. [4]. А наиболее полно результаты по этой тематике, полученные в период с 1962 по 1978 год, отражены в монографии [5]. В нее был включен и ряд материалов его докторской диссертации "Методы минимизации недифференцируемых функций и их приложения" – её защита состоялась в 1970 году [6]. Эта книга получила огромную известность, была практически сразу же переведена на английский язык известными польскими учеными К. Кивелем и А. Рущиньским и издана в издательстве "Шпрингер" в 1985 г. [7]. В монографии [5] исследованы основные модификации метода обобщенного градиентного спуска, изложены алгоритмы субградиентного типа с растяжением пространства в направлении субградиента и разности двух последовательных субградиентов, которые обладают ускоренной сходимостью, а также применения методов минимизации недифференцируемых функций при решении разнообразных задач оптимального планирования и проектирования.

О сущности проблем, связанных с недифференцируемостью и о вкладе Н. З. Шора в их разрешение замечательно написал Б.Т. Поляк [8]:

"Основные алгоритмы минимизации гладких функций – градиентный и Ньютона – были построены на использовании линейной и квадратичной аппроксимации функции, задаваемой первыми членами ряда Тейлора. Однако, для недифференцируемой функции эта идея неприменима – такая функция не может быть хорошо аппроксимирована ни линейной, ни квадратичной функциями... Поэтому разработка методов минимизации негладких функций требует привлечения новых идей. Одна из них, принадлежащая Н.З.Шору, выглядит несколько неожиданно. Пишется прямой аналог градиентного метода с заменой градиента на произвольный субградиент $g_f(x)$ функции $f(x)$:

$$x_{k+1} = x_k - \gamma_k g_f(x_k) \quad (1)$$

[...]; значения функции в методе (1) не могут убывать монотонно. Оказывается, однако, что при этом монотонно убывает другая функция – расстояние до точки минимума, и в этом-то заключается основная идея субградиентного метода (1). "

Результаты Н.З.Шора по методам негладкой оптимизации можно условно разделить на три направления. Первое – методы обобщенного градиентного спуска (ОГС) (1962 – 1971 гг.), которые положили начало новому направлению математического программирования – численным методам негладкой оптимизации. Второе – субградиентные методы с растяжением пространства в направлении субградиента, которые в сравнении с методами ОГС имеют ускоренную сходимость. Эти методы дали теории оптимизации уникальный алгоритм – метод эллипсоидов, скорость сходимости которого зависит лишь от размерности пространства. Использование метода эллипсоидов позволило решить ряд важных вопросов в теории сложности задач математического программирования. Наиболее результативными эти исследования оказались в период с 1968 по 1986 гг.

И, наконец, третье направление – это субградиентные методы с растяжением пространства в направлении разности двух последовательных субградиентов – так называемые r -алгоритмы. До настоящего времени r -алгоритмы являются одним из наиболее эффективных средств решения задач недифференцируемой оптимизации. При минимизации гладких функций они оказались конкурентноспособными с наиболее удачными реализациями методов сопряженных направлений и методов квазиньютоновского типа. Начало исследований по r -алгоритмам пришлось на 1970–1974 гг., активно продолжались в последующие годы и продолжают учениками Н. З. Шора. К настоящему времени усилиями Н. З. Шора и его учеников Н. Г. Журбенко, Л. П. Шабашовой,

В. И. Гершовича, П. И. Стецюка, А. П. Лиховида, А. В. Кунцевича и др. разработано несколько модификаций r -алгоритма применительно к решению разнообразных задач оптимизации. r -алгоритм использовался в задачах оптимизации большой размерности и в блочных задачах с различными схемами декомпозиции, при решении минимаксных и матричных задач оптимизации, для вычисления двойственных лагранжевых оценок в многоэкстремальных и комбинаторных задачах оптимизации. На практике он применялся для решения задач оптимального планирования, оптимального проектирования, синтеза и анализа сетей, восстановления изображений, эллипсоидальной аппроксимации и локализации и др. [5, 9, 10].

Чтобы кратко дать представление о результатах Н. З. Шора по этой тематике, первой в данном сборнике помещена обзорная статья [11], напечатанная в номере журнала "Кибернетика", посвященном памяти В. М. Глушкова.

Методы ОГС дали возможность решить большое число задач производственно-транспортного планирования с применением схем декомпозиции для задач большой размерности. Подробную информацию об этих задачах можно найти в [5, 12]. Для данного сборника отобран ряд результатов Н. З. Шора по решению задач блочной структуры, в частности, в стохастическом программировании. В статье [13] с использованием ОГС впервые применена декомпозиция по ограничениям. Результаты Н. З. Шора по задачам стохастического программирования изложены в нескольких разделах его докторской диссертации ([6], с. 160–180) и большом количестве статей. Им были подробно исследованы двухэтапные и многоэтапные задачи, связанные с ними вопросы декомпозиции и особенности их численного решения.

Метод ОГС также послужил основой для создания стохастического аналога обобщенного градиентного спуска [14], который имеет большое практическое применение, в частности, при решении многоэтапных задач стохастического программирования. Эта статья также включена в сборник. В ней впервые введено и исследовано понятие стохастического квазиградиента.

В данный сборник помещены также статья [15] – в ней впервые использована оригинальная методика регулировки шагового множителя, и статья [16], в которой впервые был предложен r -алгоритм. В следующей статье [17] приводится частный случай алгоритма обобщенного градиентного спуска с растяжением пространства в направлении субградиента, в котором величина шага и коэффициента растяжения пространства определяются только размерностью пространства перемен-

ных. Этот вариант алгоритма полностью соответствует методу эллипсоидов, предложенному Д. Б. Юдиным и А. С.Немировским [18], исходя из методов последовательных отсечений. Этот алгоритм использован Л. Г. Хачияном [19] для построения и обоснования первого полиномиального алгоритма решения задачи линейного программирования с рациональными коэффициентами и получил важное применение в теории сложности алгоритмов дискретной оптимизации [20]. В статье [21] схема декомпозиции по ограничениям применена для транспортной динамической задачи в сетевой постановке. Эта статья содержит также описание алгоритма усреднения, позволяющего получить решение исходной задачи на основе решения двойственной к ней задачи.

Завершающая первую часть сборника статья [22] посвящена результатам выполнения проекта "Методы решения и аппроксимации в стохастическом программировании", который выполнялся совместно с группой профессора Питера Калла из Института исследования операций Цюрихского университета при поддержке Швейцарской национальной научной организации². В результате выполнения этого проекта были разработаны программы для решения двух специальных классов задач двухэтапного стохастического программирования. В этих программах были использованы разработанные Н.З. Шором методы негладкой оптимизации и применена схема декомпозиции по переменным. Программы включены в состав системы моделирования SLP-IOR³, разработанной в Институте исследования операций Цюрихского университета.

Завершая краткую характеристику первой части сборника, отметим следующее. Разработанные Н. З. Шором подходы к решению задач блочной структуры нашли многочисленные приложения. 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. Много прикладных задач блочной структуры описано в монографиях [5, 12, 23, 24]. Так в [5, 12] рассмотрены разнообразные задачи оптимального планирования и проектирования, задачи долгосрочного и текущего производственно-транспортного планирования, многоиндексные распределительные задачи в линейной, нелинейной и динамической постановках и др. Монография Н. З. Шора и Д. И. Соломона [23] посвящена методам решения блочных задач дробно-линейного програм-

²Swiss National Science Foundation, Project number: 7UKPJ048452, September 1997 – August 1998.

³http://www.ior.uzh.ch/Pages/English/Research/StochOpt/SLP_Users_Guide.pdf

мирования большой размерности. Здесь рассмотрены такие практические задачи оптимизации: определение клиентурного плана и структуры подвижного состава автотранспортных предприятий, маршрутизация автомобильных грузовых перевозок. Монография [24] содержит применение схем декомпозиции при решении ряда задач оптимального проектирования и маршрутизации в сетях с учетом отказа отдельных компонент сети и изменения требований к потокам. Эти работы были выполнены в рамках международного проекта "Математическое и программное обеспечение оптимального проектирования структур надежных сетей" Украинского научно-технологического центра⁴ – последнего международного проекта, выполненного под руководством Н. З. Шора.

Во вторую часть данного сборника вошли четыре статьи Н. З. Шора [25]–[28], связанные с применением методов недифференцируемой оптимизации в дискретных, полиномиальных и матричных задачах. Их объединяет использование подхода Лагранжа для получения двойственных оценок. Получение этих оценок очень важно для невыпуклых квадратичных моделей, которые встречаются во многих приложениях, например, в полиномиальных задачах оптимизации и NP-трудных экстремальных задачах на графах. Такой подход дает возможность среди NP-трудных невыпуклых квадратичных задач выделить такие подклассы, для которых проблема нахождения значения глобального минимума целевой функции разрешима за полиномиальное время.

Так получилось, что наиболее полная по материалам этой тематики книга [10] вышла за рубежом на английском языке, на русский не переводилась и является библиографической редкостью в странах СНГ. Мы надеемся, что публикация отобранных статей [25]–[28] хотя бы частично восполнит этот пробел. Статья [25], опубликованная на русском языке в "Обзрении прикладной и промышленной математики", также является труднодоступной. Она содержит прекрасный обзор по задачам матричной оптимизации, которые имеют многочисленные приложения в теории устойчивости и оптимизации динамических систем. Приведены примеры матричных задач, связанных с получением оптимальной функции Ляпунова, и методы их решения с помощью алгоритмов недифференцируемой оптимизации. К матричным задачам относятся и задачи построения оптимальных по объему вписанных в многогранник и описанных вокруг многогранника эллипсоидов, имеющие многочисленные приложения в теории оценивания, распознавании образов, при построении численных методов оптимизации. Оригинальные алгоритмы для решения этих задач разработаны Н. З. Шором вместе с уче-

⁴Проект № 1625.

никами С. И. Стеценко и О. А. Березовским. Часть этих результатов вошла в монографию [9], но написанная позже статья [25], без сомнения, является наиболее полной их коллекцией.

Вторая помещенная здесь статья [26] посвящена задачам полиномиального программирования. К этому классу относят задачи с целевой функцией, заданной полиномом и ограничениями в виде полиномиальных равенств и неравенств. В виде задач полиномиального программирования можно представить широкий класс моделей линейного целочисленного программирования и разнообразные модели комбинаторной и многоэкстремальной оптимизации.

В работах Н. З. Шора (см., например, [9]) исследуются так называемые двойственные или лагранжевы оценки квадратичных оптимизационных задач, в которых используются классические множители Лагранжа. Результатом такой процедуры есть оценка снизу (сверху) для оптимального значения целевой функции исходной задачи минимизации (максимизации). Задача нахождения оптимальных множителей Лагранжа относится к классу задач недифференцируемой оптимизации, а проблемы построения алгоритмов для определения лагранжевых оценок тесно связаны с применением методов недифференцируемой оптимизации.

Используя функцию Лагранжа, можно получать двойственные оценки, которые являются оценками снизу для глобального минимума полинома. Эти оценки могут быть улучшены за счет предложенной Н. З. Шором идеи введения избыточных ограничений. Одной из основных работ, посвященных этому вопросу, является статья [27], включенная во вторую часть. Оказалось, что указанные проблемы тесно связаны с исследованиями Гильберта о представлении неотрицательных полиномов в виде суммы квадратов полиномов меньших степеней (так называемая 17-я проблема Гильберта).

Статья [28] была опубликована за рубежом, и в ней достаточно полно представлены результаты как по невыпуклым полиномиальным задачам, так и по дискретным оптимизационным задачам.

Завершает сборник список научных трудов академика Н. З. Шора. По индексу цитирования его работы, особенно книги [5] и [10], занимают лидирующие позиции. Субградиентные методы недифференцируемой оптимизации, разработанные в Институте кибернетики под его руководством, оказали огромное влияние на развитие теории и практики для многих направлений математического программирования, получили высокую оценку мирового

научного сообщества. Надеемся, что включенные в этот сборник работы будут способствовать популяризации его результатов.

За выдающиеся заслуги профессор Н. З. Шор в 1990 году был избран членом-корреспондентом, а в 1997 году – академиком Национальной академии наук Украины.

Занимаясь педагогической деятельностью в Киевском отделении Московского физико-технического института, Национальном техническом университете (КПИ), Киевском национальном университете имени Тараса Шевченко, Соломоновом университете, он всегда помогал молодым ученым внимательными и доброжелательными научными консультациями, рекомендациями. Под его руководством было подготовлено и защищено много докторских и кандидатских диссертаций. Сейчас ученики Наума Зуселевича успешно работают во многих странах в различных областях прикладной математики.

Наума Зуселевича любили и уважали все, кто имел счастье работать с ним на протяжении сорока восьми лет его научной деятельности.

Научное наследие академика Н. З. Шора оказало существенное влияние на развитие теории и численных методов оптимизации и остается актуальным для будущего.

Редакционная коллегия сердечно благодарит Е. И. Шор за неоценимую помощь в сборе и подготовке материалов и всех, кто помогал готовить этот сборник – сотрудников отдела негладкой оптимизации, работников издательства "Эврика" и многих других.

Редакционная коллегия
Киев, Кишинэу, февраль 2008

Литература

1. СЕРГІЄНКО І.В. Інформатика в Україні: становлення, розвиток, проблеми. – К.: Наук. думка, 1999. – 354 с.
2. POLYAK В.Т. History of mathematical programming in the USSR: analyzing the phenomenon // Math. Program. – Ser. B. – 2001. – P. 1–16.
3. МАЛИНОВСКИЙ Б.Н. Хранить вечно = Store eternally. – К.: Горобец, 2007. – 176 с.: ил. – Рус., англ.
4. ШОР Н.З. О структуре алгоритмов численного решения задач оптимального планирования и проектирования. Диссертация на соискание ученой степени кандидата физ.-мат. наук. Киев, ИК АН УССР, 1964. – 135 с.
5. ШОР Н.З. Методы минимизации недифференцируемых функций и их приложения. – Киев: Наук. думка, 1979. – 199 с.
6. ШОР Н.З. Методы минимизации недифференцируемых функций и их приложения. Диссертация на соискание ученой степени доктора физ.-мат. наук. Киев, ИК АН УССР, 1970. – 203 с.
7. SHOR N.Z. Minimization Methods for Non-Differentiable Functions. Berlin: Springer-Verlag, 1985. – 178 p.
8. Поляк Б.Т. Введение в оптимизацию. – М.: Наука, 1983. – 384 с.
9. ШОР Н.З., СТЕЦЕНКО С.И. Квадратичные экстремальные задачи и недифференцируемая оптимизация. – К.: Наук. думка, 1989. – 208 с.
10. SHOR N.Z. Nondifferentiable optimization and polynomial problems. – Boston; Dordrecht; London: Kluwer Academic Publishers, 1998. – 394 p.
11. ШОР Н.З., ЖУРБЕНКО Н.Г., ЛИХОВИД А.П., СТЕЦЮК П.И. Развитие алгоритмов недифференцируемой оптимизации и их приложения // Кибернетика и системный анализ. – 2003. – № 4. – С. 80–94.
12. МИХАЛЕВИЧ В.С., ТРУБИН В.А., ШОР Н.З. Оптимизационные задачи производственно-транспортного планирования. Модели, методы, алгоритмы. // – М.: Наука, – 1986. – 260 с.

13. Шор Н.З. Применение обобщенного градиентного спуска в блочном программировании // Кибернетика. – 1967. – № 3. – С. 53–55.
14. Ермольев Ю.М., Шор Н.З. Метод случайного поиска для задач двухэтапного стохастического программирования и его обобщение // Кибернетика. – 1968. – № 1. – С. 90–92.
15. Шор Н.З., ЩЕПАКИН М.Б. Алгоритм решения двухэтапной задачи стохастического программирования // Кибернетика. – 1968. – № 3. – С. 56–58.
16. Шор Н.З., ЖУРБЕНКО Н.Г. Метод минимизации, использующий операцию растяжения пространства в направлении разности двух последовательных градиентов // Кибернетика. – 1971. – № 3. – С. 51–59.
17. Шор Н.З. Метод отсечения с растяжением пространства для решения задач выпуклого программирования // Кибернетика. – 1977. – № 1. – С. 94–95.
18. Юдин Д.Б., НЕМИРОВСКИЙ А.С. Информационная сложность и эффективные методы решения выпуклых экстремальных задач // Экономика и мат. методы. – 1976. – Вып. 2. – С. 357–359.
19. Хачиян Л.Г. Полиномиальные алгоритмы в линейном программировании // ДАН СССР. – 1979. – Т. 244. – № 5. – С. 1093–1096.
20. GRÖTSCHEL M., LOVASZ L., SCHRIJVER A. Geometric Algorithms and Combinatorial Optimization. – Springer-Verlag, Berlin. – 1988. – 362 p.
21. БЕЛЯЕВА Л.В., ЖУРБЕНКО Н.Г., ШОР Н.З. О методе решения одного класса динамических распределительных задач // Экономика и матем. методы. – 1978. – Т. 14. – Вып. 1. – С. 137–146.
22. Шор Н.З., БАРДАДЫМ Т.А., ЖУРБЕНКО Н.Г., ЛИХОВИД А.П., СТЕЦЮК П.И. Использование методов негладкой оптимизации в задачах стохастического программирования // Кибернетика и системный анализ. – 1999. – № 5. – С. 33–47.
23. Шор Н.З., СОЛОМОН Д.И. Декомпозиционные методы в дробно-линейном программировании. – Кишинев: Штиинца, 1989. – 204 с.
24. Шор Н.З., СЕРГІЄНКО І.В., ШИЛО В.П., СТЕЦЮК П.І. та інші Задачі оптимального проектування надійних мереж. – К.: Наук. думка, 2005. – 230 с.

25. ШОР Н.З. Минимизация матричных функций и недифференцируемая оптимизация // Обзорение прикладной и промышленной математики. – М.: научн. изд.-во "ТВП", 1995. – Т. 2. – С. 113–138.
26. ШОР Н.З., СТЕЦЮК П.И. Использование модификации r -алгоритма для нахождения глобального минимума полиномиальных функций // Кибернетика и системный анализ. – 1997. – № 4. – С. 28–49.
27. ШОР Н.З. Роль избыточных ограничений в улучшении двойственных оценок для полиномиальных оптимизационных задач // Кибернетика и системный анализ. – 1998. – № 4. – С. 106–121.
28. SHOR N.Z. AND STETSYUK P.I. Lagrangian Bounds in Multiextremal Polynomial and Discrete Optimization Problems // Journal of Global Optimization. – 2002. – **23**. – P. 1–41.

Ч А С Т Ь I

Первая часть сборника содержит 8 работ по методам недифференцируемой оптимизации и их приложениям в задачах блочного программирования.

Первая из работ [210] содержит обзор разработанных в Институте кибернетики методов недифференцируемой оптимизации: обобщенный градиентный спуск, методы с растяжением пространства в направлении субградиента, r -алгоритмы.

Остальные семь работ помещены в хронологическом порядке и связаны с важными этапами в развитии методов недифференцируемой оптимизации в 60–80 гг.

В работе [35] впервые было обосновано использование ОГС при декомпозиции задачи линейного программирования.

Работы [45, 46] связаны с разработкой и обоснованием алгоритмов решения двухэтапной задачи стохастического программирования. В статье [45] вводится понятие стохастического квазиградиента, послужившее основой нового направления в численных методах стохастического программирования. В статье [46] предложен и обоснован оригинальный способ регулировки шага в методах ОГС.

Две следующие работы [60, 100] связаны с важными достижениями в области субградиентных методов с растяжением пространства. В [60] предложен метод минимизации, использующий операцию растяжения пространства в направлении разности двух последовательных субградиентов (r -алгоритм). В работе [100] описывается метод эллипсоидов как частный случай алгоритма с растяжением пространства в направлении субградиента.

В двух последних работах [104, 200] описаны приложения субградиентных методов с растяжением пространства для решения блочных задач: на основе схемы декомпозиции по связывающим ограничениям [104] и по связывающим переменным [200].

Развитие алгоритмов недифференцируемой оптимизации и их приложения

Н. З. Шор, Н. Г. Журбенко, А. П. Лиховид, П. И. Стецюк
Кибернетика и системный анализ. – 2003. – № 4. – С. 80–94.

1. Введение

К задачам минимизации выпуклых функций с разрывным градиентом сводится большое число проблем, возникающих при решении сложных задач математического программирования. Владение методами недифференцируемой оптимизации дает возможность гибко использовать различные схемы декомпозиции (по переменным, ограничениям, ресурсам и т. п.), учитывающие специфику задач большой размерности, позволяет эффективно находить двойственные оценки в задачах дискретного и непрерывно- дискретного программирования и для некоторых классов многоэкстремальных задач. Появляется возможность использовать негладкие функции штрафа, позволяющие при конечных значениях штрафных параметров получать задачу безусловной минимизации, полностью эквивалентную первоначальной задаче выпуклого программирования. Техничко-экономические характеристики оптимизируемых объектов часто хорошо аппроксимируются кусочно-гладкими функциями от неизвестных параметров, что порождает задачи оптимизации с негладкими функциями.

Отсутствие эффективных методов негладкой оптимизации затрудняло решение указанных классов задач и вынуждало либо изменять постановку задачи в ущерб адекватности модели реальности, либо использовать различные приемы сглаживания. Последнее не всегда приводит к успеху, так как применение сглаживающих функций вызывает плохую обусловленность минимизируемой функции, что ухудшает вычислительную устойчивость таких эффективных методов гладкой минимизации, как квазиньютоновские и методы сопряженных градиентов.

Таким образом, область приложений методов негладкой оптимизации весьма широка, поэтому большое внимание уделяется разработке вычислительных методов негладкой оптимизации. Эти методы вполне конкурентноспособны как по надежности, так и по времени счета и точности результатов с наиболее эффективными методами решения гладких плохо обусловленных задач.

Цель данного обзора – дать краткое описание разработанных в Институте кибернетики семейств алгоритмов недифференцируемой оптимизации и показать их многочисленные приложения. В обзор включены следующие методы субградиентного типа: обобщенный градиентный спуск; методы субградиентного типа с растяжением пространства в направлении субградиента; r -алгоритмы, являющиеся мощным практическим средством решения задач недифференцируемой оптимизации.

Виктор Михайлович Глушков придавал большое значение исследованиям методов негладкой оптимизации. На самом высоком правительственном уровне он способствовал их внедрению в различных областях народного хозяйства. В частности, он уделял большое внимание созданию системы автоматизированного планирования производства и распределения заказов в черной металлургии.

2. Обобщенный градиентный спуск

Пусть $f(x)$ – выпуклая функция, определенная на евклидовом пространстве E^n , X^* – множество минимумов (оно может быть и пустым), $x^* \in X^*$ – точка минимума; $\inf f(x) = f^*$; $g_f(x)$ – субградиент (произвольный) функции $f(x)$ в точке x .

Субградиент $g_f(\bar{x})$ функции f в точке \bar{x} есть вектор $g_f(\bar{x})$ такой, что

$$f(x) - f(\bar{x}) \geq (g_f(\bar{x}), x - \bar{x}) \quad \forall \quad x \in E^n .$$

Из определения субградиента следует, что если $f(x) < f(\bar{x})$, то

$$(-g_f(\bar{x}), x - \bar{x}) > 0. \tag{1}$$

Геометрически формула (1) означает, что антисубградиент в точке \bar{x} образует острый угол с произвольным направлением, проведенным из \bar{x} в сторону точки x с меньшим значением $f(x)$. Отсюда, если X^* непусто, $\bar{x} \notin X^*$, то при движении из \bar{x} в направлении $-g_f(\bar{x})$ с достаточно малым шагом расстояние до X^* убывает. Этот простой факт лежит в основе субградиентного метода, или метода обобщенного градиентного

спуска (ОГС), впервые предложенного в [1] в связи с решением сетевой транспортной задачи.

Методом обобщенного градиентного спуска (ОГС) называется процедура построения минимизирующей последовательности $\{x_k\}_{k=0}^{\infty}$, где x_0 — начальное приближение, а x_k строятся по следующей рекуррентной формуле:

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

здесь $g_f(x_k)$ — произвольный субградиент функции $f(x)$ в точке x_k , h_{k+1} — шаговый множитель. Если $\|g_f(x_k)\| = 0$, то x_k — есть точкой минимума функции $f(x)$ и процесс останавливается.

В Институте кибернетики, начиная с 1962 года, были разработаны несколько вариантов ОГС, использующих соотношение (2). Эти результаты получены в период с 1962 по 1968 год и отражены в монографии [2]. Наиболее общий результат о сходимости ОГС содержится в следующей теореме (см., например, [2]).

Теорема 1. Пусть $f(x)$ — выпуклая функция, определенная на E^n , с ограниченной областью минимумов X^* , $\{h_k\}$ ($k = 1, 2, \dots$) — последовательность чисел, обладающая свойствами:

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

Тогда последовательность $\{x_k\}$ ($k = 1, 2, \dots$), образованная по формуле (2) при произвольном $x_0 \in E^n$ обладает одним из следующих свойств: либо найдется такое $k = \bar{k}$, что $x_{\bar{k}} \in X^*$, либо $\lim_{k \rightarrow \infty} \min_{y \in X^*} \|x_k - y\| = 0$, $\lim_{k \rightarrow \infty} f(x_k) = \min_{y \in E^n} f(x) = f^*$.

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

Теорема 2. Пусть $f(x)$ — выпуклая функция, определенная на E^n , и для всех $x \in E^n$ при некотором φ ($0 \leq \varphi < \pi/2$) выполняется неравенство

$$(g_f(x), x - x^*(x)) \geq \cos \varphi \|g_f(x)\| \cdot \|x - x^*(x)\|, \quad (3)$$

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

$$h_1 \geq \begin{cases} \|x^*(x_0) - x_0\| \cos \varphi, & \pi/4 \leq \varphi < \pi/2, \\ \|x^*(x_0) - x_0\| / (2 \cos \varphi), & 0 \leq \varphi < \pi/4, \end{cases}$$

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

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

где

$$r(\varphi) = \begin{cases} \sin \varphi, & \pi/4 \leq \varphi < \pi/2, \\ 1 / (2 \cos \varphi), & 0 \leq \varphi < \pi/4, \end{cases}$$

и вычислить $\{x_k\}_{k=1}^{\infty}$ по формуле (2), то либо при некотором k^* $g_f(x_{k^*})$ и x_{k^*} принадлежит области минимумов, либо при всех $k = 1, 2, \dots$ выполняется неравенство

$$\|x_k - x^*(x_k)\| \leq \begin{cases} h_{k+1} / \cos \varphi, & \pi/4 \leq \varphi < \pi/2, \\ 2 \cos \varphi \cdot h_{k+1}, & 0 \leq \varphi < \pi/4. \end{cases}$$

Таким образом, если угол φ заранее известен, то, регулируя шаг по формулам теоремы 2, можно получить сходимость к минимуму со скоростью геометрической прогрессии со знаменателем $q = r(\varphi)$. В формуле (3) $\cos \varphi$ характеризует степень вытянутости поверхностей уровня функции $f(x)$. Если в некоторой окрестности минимума функции $f(x)$ не существует $\varphi < \pi/2$ такого, что для любого x из этой окрестности выполняется (3), то такую функцию будем называть *существенно овражной*. При минимизации существенно овражных функций приведенный в теореме 2 способ регулировки шаговых множителей неприменим. В этом случае следует использовать универсальный способ выбора шаговых множителей, указанный в теореме 1.

Сформулируем теорему, аналогичную теореме 2, непосредственно в терминах, характеризующих степень "вытянутости" поверхностей уровня.

Теорема 3. Пусть выпуклая функция $f(x)$ определена на E^n , x^* — единственная точка минимума $f(x)$ и заданы начальное приближение

x_0 и числа σ и h_1 , причем $\sigma \geq \sqrt{2}$, $h_1 \geq \|x_0 - x^*\| / \sigma$. 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

$$\|x - x^*\| \leq \sigma \|z - x^*\|,$$

то последовательность $\{x_k\}_{k=0}^{\infty}$, образованная с помощью рекуррентных формул (2), где $h_{k+1} = h_k \sqrt{\sigma^2 - 1} / \sigma$, сходится к x^* со скоростью геометрической прогрессии:

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

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

Рассмотрим еще один вариант метода ОГС, когда шаговый множитель остается в течение определенного числа шагов постоянным, а затем уменьшается в два раза [2].

Теорема 4. Пусть для выпуклой функции $f(x)$ выполняются условия теоремы 1, $\sigma \geq 2$. Рассмотрим при заданном x итеративный процесс (2), где $h_{k+1} = h_0 \cdot 2^{-\lfloor (k+1)/N \rfloor}$. Здесь $\lfloor a \rfloor$ — целая часть числа a . При достаточно большом h_0 и $N \geq 3\sigma^2 + 1$ выполняется неравенство

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

Регулировка шага, согласно теореме 4, была использована в методе ОГС, который был предложен в 1962 году в связи с решением транспортной задачи в сетевой форме [1]. Фактически работа [1] была первым примером использования субградиентного процесса для минимизации выпуклых недифференцируемых функций.

Методы ОГС дали возможность решить большое число задач производственно-транспортного планирования с применением схем декомпозиции (по переменным и ограничениям) для задач большой размерности. Подробную информацию об этих задачах можно найти в [3]. Метод ОГС также послужил основой для создания стохастического аналога обобщенного градиентного спуска [4], который имеет большое практическое применение, в частности при решении многоэтапных задач стохастического программирования. В [3] описано применение метода обобщенного стохастического градиента к решению двухэтапной стохастической транспортной задачи, связанной с определением объемов складов однородной продукции при случайном спросе.

Результаты по методам ОГС, полученные в Институте кибернетики, получили развитие в работах [5, 6] для решения задач выпуклого программирования с ограничениями. До 1974 года работы по ОГС были малоизвестны за рубежом, так как публиковались на русском языке в малодоступных изданиях. Они получили известность после публикации в [7] подробного обзора результатов и библиографии работ по недифференцируемой оптимизации, выполненных в СССР.

3. Субградиентные методы с растяжением пространства в направлении субградиента

При анализе алгоритмов ОГС, сходящихся со скоростью геометрической прогрессии, существенную роль играли верхние границы синусов углов между направлением антиградиента в данной точке и направлением из нее на точку минимума. Медленная сходимость ОГС непреодолима в рамках этого метода в овражных задачах, когда верхняя граница указанных углов равна $\pi/2$.

Можно изменить ситуацию, используя линейные неортогональные преобразования пространства аргументов для улучшения обусловленности задачи. В случае, когда антиградиенты образуют угол, близкий к $\pi/2$, с направлением на точку минимума, разумно применить операцию растяжения пространства в направлении градиента для уменьшения его "поперечной" составляющей. Эти эвристические соображения послужили основой создания семейства методов субградиентного типа с растяжением пространства.

Операция растяжения пространства в направлении градиента первоначально введена Н.З.Шором [8]–[10] как эвристическая процедура для улучшения свойств обусловленности задачи. Она реализуется посредством оператора растяжения пространства, который в векторной форме имеет вид

$$R_\alpha(\xi) = I_n + (\alpha - 1)\xi\xi^T, \quad \xi \in E^n, \quad \|\xi\| = 1, \quad \alpha > 1,$$

где $(\cdot)^T$ означает транспонирование, $\|\cdot\|$ – евклидова норма, I_n – единичная матрица порядка n , α – коэффициент растяжения пространства, ξ – направление растяжения. Подробно свойства оператора $R_\alpha(\xi)$ изложены в монографии [2]. При описании алгоритмов используется оператор $R_\beta(\xi)$, обратный к оператору растяжения пространства $R_\alpha(\xi)$. Он имеет

следующую форму ([2]):

$$R_\beta(\xi) = R_\alpha^{-1}(\xi) = I_n + (\beta - 1)\xi\xi^T, \quad \beta = \frac{1}{\alpha} < 1.$$

Опишем принципиальную схему алгоритмов субградиентного типа с растяжением пространства в направлении субградиента для минимизации функции $f(x)$.

Заданы $x_0 \in E^n$, $B_0 = A_0^{-1} = I_n$ (единичная матрица размера $n \times n$). После k шагов имеем $x_k \in E^n$, $B_k = A_k^{-1}$, A_k – матрица размерности $n \times n$ преобразования пространства после k шагов.

1. Вычисляем $g_f(x_k)$ (если $g_f(x_k) = 0$, процесс останавливается).

2. Определяем $\tilde{g}_k = g_{\varphi_k}(y_k) = B_k^* g_f(x_k)$, где $\varphi_k(y) = f(B_k y)$; $y_k = A_k x_k$; \tilde{g}_k – обобщенный градиент для функции $\varphi_k(y)$, определенной в "растянутом" пространстве.

3. Находим

$$\xi_k = \tilde{g}_k / \|\tilde{g}_k\|; \quad x_{k+1} = x_k - h_{k+1} B_k \xi_k. \quad (4)$$

Формуле (4) соответствует движение по антисубградиенту в "растянутом" пространстве: $A_k x_{k+1} = y_k - h_{k+1} \xi_k$.

4. Вычисляем

$$B_{k+1} = A_{k+1}^{-1} = B_k R_{\beta_{k+1}}(\xi_k), \quad \beta_{k+1} = 1/\alpha_{k+1}. \quad (5)$$

Формуле (5) соответствует растяжение преобразованного пространства в направлении ξ_k : $A_{k+1} = R_{\alpha_{k+1}}(\xi_k) A_k$, $\alpha_{k+1} > 1$.

5. Переходим к следующему шагу: $k + 1 \rightarrow k + 2$.

Основная сложность при конструировании работоспособного алгоритма состояла в выборе коэффициентов растяжения пространства α_k и стратегии изменения шаговых множителей h_k . Первые эксперименты показали, что, выбирая $\alpha = 2$ и $h_k = \text{const}$, для многих примеров выпуклых овражных функций можно получить хорошие результаты [8]. К сожалению, такой простой способ не всегда приводит к цели. При построении других вариантов алгоритмов, которые удалось теоретически обосновать, шаговый множитель и коэффициенты растяжения пространства выбирались таким образом, чтобы последовательность расстояний до точки минимума в соответствующих преобразованных пространствах не возрастала. Этот принцип гарантирует сходимость со скоростью геометрической прогрессии по значению функции. Для реализации указанного принципа необходимы некоторая дополнительная информация о функции $f(x)$ – значение функции в точке минимума f^* и так называемые постоянные роста M и N .

Теорема 5. Пусть $f(x)$ – выпуклая функция, определенная на E^n и в некоторой сферической окрестности S_d точки минимума x^* : $S_d = \{x : \|x - x^*\| \leq d\}$ субградиент удовлетворяет двустороннему неравенству

$$N(f(x) - f(x^*)) \leq (g_f(x), x - x^*) \leq M(f(x) - f(x^*)), \quad (6)$$

где $M > N$ – положительные константы. Тогда, если в алгоритме принять:

- 1) $x_0 \in S_d$,
- 2) $h_{k+1} = \frac{2MN}{M+N} \frac{f(x_k) - f(x^*)}{\|g_k\|}$,
- 3) $1 < \alpha_{k+1} \leq \frac{(M+N)}{(M-N)}$, $k = 0, 1, 2, \dots$,

то для всех $k = 0, 1, \dots$ справедливо неравенство

$$\|A_k(x_k - x^*)\| \leq d. \quad (7)$$

Из неравенства (7) следует локализация x^* в эллипсоиде Φ_k с центром в точке x_k . Отношение объемов эллипсоидов Φ_{k+1} и Φ_k задается следующим равенством

$$\frac{\text{vol}(\Phi_{k+1})}{\text{vol}(\Phi_k)} = \beta_k = \frac{M - N}{M + N}.$$

Для квадратичной положительно определенной функции в неравенстве (6) можно выбирать $M = N = 2$. Для кусочно-линейной функции, надграфик которой представляет собой конус с вершиной в точке (x^*, f^*) , можно выбирать $M = N = 1$. Для этих случаев $\beta_{k+1} = \beta = 0$ и алгоритм сходится за число шагов, не превышающее n .

Решение невырожденной системы n линейных уравнений с n неизвестными $(a_i, x) + b_i = 0$, $i = 1, \dots, n$, можно заменить нахождением минимума $f(x) = \max_{1 \leq i \leq n} |(a_i, x) + b_i|$. Беря $f^* = 0$, $\beta_k = 0$ и применяя метод (4), (5), получаем алгоритм, соответствующий известной конечной процедуре решения линейных алгебраических систем - методу ортогонализации градиентов.

Получены обобщения теоремы 5 и для некоторых классов невыпуклых функций, возникающих при решении систем нелинейных уравнений $f_i(x) = 0$, $i = 1, \dots, n$. Для $f(x) = \max |f_i(x)|$ можно показать, что если x^* (решение системы) – регулярная точка (т. е. функции $f_i(x)$ непрерывно дифференцируемы в этой точке и якобиан системы $I(x^*)$

отличен от нуля), то для любого $\delta > 0$ найдется достаточно малая окрестность $S_d(x^*)$, такая, что постоянные M и N в (6) можно выбирать соответственно $M = 1 + \delta$; $N = 1 - \delta$; $\beta = \frac{M-N}{M+N} = \delta$. Как показано в работе [2], если применять предельный вариант алгоритма с $\beta = 0$ и восстановлением после каждых n итераций (большой цикл), то при обычных предположениях гладкости и регулярности для решения систем нелинейных уравнений можно получить квадратичную скорость сходимости (относительно больших циклов).

Семейство алгоритмов с растяжением пространства в направлении субградиента содержит как частный случай так называемый метод эллипсоидов. Метод эллипсоидов был предложен Д.Б.Юдиным и А.С.Немировским [11], исходя из методов последовательных отсечений, а также, независимо, Н.З.Шором [12], как частный случай алгоритма с растяжением пространства в направлении субградиента. Такой алгоритм использует следующие параметры: коэффициент растяжения пространства выбирается постоянным и равным

$$\alpha_{k+1} = \alpha = \sqrt{\frac{n+1}{n-1}},$$

а регулировка шага осуществляется по правилу:

$$h_1 = \frac{r}{n+1}; \quad h_{k+1} = h_k \frac{n}{\sqrt{n^2-1}}; \quad k = 1, 2, \dots,$$

где n – размерность пространства, r – радиус шара с центром в точке x_0 , содержащий точку x^* .

Метод эллипсоидов сходится со скоростью геометрической прогрессии по отклонению наилучшего достигнутого на данном шаге значения $f(x)$ от оптимального, при этом знаменатель геометрической прогрессии зависит только от размерности пространства n асимптотически

$$q_n \approx 1 - \frac{1}{2n^2}.$$

Метод эллипсоидов использован Л.Г.Хачияном [13] для построения и обоснования первого полиномиального алгоритма для решения задачи линейного программирования с рациональными коэффициентами. Кроме того, этот метод эллипсоидов получил важное применение в теории сложности алгоритмов дискретной оптимизации [14]. На конгрессе по математическому программированию, который проходил в Бонне в 1982 году, методу эллипсоидов и его приложениям было уделено большое внимание. В частности, в [15] был опубликован обзорный доклад

Н.З.Шора по методам негладкой оптимизации, разработанным в Институте кибернетики [16].

Опыт применения алгоритмов с растяжением пространства в направлении градиента показал возможность существенного ускорения субградиентных процессов при использовании операторов, изменяющих метрику пространства. В то же время трудности выбора шаговых множителей стимулировали поиск новых методов негладкой оптимизации с переменной метрикой, в которых выбор шагового множителя связан с поиском минимума по направлению. Этот класс алгоритмов описан в следующем разделе.

4. Субградиентные методы с растяжением пространства в направлении разности двух последовательных субградиентов

При решении сложных задач недифференцируемой оптимизации средней размерности (до нескольких сот переменных) особенно эффективны оказались алгоритмы субградиентного типа с растяжением пространства в направлении разности двух последовательных субградиентов (r -алгоритмы), предложенные в 1971 году в работе [17].

По своей структуре и трудоемкости итерации r -алгоритмы близки к методам с растяжением пространства в направлении субградиента. Но между ними есть важное различие: методы ОГС с растяжением пространства в направлении субградиента в принципе не могут быть монотонными, в то же время r -алгоритмы при определенной регуировке шаговых множителей и коэффициентов растяжения пространства могут стать монотонными. Это обстоятельство связано с простым геометрическим фактом: если точка x на границе двух "кусков" кусочно-гладкой поверхности уровня, причем градиенты к этим гладким "кускам", вычисленные в данной точке, образуют тупой угол, то любое растяжение пространства в направлении одного из градиентов (или последовательное растяжение попеременно в направлении двух указанных градиентов) не может превратить этот угол в острый, он может лишь приближаться к $\pi/2$, оставаясь тупым. Применяя метод с растяжением пространства в направлении субградиента, нельзя получить направление убывания функции в виде антиградиента к одному из кусков в растянутом пространстве. В то же время растяжение пространства в направлении разности двух указанных градиентов с достаточным ко-

эффицентом растяжения превращает тупой угол между градиентами в острый, т.е. соответствующие образы этих антиградиентов в растянутом пространстве становятся направлениями убывания функции.

Приведем общую схему r -алгоритмов для минимизации выпуклой функции $f(x)$, определенной на E^n . Будем предполагать, что $f(x)$ имеет ограниченную область минимумов X^* , так что $\lim_{\|x\| \rightarrow \infty} f(x) = +\infty$,

Выбираем начальное приближение $x_0 \in E^n$ и неособенную матрицу B_0 (чаще всего B_0 совпадает с единичной матрицей I_n или с диагональной матрицей D_n с положительными элементами на диагонали, с помощью которой осуществляется масштабирование переменных).

Первый шаг алгоритма производим по формуле $x_1 = x_0 - h_0 \eta_0$, где $\eta_0 = B_0 B_0^T g_f(x_0)$, h_0 – некоторый шаговый множитель, выбираемый из условия существования в точке x_1 субградиента $g_f(x_1)$, такого, что $(g_f(x_1), \eta_0) \leq 0$. При $B_0 = I_n$ имеем $\eta_0 = g_f(x_0)$ и первый шаг совпадает с итерацией субградиентного процесса. Пусть в результате вычислений после k ($k = 1, 2, \dots$) шагов процесса получены определенные значения $x_k \in E^n$ и матрицы B_k размерности $n \times n$. Опишем $(k + 1)$ -й шаг процесса.

1. Вычисляем следующие величины: $g_f(x_k)$ – субградиент функции $f(x)$ в точке x_k ; $r_k = B_k^T (g_f(x_k) - g_f(x_{k-1}))$ – вектор разности двух последовательных субградиентов в преобразованном пространстве.

Переход от начального пространства к преобразованному задается формулой $y = A_k x$, где $A_k = B_k^{-1}$. Определим функцию $\varphi_k(y) = f(B_k y)$, тогда $g_{\varphi_k}(y) = B_k^T g_f(x)$. Таким образом, r_k есть разность двух субградиентов от функции $\varphi_k(y)$, вычисленных в точках $y_k = A_k x_k$ и $\tilde{y}_k = A_k x_{k-1}$.

2. Определяем $\xi_k = r_k / \|r_k\|$.

3. Задаем величину β_k , обратную коэффициенту растяжения пространства α_k перед $(k + 1)$ -м шагом.

4. Вычисляем $B_{k+1} = B_k R_{\beta_k}(\xi_k)$, где $R_{\beta_k}(\xi_k)$ – оператор растяжения пространства на $(k + 1)$ -м шаге. Заметим, что $B_{k+1} = A_{k+1}^{-1}$.

5. Находим $\tilde{g}_k = B_{k+1}^T g_f(x_k)$, субградиент функции $\varphi_{k+1} = f(B_{k+1} y)$ в точке $y_{k+1} = A_{k+1} x_k$.

6. Определяем

$$x_{k+1} = x_k - h_k B_{k+1} \tilde{g}_k / \|\tilde{g}_k\|. \quad (8)$$

Шаг алгоритма (8) соответствует шагу обобщенного градиентного спуска в преобразованном под воздействием оператора A_{k+1} пространстве. Действительно, применив к обеим частям формулы (8) оператор A_{k+1}

получим

$$y_{k+1} = A_{k+1}x_{k+1} = \bar{y}_k - h_k \tilde{g}_k / \|\tilde{g}_k\|, \quad (9)$$

где $\bar{y}_k = A_{k+1}x_k$.

7. Переходим к следующему шагу или заканчиваем работу алгоритма при выполнении некоторых условий останова.

Практическая эффективность алгоритма во многом зависит от выбора шагового множителя h_k . В r -алгоритме h_k выбирается из условия приближенного поиска минимума $f(x)$ по направлению, при этом при минимизации выпуклых функций должно соблюдаться условие $h_k \geq h_k^*$, (h_k^* – значение шагового множителя, соответствующего минимуму по направлению). В общем случае необходимо, чтобы направление субградиента в точке x_{k+1} образовывало нетупой угол с направлением спуска из точки x_k .

При минимизации негладких выпуклых функций, определенных на E^n , наиболее удачными оказались следующие варианты алгоритма при проведении экспериментальных и практических расчетов. Коэффициенты растяжения пространства α_k выбираются в пределах 2-3, для шагового множителя h_k применяется адаптивный способ регулировки. задается некоторое натуральное число m , постоянные $q > 1$ и $t_k^0 > 0$. После k шагов получаем постоянную t_k^0 . Двигаемся из точки x_k в направлении спуска с шагом t_k^0 до тех пор, пока не будет выполнено условие завершения спуска по направлению, либо число шагов не станет равным m . Условие завершения спуска может состоять в том, что значение функции в очередной точке не меньше, чем значение функции в предыдущей точке; другой вариант такого условия – производная по направлению спуска в данной точке неотрицательна. Если прошло m шагов, а условие завершения спуска не выполнено, то вместо t_k^0 запоминаем $t_k^1 = qt_k^0$, где $q > 1$, и продолжаем спуск в том же направлении с большим шагом. Если после очередных m шагов условие завершения спуска не выполнено, то вместо t_k^1 берем $t_k^2 = qt_k^1$ и т. д. Так как предполагаем, что $\lim_{\|x\| \rightarrow \infty} f(x) = +\infty$, то после конечного числа шагов в определенном направлении обязательно выполнится условие завершения спуска. Постоянная шага $t_k^{p_k} = q^{p_k} t_k^0$ ($p \in \{0, 1, 2, \dots\}$), которая использовалась на последнем шаге, принимается в качестве начальной при спуске в новом направлении из точки x_{k+1} , т.е. $t_{k+1}^0 = t_k^{p_k}$.

Приведенный выше способ регулировки шагового множителя основан на следующих соображениях. Допустим, что мы применили r -алгоритм с постоянным коэффициентом растяжения пространства α . Тогда за n итераций произвольное направление растянется в "среднем" в α

раз, т.е. если идти в растянутом пространстве с постоянным шагом, то в первоначальном пространстве за n итераций шаг уменьшится примерно в α раз. Если расстояние до точки минимума при этом будет уменьшаться с такой же средней скоростью, т.е. примерно в α раз за n итераций, то число шагов по направлению будет ограниченным. При более медленном темпе сходимости шаг в первоначальном пространстве "в среднем" будет уменьшаться быстрее, чем расстояние до точки минимума и число шагов по направлению может неограниченно возрастать. Для предотвращения этого вводится адаптивная процедура увеличения шагового множителя h , стабилизирующая число шагов по направлению от одного до нескольких единиц. Данная процедура позволяет быстрее выявить направления, в которых функция неограниченно убывает (такие случаи часто встречаются на практике, когда задачу линейного программирования решают в пространстве двойственных переменных с использованием схемы декомпозиции по ограничениям, а задача оказывается несовместной). В этом случае шаговый множитель резко возрастает, что служит сигналом к остановке процесса спуска.

Как показали многочисленные вычислительные эксперименты и практические расчеты, в большинстве случаев при $\alpha \in [2, 3]$ и $m = 3$ и указанном выше способе регулировки h число шагов по направлению в среднем редко превосходит 2, при этом за n шагов r -алгоритма точность по функционалу, как правило, улучшается в 3-5 раз.

В случае минимизации гладкой функции для ускорения сходимости можно применять более тонкие способы поиска минимума по направлению, например, квадратичную аппроксимацию по трем точкам, процесс "золотого сечения" и др. В гладком случае хорошо зарекомендовал себя адаптивный способ регулировки шага по направлению, подобный приведенному выше, с небольшим изменением: если на данной итерации функция приняла уже после первого шага большее значение, то шаговый множитель умножается на заданное число, меньшее единицы (порядка 0,8-0,95). Это связано с тем, что в гладком случае скорость сходимости может оказаться более быстрой при более точном нахождении минимума по направлению, а дополнительное измельчение шага способствует увеличению точности поиска минимума по направлению.

Применительно к задачам гладкой оптимизации r -алгоритм по своей формальной структуре близок к алгоритмам квазиньютоновского типа с переменной метрикой. Так предельный вариант r -алгоритма с бесконечным коэффициентом растяжения (т.е. $\beta = 0$) является проективным вариантом метода сопряженных градиентов. В работе [17] показано, что предельный вариант r -алгоритма обладает квадратич-

ной скоростью сходимости при обычных условиях гладкости и регулярности. Заметим, что на основе использования операторов растяжения пространства можно строить семейство квазиньютоновских методов и при конечных значениях коэффициентов растяжения [18, 19]. Эти алгоритмы характеризуются численной устойчивостью по отношению к точности поиска минимума по направлению.

Далее рассмотрим монотонную модификацию r -алгоритма [20]. Данная модификация отличается от других вариантов r -алгоритмов как выбором направления движения, так и регулировкой величины шагового множителя. В общем случае монотонная модификация r -алгоритма состоит из последовательности итераций, на каждой из которых производятся следующие действия (ниже $x_0, x_1, \dots, x_k, \dots$ минимизирующая последовательность точек пространства E^n , на котором определена функция $f(x)$). Пусть задана некоторая начальная точка $x_0 \in E^n$, неособая матрица B_0 размерности $n \times n$. Начальная итерация соответствует первому шагу приведенной выше схемы r -алгоритма.

Пусть проделано \bar{k} итераций ($\bar{k} > 1$), в результате которых получена точка $x_{\bar{k}} \in E^n$, обратная матрица $B_{\bar{k}}$ преобразования пространства и субградиент $g(x_{\bar{k}})$, так что направление движения из точки $x_{\bar{k}}$ определяется по формуле:

$$x_{\bar{k}+1}(h_{\bar{k}}) = x_{\bar{k}} - h_{\bar{k}} B_{\bar{k}} B_{\bar{k}}^T g_{\bar{k}}$$

$h_{\bar{k}}$ – неизвестный шаговый множитель, который определяется из условия минимума по направлению:

$$h_{\bar{k}}^{\min} = \arg \min f(x_{\bar{k}}(h_{\bar{k}}))$$

В частности, $h_{\bar{k}}^{\min}$ может равняться 0, а в этом случае $x_{\bar{k}+1} = x_{\bar{k}}$.

Шаговый множитель при движении в основном пространстве выбирается из условия приближенного минимума по направлению движения, но в отличие от известных модификаций в определенных обстоятельствах может оказаться, что шаговый множитель равен 0. В этом случае остаемся в текущей точке и на соответствующей итерации изменяется только обратная матрица.

Если $h_{\bar{k}}^{\min} > 0$, то

$$x_{\bar{k}+1} = x_{\bar{k}} - h_{\bar{k}}^{\min} B_{\bar{k}} B_{\bar{k}}^T g_{\bar{k}}.$$

В обоих случаях среди субградиентов функции f в точке $x_{\bar{k}+1}$ находим субградиент, который образует с направлением движения угол,

больший или равный $\frac{\pi}{2}$. Если искомый субградиент можно выбрать неоднозначно, то выбирается тот, проекция которого на направление спуска минимальна.

Последовательность $f(x_0), f(x_1), \dots, f(x_k)$ является невозрастающей, так как следующая точка получается из предыдущей в результате операции нахождения приближенного минимума по некоторому направлению, применение которой обеспечивает выполнение неравенства $f(x_k) \geq f(x_{k+1})$. Отсюда ясно, что последовательность значений функции $\{f(x_k)\}_{k=0}^{\infty}$ имеет предел $f^* \geq f(x^*)$.

В алгоритм вводится новая возможность выбора направления движения, связанная с решением вспомогательных задач выпуклого квадратичного программирования. Рассмотрим случай минимизации функции максимума гладких выпуклых функций. Отметим, что гладкая выпуклая функция аппроксимируется снизу в любой точке \bar{x} линейной функцией вида $(g_f(\bar{x}), x - \bar{x})$.

Пусть $\varphi(x) = \max_{i \in I} f_i(x)$, где I – конечное множество индексов, $f_i(x)$ – непрерывно дифференцируемые выпуклые функции, определенные на всем E^n . Для упрощения выкладок будем считать, что $\varphi(x)$ имеет единственный минимум – искомую точку $x^* \in E^n$.

Пусть \bar{x} – произвольная точка из E^n . Обозначим $I_0(\bar{x}) = \{i \in I \mid \varphi(\bar{x}) = f_i(\bar{x})\}$, $I_\varepsilon(\bar{x}) = \{i \in I \mid \varphi(\bar{x}) \geq f_i(\bar{x}) \geq \varphi(\bar{x}) - \varepsilon\}$. $I_\varepsilon(\bar{x})$ будем называть ε -активным множеством индексов для точки \bar{x} .

По сравнению с обычным г-алгоритмом при выборе направления спуска используется следующая специальная квадратичная задача: найти

$$\rho^* = \min_{\{\alpha: \{\alpha_i\}_{i \in I_\varepsilon(\bar{x})}\}} \sum_{p,l} \alpha_p \alpha_l (g_{f_p}(\bar{x}), g_{f_l}(\bar{x})) \quad (10)$$

при ограничениях

$$\alpha_i \geq 0 \text{ для всех } i \in I_\varepsilon(\bar{x}) \quad (11)$$

$$\sum_{i \in I_\varepsilon(\bar{x})} \alpha_i = 1 \quad (12)$$

Если $\rho^* = 0$, то $\varphi(x^*) \geq \varphi(\bar{x}) - \varepsilon$, т.е. значение $\varphi(\bar{x}) - \varepsilon$ является нижней оценкой. Имея значения рекорда и нижней оценки минимизируемой функции в точке \bar{x} , можно оценить точность определения минимума в текущей точке. Если $\rho^* > 0$, то вектор, противоположный вектору $\sum_{i \in I_\varepsilon(\bar{x})} \alpha_i^* g_{f_i}(\bar{x})$, дает направление движения из точки \bar{x} , уменьшающее значение $\varphi(\bar{x})$ не менее, чем на ε .

Отметим, что вспомогательная задача (10)–(12) является относительно λ задачей квадратичного выпуклого программирования с линейными ограничениями, и она легко сводится с использованием негладких штрафных функций к задаче безусловной минимизации негладкой кусочно-квадратичной функции. Для решения этой задачи можно применить r -алгоритм в сочетании с методом негладких штрафных функций. Вычислительные эксперименты показали для больших примеров сокращение времени получения решения для вспомогательной задачи по сравнению с известной программой LOQO, которая реализует подход, основанный на методе внутренних точек.

Таким образом, реализация монотонной модификации r -алгоритма в общем случае состоит из конечного числа стадий, каждая из которых состоит из нескольких блоков:

(а) вычисление последовательных значений $x_{\bar{k}}, x_{\bar{k}+1}, \dots$ с невозрастающими значениями минимизируемой функции при постоянном $\alpha > 1$ и соответствующих матриц $B_{\bar{k}}, B_{\bar{k}+1}$;

(б) контроль вырожденности матриц $B_{\bar{k}}$ и их "восстановление" в случае необходимости;

(с) решение специальных вспомогательных квадратичных задач для получения оценок снизу минимизируемой функции.

Монотонная модификация r -алгоритма применялась для решения широкого класса оптимизационных задач, включая минимаксные задачи, специальные квадратичные задачи, задачи максимального разреза графа, задачи оптимального управления с дискретным временем, полиномиальные многоэкстремальные задачи [20].

5. Приложения алгоритмов недифференцируемой оптимизации

Приведем основные источники, порождающие задачи негладкой оптимизации.

Во-первых, это задачи математического программирования большой 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. Пусть задано параметрическое семейство выпуклых функций, определенных на E^n , $\{f_\alpha(x)\}_{\alpha \in A}$. Основным источником получения негладких функций в выпуклом программировании является операция взятия поточечного максимума по параметру α , т.е. построение функции максимума:

$$F(x) = \sup_{\alpha \in A} f_\alpha(x).$$

Область определения функции $F(x)$ ($\text{dom } F$) совпадает с такими значениями $x \in E^n$, при которых $\{f_\alpha(x)\}$ ограничена сверху по α . Для каждого $\bar{x} \in \text{dom } F$ определим подмножество индексов

$$I(\bar{x}) = \{\alpha \in A : f_\alpha(\bar{x}) = F(\bar{x})\}$$

Субградиентное множество $G_F(\bar{x})$ функции F в точке \bar{x} определяется формулой

$$G_F(\bar{x}) = \overline{\text{conv}} \{ \cup_{\alpha \in I(\bar{x})} G_{f_\alpha}(\bar{x}) \}, \quad (13)$$

где $\overline{\text{conv}} \{M\}$ обозначает операцию нахождения минимального выпуклого замкнутого множества, содержащего M , $G_{f_\alpha}(\bar{x})$ – субградиентные подмножества функций f_α в точке \bar{x} , $\alpha \in I(\bar{x})$. Если все функции f_α ($\alpha \in I(\bar{x})$) являются дифференцируемыми в точке \bar{x} , то $G_{f_\alpha}(\bar{x})$ состоит из единственной точки, совпадающей с градиентом $g_{f_\alpha}(\bar{x})$, и формула принимает следующий вид:

$$G_F(\bar{x}) = \overline{\text{conv}} \{ \cup_{\alpha \in I(\bar{x})} g_{f_\alpha}(\bar{x}) \}.$$

В случае, когда $I(\bar{x})$ – конечное множество, то все крайние точки множества $G_F(\bar{x})$ являются градиентами некоторых функций f_α , $\alpha \in I(\bar{x})$, в точке \bar{x} , и $G_F(\bar{x})$ представляет собой выпуклый многогранник соответствующей размерности.

Третий источник негладких задач – лагранжевы оценки в задачах математического программирования. Рассмотрим достаточно общую задачу математического программирования, ограничения которой разбиты на две части, часть из которых имеет вид условия принадлежности $x \in X \subseteq E^n$, а другая часть определяется системой равенств: найти

$$f_0^* = \inf_{x \in X} f_0(x), \quad X \subseteq E^n, \quad (14)$$

при ограничениях

$$f_i(x) = 0, \quad i = 1, \dots, m. \quad (15)$$

Предположим, что X – замкнутое подмножество n -мерного евклидова пространства, $f_i(x)$ – непрерывные функции, определенные на X .

Для оценки f_0^* введем функцию Лагранжа

$$L(x, u) = f_0(x) + \sum_{i=1}^m u_i f_i(x),$$

где $u = \{u_1, \dots, u_m\}$ – вектор множителей Лагранжа. Рассмотрим оценку

$$\psi(u) = \inf_x L(x, u).$$

При любом допустимом x и произвольном u имеем $\psi(u) \leq f_0(x)$, откуда следует что $\psi(u) \leq f_0^*$.

Задача нахождения наилучшей оценки для оптимального значения задачи (14), (15) в данном классе лагранжевых оценок сводится к решению координирующей задачи: найти

$$\psi^* = \sup_u \psi(u).$$

Функция $\psi(u)$ является вогнутой как результат операции минимизации по $x \in X$ параметрического (u -параметр) семейства линейных по u функций. Предположим, что $\psi(u)$ – собственная вогнутая функция с непустой областью $\text{dom } \psi$, имеющей внутренние точки. Пусть \bar{x} – некоторая внутренняя точка $\text{dom } \psi$, т.е. $\bar{x} \in \text{int } \text{dom } \psi$.

Тогда по правилам вычисления субградиента от функции максимума $\varphi(u) = -\psi(u)$ субградиентное множество $G_\varphi(\bar{x})$ определяется следующим образом:

$$G_\varphi(\bar{x}) = \overline{\text{conv}} \left\{ \bigcup_{x \in X(\bar{x})} F(x(u)) \right\},$$

где $X(u)$ – множество всевозможных решений локальной задачи $\inf_{x \in X} L(x, \bar{u})$; $F(x(\bar{u}))$ – вектор "невязок", соответствующий решению $x(\bar{u})$, $F(x(\bar{u})) = \{f_1(x(\bar{u})), \dots, f_m(x(\bar{u}))\}$. Таким образом, если в точке \bar{x} локальная задача имеет единственное решение, то φ в соответствующей точке дифференцируема и ее градиент совпадает с вектором невязок $\{-f_i(x(\bar{u}))\}_{i=1}^m$. В противном случае, как правило, градиент функции $\varphi(u)$ в соответствующей точке терпит разрыв.

Четвертый источник – 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).

Пятый источник составляют 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 $+\infty$.

Шестой источник – 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 рассматривать как специальные задачи нелинейного программирования, для решения которых применимы схемы декомпозиции или метод негладких штрафных функций.

Седьмой источник составляют 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.

Следует 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 вычислительные методы, разработанные для решения задач негладкой оптимизации, оказываются эффективными и для оптимизации "плохих" гладких функций (например, функций овражного типа). Многочисленные приложения алгоритмов негладкой оптимизации для указанных классов задач можно найти в монографиях Н. З. Шора [2, 3, 21, 22].

6. Заключение

Работа содержит краткое описание разработанных в Институте кибернетики методов недифференцируемой оптимизации.

1. Методы обобщенного градиентного спуска, положившие начало новому направлению математического программирования – численным методам негладкой оптимизации, которому в настоящее время посвящены многочисленные научные статьи и монографии.

2. Субградиентные методы с растяжением пространства в направлении субградиента, которые имеют ускоренную сходимость в сравнении с методами обобщенного градиентного спуска. Эти методы дали теории оптимизации уникальный алгоритм – метод эллипсоидов, скорость сходимости которого зависит лишь от размерности пространства. Использование метода эллипсоидов позволило решить ряд важных вопросов в теории сложности задач математического программирования.

3. Субградиентные методы с растяжением пространства в направлении разности двух последовательных субградиентов – r -алгоритмы. В рамках этого семейства методов получены достаточно эффективные реализации r -алгоритмов. Число итераций для нахождения оптимального значения f^* с ε -точностью для функций от n переменных эмпирически оценивается как $N = O(n \log \frac{1}{\varepsilon})$. Разработанные модификации r -алгоритма являются эффективным средством минимизации выпуклых негладких функций. При минимизации гладких функций они оказались конкурентноспособными с наиболее удачными реализациями методов сопряженных направлений и методов квазиньютоновского типа.

r -алгоритм использовался в задачах оптимизации большой размерности и в квази-блочных задачах с различными схемами декомпозиции, для вычисления двойственных лагранжевых оценок в многоэкстремальных и комбинаторных задачах оптимизации. На практике он применялся для решения задач оптимального планирования, оптимального проектирования, синтеза сетей, восстановления изображений, эллипсоидальной аппроксимации и локализации, и др.

В.М.Глушков, под руководством которого сложилась школа негладкой оптимизации в Институте кибернетики, придавал большое значение развитию численных методов оптимизации, отмечая важность их практических приложений. В настоящее время развитие методов недифференцируемой оптимизации активно продолжается.

Литература

1. ШОР Н.З. Применение метода градиентного спуска для решения сетевой транспортной задачи. // Материалы науч. семинара по теорет. и прикл. вопр. кибернетики и исследования операций. – Киев: Науч. совет по кибернетике АН УССР, 1962. – Вып. 1. – С. 9–17.
2. ШОР Н.З. Методы минимизации недифференцируемых функций и их приложения. – Киев: Наук. думка, 1979. – 199 с.
3. МИХАЛЕВИЧ В.С., ТРУБИН В.А., ШОР Н.З. Оптимизационные задачи производственно-транспортного планирования. Модели, методы, алгоритмы. – М.: Наука, 1986. – 260 с.
4. ЕРМОЛЬЕВ Ю.М., ШОР Н.З. Метод случайного поиска для задач двухэтапного стохастического программирования и его обобщение. Кибернетика. – 1968. – № 1. – С. 90–92.
5. ЕРЕМИН И.И. О методе "штрафов" в выпуклом программировании. – Кибернетика. – 1966. – № 4. – С. 63–67.
6. ПОЛЯК Б.Т. Один общий метод решения экстремальных задач. – Докл. АН СССР. – 1967. – 174. – № 1. – С. 33–36.
7. BALINSKI M.L., WOLFE P. (EDS). Nondifferentiable optimization // Math. Programming Study, 3. – Amsterdam: North-Holland, 1975. – 178 p.
8. ШОР Н.З., БИЛЕЦКИЙ В.И. Метод растяжения пространства для ускорения сходимости в задачах овражного типа. – Тр. семинара .Науч. совета АН УССР по кибернетике "Теория оптимальных решений". – Киев, 1969. – № 2. – С. 3–18.
9. ШОР Н.З. Использование операций растяжения пространства в задачах минимизации выпуклых функций // Кибернетика. – 1970. – № 1. – С. 6–12.
10. ШОР Н.З. О скорости сходимости обобщенного градиентного спуска с растяжением пространства // Кибернетика. – 1970. – № 2. – С. 80–85.
11. ЮДИН Д.Б., НЕМИРОВСКИЙ А.С. Информационная сложность и эффективные методы решения выпуклых экстремальных задач // Экономика и мат. методы. – 1976. – Вып. 2. – С. 357–359.

12. ШОР Н.З. Метод отсечения с растяжением пространства для решения задач выпуклого программирования // Кибернетика. – 1977. – № 1. – С. 94–95.
13. ХАЧИЯН Л.Г. Полиномиальные алгоритмы в линейном программировании // ДАН СССР. – 1979. – **244**. – № 5. – С. 1093–1096.
14. GROTSCHEL M., LOVASZ L., SCHRIJVER A. Geometric algorithms and combinatorial optimization. – Berlin: Springer-Verlag, 1988. – 362 p.
15. BACHEM A., GROTSCHEL M., KORTE B. (EDS). Mathematical programming the state of art. – Berlin: Springer-Verlag, 1983. – 655 p.
16. SHOR N.Z. Generalized gradient methods of nondifferentiable optimization employing space dilatation operations. – Berlin: Springer-Verlag, 1982. – P. 501–529.
17. ШОР Н.З., ЖУРБЕНКО Н.Г. Метод минимизации, использующий операцию растяжения пространства в направлении разности двух последовательных градиентов // Кибернетика. – 1971. – № 3. – С. 51–59.
18. ЖУРБЕНКО Н.Г. Построение семейства методов сопряженных направлений на основе использования оператора растяжения пространства // Теория и приложения методов оптимизации. – Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 1998. – С. 12–18.
19. ЖУРБЕНКО Н.Г. Квазиньютоновские алгоритмы минимизации на основе использования оператора растяжения пространства // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 1999. – С. 45–50.
20. ШОР Н.З. Монотонные модификации r -алгоритмов и их приложения // Кибернетика и систем. анализ. – 2002. – № 6. – С. 74–95.
21. ШОР Н.З., СТЕЦЕНКО С.И. Квадратичные экстремальные задачи и недифференцируемая оптимизация. – Киев: Наук. думка, 1989. – 208 с.
22. SHOR, N.Z. Nondifferentiable optimization and polynomial problems. Boston; Dordrecht; London: Kluwer Academic Publishers, – 1998. – 412 p.

Применение обобщенного градиентного спуска в блочном программировании

Н. З. Шор

Кибернетика. — 1967. — № 3. — С. 53–55.

1. Метод обобщенного градиентного спуска (о. г. с.) был развит для минимизации выпуклых (максимизации вогнутых) недифференцируемых функций [1-3]. Этот метод отличается от обычного градиентного спуска формально тем, что в точках, где минимизируемая функция недифференцируема (в этих точках направление градиента не определено), направление градиента заменяется направлением нормали к произвольной опорной гиперплоскости поверхности уровня в этой точке. В отличие от методов возможных направлений при использовании о. г. с. не всегда можно добиться, чтобы последовательность значений минимизируемой функции, получаемая при реализации спуска, была монотонно убывающей. В связи с этим при практическом использовании о. г. с. нужно применять специальную методику регулировки шага [1,2].

Впервые метод обобщенного градиентного спуска был применен в 1962 г. для решения транспортной задачи линейного программирования в сетевой форме [2]. Обоснование метода о. г. с. для произвольной выпуклой функции дано в диссертации автора [1], а также в обзорной статье [3].

Доклад Б. Т. Поляка на V Международном математическом Конгрессе в Москве был посвящен некоторым обобщениям этого метода.

В данной статье рассмотрим возможность применения о. г. с. в блочном линейном программировании и, в частности, для решения двухэтапных задач стохастического программирования [4,6].

2. Запишем общую задачу линейного программирования в форме, когда система ограничений некоторым образом разбита на две части. Найти

$$\max L(X) = \max (C, X) = \max \sum_{j=1}^n c_j x_j; \quad C \in E^n; \quad X \in E^n; \quad (1)$$

$$C = \{c_j\}; \quad X = \{x_j\}; \quad j = 1, \dots, n$$

при ограничениях:

$$x_j \geq 0; \quad j = 1, \dots, n; \quad (2)$$

$$\sum_{j=1}^n a_{ij}^{(1)} x_j \leq b_i^{(1)}; \quad i = 1, \dots, m_1; \quad (3)$$

$$\sum_{j=1}^n a_{kj}^{(2)} x_j \leq b_k^{(2)}, \quad k = 1, \dots, m_2; \quad (4)$$

Будем считать, что:

- 1) задача (1)–(4) имеет оптимальное решение;
- 2) система ограничений (2), (4) вырезает в пространстве ограниченный многогранник D (если это не так, то мы всегда можем к системе неравенств (4) добавить неравенство вида $\sum_{j=1}^n x_j \leq M$ и M подобрать настолько большим, чтобы хотя бы одно из оптимальных решений задачи (1)–(4) удовлетворяло этому неравенству).

Составим для задачи (1), (3) функцию Лагранжа

$$L\{\Lambda, X\}; \quad \Lambda \in E^{m_1}; \quad \Lambda = \{\lambda_i\}; \quad i = 1, \dots, m_1.$$

Рассмотрим задачу. Найти

$$\begin{aligned} \min_{\{\Lambda \geq 0\}} \max_{\{X \in D\}} L(\Lambda, X) &= \\ &= \min_{\{\Lambda \geq 0\}} \max_{\{X \in D\}} \left[\sum_{j=1}^n c_j x_j + \sum_{i=1}^{m_1} \lambda_i \left(b_i^{(1)} - \sum_{j=1}^n a_{ij}^{(1)} x_j \right) \right]. \quad (5) \end{aligned}$$

Изучим функцию:

$$L^* \{\Lambda\} = \max_{\{X \in D\}} \left[\sum_{j=1}^n c_j x_j + \sum_{i=1}^{m_1} \lambda_i \left(b_i^{(1)} - \sum_{j=1}^n a_{ij}^{(1)} x_j \right) \right]. \quad (6)$$

Так как многогранник D ограничен, то функция $L^* \{\Lambda\}$ определена для любого Λ . Как результат применения операции максимума к выпуклому

множеству линейных форм, функция $L^* \{ \Lambda \}$ является выпуклой. Так как D – многогранник, то $L^* \{ \Lambda \}$ – кусочно-линейная функция.

Для $\Lambda \geq 0$ и для $X = \bar{X}$, где $\bar{X} = \{ \bar{x}_1, \dots, \bar{x}_n \}$ – допустимое решение задачи (1)–(4),

$$L^* \{ \Lambda \} \geq L \{ \Lambda, \bar{X} \} = \sum_{j=1}^n c_j \bar{x}_j + \sum_{i=1}^{m_1} \lambda_i \left(b_i^{(1)} - \sum_{j=1}^n a_{ij}^{(1)} \bar{x}_j \right) \geq \sum_{j=1}^n c_j \bar{x}_j. \quad (7)$$

Таким образом, $L^* \{ \Lambda \}$ ограничена снизу для $\Lambda \geq 0$, а так как $L^* \{ \Lambda \}$ – кусочно-линейная, то существует Λ^* такое, что

$$L^* \{ \Lambda^* \} = \min_{\{ \Lambda \geq 0 \}} L^* \{ \Lambda \}. \quad (8)$$

Рассмотрим задачу, двойственную к задаче (1)–(4).

Найти

$$\min \left(\sum_{i=1}^{m_1} b_i^{(1)} \lambda_i + \sum_{k=1}^{m_2} b_k^{(2)} \sigma_k \right) \quad (9)$$

при ограничениях:

$$\lambda_i \geq 0; \quad \sigma_k \geq 0; \quad i = 1, \dots, m_1; \quad k = 1, \dots, m_2; \quad (10)$$

$$\sum_{i=1}^{m_1} a_{ij}^{(1)} \lambda_i + \sum_{k=1}^{m_2} a_{kj}^{(2)} \sigma_k \geq c_j; \quad j = 1, \dots, n. \quad (11)$$

Лемма. Любому оптимальному решению задачи (5) $\Lambda^* = \{ \lambda_1^*, \dots, \lambda_{m_1}^* \}$ соответствует оптимальное решение задачи (9)–(11) вида

$$(\Lambda^*; \Sigma^*) = \{ \lambda_1^*, \dots, \lambda_{m_1}^*; \sigma_1^*, \dots, \sigma_{m_2}^* \}.$$

Доказательство. Рассмотрим задачу нахождения $L^* \{ \Lambda \}$ (6). Это – задача линейного программирования. Из теоремы двойственности узнаем, что найдутся такие $\sigma_k \geq 0$, что

$$\sum_{k=1}^{m_2} \sigma_k a_{kj}^{(2)} \geq c_j - \sum_{i=1}^{m_1} \lambda_i a_{ij}^{(1)} \quad (j = 1, \dots, n) \quad (12)$$

или

$$\sum_{i=1}^{m_1} \lambda_i a_{ij}^{(1)} + \sum_{k=1}^{m_2} \bar{\sigma}_k a_{kj}^{(2)} \geq c_j \quad (13)$$

Кроме того,

$$L^* \{ \Lambda \} = \sum_{i=1}^{m_1} \lambda_i b_i^{(1)} + \sum_{k=1}^{m_2} \bar{\sigma}_k b_k^{(2)}. \quad (14)$$

Таким образом, для любого $\Lambda \geq 0$ найдется $\bar{\Sigma} = \{ \bar{\sigma}_k \} \geq 0$, что $(\Lambda, \bar{\Sigma})$ образует допустимое решение задачи (9)–(11).

Далее,

$$\begin{aligned} L \{ \Lambda, X \} &= \sum_{j=1}^n x_j \left(c_j - \sum_{i=1}^{m_1} \lambda_i a_{ij}^{(1)} \right) + \sum_{i=1}^{m_1} \lambda_i b_i^{(1)} \leq \\ &\leq \sum_{j=1}^n x_j \sum_{k=1}^{m_2} \bar{\sigma}_k a_{kj}^{(2)} + \sum_{i=1}^{m_1} \lambda_i b_i^{(1)} = \\ &= \sum_{i=k}^{m_2} \bar{\sigma}_k \sum_{j=1}^n a_{kj}^{(2)} x_j + \sum_{i=1}^{m_1} \lambda_i b_i^{(1)} \leq \sum_{k=1}^{m_2} \bar{\sigma}_k b_k^{(2)} + \sum_{i=1}^{m_1} \lambda_i b_i^{(1)} \end{aligned} \quad (15)$$

для

$$X \in D, \quad \Lambda \geq 0; \quad \bar{\Sigma} \geq 0; \quad \sum_{i=1}^{m_1} \lambda_i a_{ij}^{(1)} + \sum_{k=1}^{m_2} \bar{\sigma}_k a_{kj}^{(2)} \geq c_j; \quad j = 1, \dots, n.$$

Возьмем $(\Lambda; \Sigma) = (\hat{\Lambda}; \hat{\Sigma})$, где $(\hat{\Lambda}; \hat{\Sigma})$ – оптимальное решение задачи (9)–(11).

Тогда

$$L^* \{ \Lambda^* \} \leq L^* \left(\hat{\Lambda} \right) \leq \sum_{k=1}^{m_2} \hat{\sigma}_k b_k^{(2)} + \sum_{i=1}^{m_1} \hat{\lambda}_i b_i^{(1)}. \quad (16)$$

Сопоставив (16) с (13), (14), где нужно взять $(\Lambda; \bar{\Sigma}) = (\Lambda^*, \Sigma^*)$, получим доказательство леммы.

Из доказанной леммы следует, что решение задачи (9)–(11), двойственной к задаче (1)–(4), сводится к минимизации выпуклой кусочно-линейной функции $L^* (\Lambda)$ при простейших ограничениях $\Lambda \geq 0$.

3. Из выражения для $L^* \{ \Lambda \}$

$$L^* \{ \Lambda \} = \sum_{j=1}^n c_j x_j^* (\Lambda) + \sum_{i=1}^{m_1} \lambda_i \left(b_i^{(1)} - \sum_{j=1}^n a_{ij}^{(1)} x_j^* (\Lambda) \right), \quad (17)$$

где $X^* \{ \Lambda \} = \{ x_j^* (\Lambda) \}$ – оптимальный вектор задачи (6), вытекает, что направление обобщенного градиентного спуска в точке Λ совпадает с направлением, противоположным направлению вектора

$$\bar{\sigma}_\Lambda = \left\{ b_i^{(1)} - \sum_{j=1}^n a_{ij}^{(1)} x_j^* (\Lambda) \right\}; \quad i = 1, \dots, m_1. \quad (18)$$

Отсюда следует итерационный алгоритм для решения задачи (9)–(11), r -ый шаг которого описывается следующим образом:

1) при фиксированном значении $\Lambda^{(r)}$, полученном на предыдущем шаге, решаем задачу (6) ($\Lambda^{(1)}$ можно взять равным 0); получаем вектор $X^* \{ \Lambda^{(r)} \}$ и вектор оптимального решения двойственной задачи $\Sigma \{ \Lambda^{(r)} \}$;

2) вычисляем вектор $\bar{\sigma}_r = \left\{ b_i^{(r)} - \sum_{j=1}^n a_{ij}^{(1)} x_j^* (\Lambda^{(r)}) \right\}$;

3) находим $\Lambda^{(r+1)}$ следующим образом:

$$\Lambda^{(r+1)} = P \left\{ \Lambda^{(r)} - k_r \frac{\bar{\sigma}_r}{|\bar{\sigma}_r|} \right\},$$

где $k_r > 0$ – величина шага, P – оператор, оставляющий неотрицательные координаты без изменения и обращающий отрицательные координаты в 0.

Если выбрать k_r так, чтобы $k_r \xrightarrow{r \rightarrow \infty} 0$ и $\sum_{r=1}^{\infty} k_r = \infty$, то этим гарантируется сходимость $L \{ \Lambda^{(r)} \}$ [1,3] к оптимальному значению задачи (5):

$$L^* \{ \Lambda^{(r)} \} \xrightarrow{(r \rightarrow \infty)} L^* \{ \Lambda^* \},$$

при этом пара векторов $(\Lambda^{(r)}, \Sigma \{ \Lambda^{(r)} \})$ сходится к оптимальному решению задачи (9)–(11).

4. Метод о. г. с. особенно удобно применять к таким задачам линейного программирования, двойственные к которым имеют блочную структуру. Именно такой задачей является двухэтапная задача стохастического программирования с конечным числом значений, которые может принимать случайный вектор правой части ограничений [4, 6].

В настоящее время в ИК АН УССР разработана программа решения этой задачи методом о. г. с. для ЭВМ. Подробности алгоритма будут рассмотрены в статье [5].

Литература

1. ШОР Н.З. О структуре алгоритмов численного решения задач оптимального планирования и проектирования, Автореферат диссертации, К., 1964.
2. ШОР Н.З. Применение метода градиентного спуска для решения сетевой транспортной задачи. Материалы научного семинара по теоретическим и прикладным вопросам кибернетики. Экономическая кибернетика и исследование операций, вып. 1, К., 1962.
3. ЕРМОЛЬЕВ Ю.М. Методы решения нелинейных экстремальных задач // Кибернетика. – 1966. – № 4.
4. ГОЛЬШТЕЙН Е.Г., ЮДИН Д.Б. Новые направления в линейном программировании. – М.: Советское радио, 1966.
5. ШОР Н.З., ЩЕПАКИН М.Б., Алгоритм решения двухэтапной задачи стохастического программирования // Кибернетика. – 1968. – № 3. – С. 56–58.
6. DANTZIG G., MADANSKY A. On the Solution of Two - Stage Linear Programming Problems under Uncertainty. Proc. Fourth Berkeley Symposium on Mathematical Statistics and Probability, № 1, 1961.

Метод случайного поиска для двухэтапной задачи стохастического программирования и его обобщение

Ю. М. Ермольев, Н. З. Шор
Кибернетика. — 1968. — № 1. — С. 90–92.

Двухэтапное стохастическое программирование может применяться тогда, когда требуется составить план на некоторый интервал времени при не полностью определенном будущем. При реализации принятого в такой ситуации плана возникают "невязки", ликвидация которых связана с определенными затратами. Необходимо найти такой план, стоимость реализации которого с учетом ликвидации "невязок" является в среднем минимальной.

Математическая задача ставится следующим образом [1]: пусть заданы:

1) система линейных неравенств вида

$$Ax + Dy \geq B_\omega, \quad (1)$$

$$x \geq 0, \quad y \geq 0, \quad (2)$$

где A — матрица $m \times n_1$, D — матрица $m \times n_2$, x, y — соответственно n_1 - и n_2 -мерные вектора, B_ω — m -мерный случайный вектор с распределением $dP(\omega)$ с ограниченной дисперсией;

2) две линейные формы: $L_1(x) = (C^1, x)$, $L_2(x) = (C^2, y)$, где C^1 — n_1 -мерный, C^2 — n_2 -мерный вектор.

Обозначим через $\Phi(x, B_\omega)$ значение

$$\min (C^2, y) \quad (3)$$

при ограничениях

$$Dy \geq B_\omega - Ax, \quad (4)$$

$$y \geq 0. \quad (5)$$

Будем считать, что для любого $x \geq 0$ и B_ω существует решение задачи (3)–(5). Пусть $E(x) = \int \Phi(x, B_\omega) dP(\omega)$. Требуется найти

$$\min_{x \geq 0} F(x) = \min_{x \geq 0} \left[(C^1, x) + E(x) \right]. \quad (6)$$

1. МЕТОД РЕШЕНИЯ

Построим для решения этой задачи итерационный процесс с использованием случайного поиска. Очевидно, функция $F(x)$ выпуклая, но не обязательно непрерывно-дифференцируемая, поэтому найдем выражение для внутренней нормали опорной гиперплоскости к телу $\{x : F(x) \leq F(x^0)\}$ в некоторой точке x^0 , т. е. направление обобщенного градиента [2]. Рассмотрим двойственную к (3)–(5) задачу: найти

$$G(x, B_\omega) = \max \left[(\lambda, B_\omega) - (x, A^* \lambda) \right] \quad (7)$$

при ограничениях

$$D^* \lambda \leq C^2, \quad (8)$$

$$\lambda \geq 0. \quad (9)$$

Из теоремы двойственности вытекает, что

$$G(x, B_\omega) = \Phi(x, B_\omega). \quad (10)$$

Нетрудно заметить, что $G(x, B_\omega)$ – выпуклая функция от x . Пусть $\lambda(x, B_\omega)$ – точка, в которой достигается максимум в (7). Будем считать область R , высекаемую ограничениями (8)–(9), ограниченной. Тогда $\lambda(x, B_\omega)$ можно выбирать совпадающей с одной из вершин R . В случае неоднозначности решения задачи (7)–(9) $\lambda(x, B_\omega)$ можно выбирать, например, следующим образом: перенумеруем условно вершины R некоторым образом и будем выбирать $\lambda(x, B_\omega)$ равной вершине с наименьшим номером из вершин, являющихся решением этой задачи. При таком определении $\lambda(x, B_\omega)$ будет представлять кусочно-постоянную вектор-функцию от x и B_ω , $G(x, B_\omega)$ – кусочно-линейную функцию от x , измеримую по мере Лебега – Стильтьеса $dP(\omega)$. Зафиксируем некоторую точку $x = x^0$. Тогда справедливы соотношения:

$$G(x, B_\omega) \geq \left[(\lambda(x^0, B_\omega), B_\omega) - (x, A^* \lambda(x^0, B_\omega)) \right],$$

$$G(x^0, B_\omega) = \left[(\lambda(x^0, B_\omega), B_\omega) - (x^0, A^* \lambda(x^0, B_\omega)) \right].$$

Отсюда следует, что

$$E(x) \geq \int (\lambda(x^0, B_\omega), B_\omega) dP(\omega) - \int (x, A^* \lambda(x^0, B_\omega)) dP(\omega),$$

$$E(x^0) = \int (\lambda(x^0, B_\omega), B_\omega) dP(\omega) - \int (x^0, A^* \lambda(x^0, B_\omega)) dP(\omega)$$

или

$$E(x) - E(x^0) \geq - \int (x - x^0, A^* \lambda(x^0, B_\omega)) dP(\omega). \quad (11)$$

Рассмотрим случайный вектор

$$C^1 - A^* \lambda(x^0, B_\omega).$$

В соответствии с неравенством (11) математическое ожидание этого вектора и является искомым вектором внутренней нормали опорной гиперплоскости к телу $\{x : F(x) \leq F(x^0)\}$.

Таким образом, вычисление направления обобщенного градиента функции $F(x)$ сводится к вычислению интегралов (11), что при известном $dP(\omega)$ приводит к решению параметрической задачи линейного программирования. В частном случае, когда B_ω принимает конечное число значений, получаем алгоритм, описанный в [3]. Однако, решение параметрической задачи вызывает большие трудности. Кроме того, вектор B_ω часто получается путем случайных испытаний и распределение $dP(\omega)$ а priori не известно. Поэтому, в общем случае предлагается алгоритм случайного поиска следующего вида.

Пусть на s -ом шаге получено значение $x^{(s)}$. Тогда $(s+1)$ -й шаг описывается следующим образом:

- а) выбираем случайную реализацию $B_\omega^{(s)}$ в соответствии с $dP(\omega)$;
- б) находим $\lambda(x^{(s)}, B_\omega^{(s)})$, решая задачу (7)–(9) при $x = x^s, B_\omega = B_\omega^{(s)}$;
- в) находим

$$x^{(s+1)} = \max \left\{ \mathbf{0}, x^{(s)} + \varrho_s \left[C^1 - A^* \lambda \left(x^{(s)}, B_\omega^{(s)} \right) \right] \right\}, \quad (12)$$

где ϱ_s – величина шага. Движение точки $x^{(s)}$ при этом происходит в случайном направлении, математическое ожидание которого совпадает с направлением обобщенного градиентного спуска.

2. СХОДИМОСТЬ МЕТОДА И ЕГО ОБОБЩЕНИЕ

Сущность предлагаемого метода случайного поиска в общем случае заключается в следующем. Пусть требуется минимизировать выпуклую вниз функцию $F(x) = F(x_1, \dots, x_n)$. Рассматривается итеративный процесс

$$x^{(s+1)} = x^{(s)} + \varrho_s \xi^{(s)}, \quad s = 0, 1, \dots, \quad (13)$$

где $x^{(0)}$ — произвольная точка, ϱ_s — величина шага, $\xi^{(s)}$ — случайное направление такое, что

$$M \left\{ \xi^{(s)} \mid x^{(1)}, \dots, x^{(s)} \right\} = \widehat{F}_x \left(x^{(s)} \right), \quad (14)$$

где $\widehat{F}(x)$ удовлетворяет неравенству

$$F(y) - F(x) \geq \left(\widehat{F}_x(x), y - x \right). \quad (15)$$

Пусть $\min F(x) = F(x^*) > -\infty$. Без ограничения общности считаем, что точка x^* единственная.

Теорема 6. Пусть

$$(i) \quad \varrho_s > 0, \quad \sum_{s=1}^{\infty} \varrho_s = \infty, \quad \sum_{s=1}^{\infty} \varrho_s^2 < \infty;$$

$$(ii) \quad M \left(\left\| \xi^{(s)} \right\|^2 \mid x^{(1)}, \dots, x^{(s)} \right) \leq c < \infty.$$

Тогда $\lim \|x^* - x^{(s)}\| = 0$, $s \rightarrow \infty$ с вероятностью 1.

Очевидно,

$$\left\| x^* - x^{(s+1)} \right\|^2 = \left\| x^* - x^{(s)} \right\|^2 + 2\varrho_s \left(\xi^{(s)}, x^* - x^{(s)} \right) + \varrho_s^2 \left\| \xi^{(s)} \right\|^2. \quad (16)$$

Возьмем от обеих частей этого равенства условное математическое ожидание:

$$\begin{aligned} M \left\{ \left\| x^* - x^{(s+1)} \right\|^2 \mid x^{(1)}, \dots, x^{(s)} \right\} &= \left\| x^* - x^{(s)} \right\|^2 + \\ + 2\varrho_s \left(\widehat{F}_x \left(x^{(s)} \right), x^* - x^{(s)} \right) &+ \varrho_s^2 M \left\{ \left\| \xi^{(s)} \right\|^2 \mid x^{(1)}, \dots, x^{(s)} \right\}. \end{aligned}$$

Из неравенства (15) следует, что

$$\left(\widehat{F}_x \left(x^{(s)} \right), x^* - x^{(s)} \right) \leq 0,$$

поэтому

$$M \left\{ \left\| x^* - x^{(s+1)} \right\|^2 \mid x^{(1)}, \dots, x^{(s)} \right\} \leq \left\| x^* - x^{(s)} \right\|^2 + c \varrho_s^2. \quad (17)$$

Положим $z^{(s)} = \left\| x^* - x^{(s)} \right\|^2 + c \sum_{k=s}^{\infty} \varrho_k^2$. Тогда неравенство (17) равносильно следующему:

$$M \left\{ z^{(s+1)} \mid z^{(1)}, \dots, z^{(s)} \right\} \leq z^{(s)}.$$

Следовательно, $z^{(s)}$ – полумартингал [4], и поэтому последовательность $\{z^{(s)}\}$ сходится к пределу с вероятностью 1, а значит в силу (i) и последовательность $\left\| x^* - x^{(s)} \right\|$ сходится к пределу с вероятностью 1.

Покажем теперь, что $\lim \left\| x^* - x^{(s)}(\omega) \right\|^2 = 0$ при $s \rightarrow \infty$. Имеем

$$\begin{aligned} M \left\{ \left\| x^* - x^{(s+1)} \right\|^2 \mid x^{(1)}(\omega), \dots, x^{(s)}(\omega) \right\} &\leq \left\| x^* - x^{(0)} \right\|^2 + \\ &+ 2 \sum_{k=1}^s \varrho_k \left(\widehat{F}_x \left(x^{(k)}(\omega) \right), x^* - x^{(k)}(\omega) \right) + c \sum_{k=1}^s \varrho_k^2. \end{aligned}$$

Левая часть этого неравенства ограничена, поэтому из условия (i) следует, что

$$\sum_{s=1}^{\infty} \varrho_s \left(\widehat{F}_x \left(x^{(s)}(\omega) \right), x^* - x^{(s)}(\omega) \right) > -\infty,$$

т. е., $\left(\widehat{F}_x \left(x^{(s)}(\omega) \right), x^* - x^{(s)}(\omega) \right) > -\infty$ при $s \rightarrow \infty$, а поэтому и

$$\lim \left\| x^* - x^{(s)}(\omega) \right\| = 0.$$

В случае, когда требуется минимизировать $F(x)$ при условии $x \geq 0$, процесс (13) видоизменяется следующим образом:

$$x^{(s+1)} = \max \left\{ 0, x^{(s)} + \varrho_s \xi^{(s)} \right\}.$$

Покажем теперь, что для процесса (12) условия теоремы выполняются. Действительно, из предположения об ограниченности области R , вытекающей из неравенств (8)–(9), следует ограниченность нормы случайного вектора $C^1 - A^* \lambda(x, B_\omega)$. Таким образом, для сходимости процесса (12) достаточно только условия (i).

Рассмотрим теперь задачу двухэтапного стохастического программирования более общего вида. Пусть имеются функции $f^\nu(x, y, \omega)$, $\nu = 0, 1, \dots, m$, зависящие от плана x , его "коррекции" y и случайного события ω , выпуклые вниз при любом ω по x, y . При фиксированных x, ω можно найти вектор $y(x, \omega)$, минимизирующий

$$f^0(x, y, \omega) \quad (18)$$

при условиях

$$f^i(x, y, \omega) \leq 0, \quad i = 1, \dots, m. \quad (19)$$

Особенно легко это сделать, если $f^\nu(x, y, \omega)$ линейны по y . Пусть $\lambda(x, \omega)$ – множители Лагранжа, при которых пара $(y(x, \omega), \lambda(x, \omega))$ образует седловую точку задачи (18)–(19). Тогда можно показать, что математическое ожидание случайного вектора

$$\xi(x, \omega) = f_x^0(x, y(x, \omega), \omega) + \sum_{i=1}^m \lambda_i(x, \omega) \times f_x^i(x, y(x, \omega), \omega)$$

совпадает с вектором обобщенного градиента функции

$$F(x) = M f^0(x, y(x, \omega), \omega). \quad (20)$$

Иначе говоря, для минимизации (20) можно предложить метод случайного поиска, аналогичный (12). Сходимость его следует из общей теоремы [6]. Интересно отметить, что минимизацию (20) можно рассматривать также как стохастическую задачу нелинейного параметрического программирования, а процесс случайного поиска вида (12) – как случайный итеративный процесс для решения этой нелинейной параметрической задачи.

Литература

1. ДАНЦИГ ДЖ. Линейное программирование, его применения и обобщения. – М.: "Прогресс", 1966.
2. ЕРМОЛЬЕВ Ю.М. Методы решения нелинейных экстремальных задач // Кибернетика. – 1966. – № 4. – С. 1–7.
3. ШОР Н.З., ЩЕПАКИН М.Б. Алгоритм решения двухэтапной задачи стохастического программирования // Теория оптимальных решений. – К.: Ин-т кибернетики АН УССР. – 1967. – № 2. – С. 38–44.
4. ДУБ Д. Вероятностные процессы. – М.: ИЛ, 1956.
5. ЕРМОЛЬЕВ Ю.М., НЕКРЫЛОВА З.В. О некоторых методах стохастической оптимизации // Кибернетика. – 1966. – № 6. – С. 96–98.
6. ЕРМОЛЬЕВ Ю.М., НЕКРЫЛОВА З.В. Метод стохастических градиентов и его применение // Теория оптимальных решений. – К.: Ин-т кибернетики АН УССР. – 1967. – № 1. – С. 24–47.

$$x \geq 0; \quad z_r \geq 0; \quad r = 1, \dots, k. \quad (8)$$

Задача, двойственная к задаче (6)–(8), записывается так.

Найти

$$\max \sum_{r=1}^k (b_r, y_r) \quad (9)$$

при ограничениях

$$A^T y_1 + \dots + A^T y_k \leq c; \quad (10)$$

$$\left\{ \begin{array}{l} D^T y_1 \leq p_1 h; \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ D^T y_k \leq p_k h; \end{array} \right. \quad (11)$$

$$y_r \geq 0; \quad r = 1, \dots, k. \quad (12)$$

Как видим, задача (9)–(12) представляет удачный объект для применения блочного программирования. В [3] был предложен алгоритм, основанный на применении обобщенного градиентного спуска, для решения задачи линейного программирования, двойственная к которой имеет блочную структуру. Но задачей, двойственной к задаче (9)–(12), будет задача (6)–(8), т. е. исходная задача стохастического программирования.

Следуя статье [3], рассмотрим следующую задачу.

Найти

$$\min_{x \geq 0} \max_{y \in R} \left[\sum_{r=1}^k (b_r, y_r) + \left(x, c - \sum_{r=1}^k A^T y_r \right) \right], \quad (13)$$

где $y = \{y_1, \dots, y_r, \dots, y_k\}$, R – область, вырезаемая ограничениями (11).

Предположим, что R не пусто и является ограниченным многогранником. Тогда для произвольного $x \geq 0$ существует $y^*(x) = \{y_1^*, \dots, y_k^*\}$, являющийся оптимальным планом задачи линейного программирования:

найти

$$\Phi(x) = \max_{y \in R} \left[\sum_{r=1}^k (b_r, y_r) + \left(x, c - \sum_{r=1}^k A^T y_r \right) \right]. \quad (14)$$

Отсюда

$$\Phi(x) = \sum_{r=1}^k (b_r, y_r^*) + \left(x, c - \sum_{r=1}^k A^T y_r^* \right) \quad (15)$$

– выпуклая кусочно-линейная функция.

Для нахождения $\min_{x \geq 0} \Phi(x)$ применим метод обобщенного градиентного спуска, $(s+1)$ -й шаг которого состоит из следующих элементов:

1) для фиксированного значения $x = x^{(s)}$, полученного на предыдущем шаге, решаем задачу (14);

2) получаем

$$x^{(s+1)} = P^+ \left\{ x^{(s)} - g_{s+1} \frac{c - \sum_{r=1}^k A^T y_r}{\left| c - \sum_{r=1}^k A^T y_r \right|} \right\}, \quad (16)$$

где g_{s+1} – величина шага, P^+ – оператор, обращающий в 0 отрицательные компоненты вектора и оставляющий без изменения остальные.

Ниже будет дано обоснование одного из способов регулировки величины шага g_s , обеспечивающего сходимость последовательности $\Phi(x^{(s)})$ к $\min_{x \geq 0} \Phi(x)$. Суть этого способа состоит в том, что g_s выбирается равным $g_0 \cdot 2^{-\lfloor \frac{s}{t} \rfloor}$, где g_0 , t – достаточно большие положительные числа (t – целое число).

Лемма. Пусть $F(x)$ – выпуклая функция, определенная на E_n , M^* – множество значений, на котором $F(x)$ принимает минимальное значение. Предположим, что M^* – не пустое, $\varrho(x, y)$ – евклидово расстояние между точками в E_n .

Пусть $x^{(0)}, x^{(1)}, \dots, x^{(s)}, \dots$ – последовательность значений x , получающаяся при применении обобщенного градиентного спуска к функции $F(x)$, начиная с точки $x^{(0)}$, с постоянным шагом g . Тогда для любого $x^* \in M^*$ найдутся s и y такие, что:

$$1) F(x^{(s)}) \leq F(y);$$

$$2) s \leq \left\lceil \frac{\varrho^2(x^{(0)}, x^*)}{g^2} \right\rceil;$$

$$3) \varrho(y, x^*) \leq g.$$

Доказательство. Обозначим через Π_x опорную плоскость к телу $M_x = \{y : F(y) \leq F(x)\}$; ($x \in M^*$). Возьмём точку $x^{(r)}$ из заданной последовательности $\{x^{(i)}\}$, $i = 0, 1, \dots$. Пусть $x^{(r)} \in M^*$; $p_{x^*}^{(r)}$ — проекция точки $x^* \in M^*$ на $\Pi_{x^{(r)}}$. Из определения обобщенного градиентного спуска следует, что

$$\begin{aligned} \varrho^2(x^{(r+1)}, x^*) &= \left(\varrho(p_{x^*}^{(r)}, x^*) - g\right)^2 + \varrho^2(x^{(r)}, x^*) - \varrho^2(p_{x^*}^{(r)}, x^*) = \\ &= \varrho^2(x^{(r)}, x^*) - 2g\varrho(p_{x^*}^{(r)}, x^*) + g^2. \end{aligned} \quad (17)$$

Возможны два случая.

1. Для некоторого $s < \left\lfloor \frac{\varrho^2(x^{(0)}, x^*)}{g^2} \right\rfloor$ $\varrho(p_{x^*}^{(s)}, x^*) \leq g$. Ясно, что $F(x^{(s)}) \leq F(p_{x^*}^{(s)})$. Возьмем $y = p_{x^*}^{(s)}$, и утверждение леммы будет справедливым.

2. Для всех $s < \left\lfloor \frac{\varrho^2(x^{(0)}, x^*)}{g^2} \right\rfloor = m$ $\varrho(p_{x^*}^{(s)}, x^*) > g$. Но в этом случае из (17) вытекает:

$$\varrho^2(x^{(s+1)}, x^*) \leq \varrho^2(x^{(s)}, x^*) - g^2,$$

т. е.

$$\varrho^2(x^{(m)}, x^*) \leq \varrho^2(x^{(0)}, x^*) - g^2 \left\lfloor \frac{\varrho^2(x^{(0)}, x^*)}{g^2} \right\rfloor \leq g^2,$$

и если взять $s = m$, $x^{(s)} = y$, то утверждение леммы будет справедливым.

Лемма доказана.

Замечание. Лемма остается справедливой также для решения задачи: найти $\min_{x \geq 0} F(x)$ с применением оператора P^+ (см. (16)), так как для $y \geq 0$

$$\varrho[P^+(x), y] \leq \varrho(x, y).$$

Пусть для произвольной пары точек $y_1, y_2 \in M^*$ и таких, что $F(y_1) = F(y_2)$, выполняются соотношения:

$$\frac{\varrho(y_1, M^*)}{\varrho(y_2, M^*)} \leq \mu; \quad \left(\varrho(x, M^*) = \inf_{x^* \in M^*} \varrho(x, x^*) \right). \quad (18)$$

Для заданного $x^{(0)}$ выберем g_0 такое, что $\varrho(x^{(0)}, M^*) \leq 2g_0\mu$, и будем производить обобщенный градиентный спуск с шагом $g = g_0$. Из доказательства леммы видно, что либо найдется $s < [3\mu^2] + 1$, что $F(x^{(s)}) \leq F(z)$ и $\varrho(z, M^*) \leq g_0$ либо для $t = [3\mu^2] + 1$, $\varrho^2(x^{(t)}, M^*) \leq 4g_0^2\mu^2 - g_0^2 3\mu^2 = g_0^2\mu^2$. Используя (17), получим, что в обоих случаях найдется $s \leq t$, что

$$\varrho(x^{(s)}, M^*) \leq g_0\mu.$$

Пусть $x^{(s)} = x_1$ (s можно брать равным индексу, на котором достигается $\min_i F(x^{(i)})$, $i=0, \dots, t$).

Если начать с точки x_1 , то, проделав t шагов спуска с шагом $g = g_0/2$, мы найдем точку $x_1^{(s_1)} = x_2$ такую, что $\varrho(x_2, M^*) \leq \frac{g_0\mu}{2}$. Таким образом, можно построить $\{x_\varrho\}$ для любого ϱ такие, что

$$\varrho(x_\varrho, M^*) \leq \frac{g_0\mu}{2\varrho}. \quad (19)$$

Если F – кусочно – линейная функция, то условие (18) всегда выполняется.

Таким образом, мы доказали, что для любой начальной точки $x^{(0)}$ можно найти такие g_0 и t , что если брать величину s -го шага $g_s = \frac{g_0}{2^{[s/t]}}$, то процесс обобщенного градиентного спуска для функции $\Phi(x)$ сходится; при этом обеспечивается экспоненциальная сходимость. (Сходимость понимается в таком смысле: $\varrho(x^{(s)}, M^*) \xrightarrow{s \rightarrow \infty} 0$.)

Из результатов [3] следует, что значения x^* , дающие $\min_{x \geq 0} \Phi(x)$, являются оптимальными планами для первого этапа двухэтапной задачи стохастического программирования, а значения y^* , дающие оптимальный план задачи (14) при $x = x^*$, являются наборами оптимальных планов для второго этапа этой задачи при различных b_r .

В приложении приводится программа на языке "АЛГОЛ" для решения задачи (6)–(8), основанная на нахождении $\min_{x \geq 0} \Phi(x)$ методом обобщенного градиентного спуска.

ПРИЛОЖЕНИЕ

Приведенная ниже программа решает двухэтапную задачу стохастического программирования с конечным числом значений случайного вектора b_r . В программе используются следующие исходные данные: транспонированные матрицы коэффициентов A и D (массивы a и d), векторы b_r (массив b), вектор стоимости c и вектор "штрафов" h (массивы c и h), а также вероятности значений $b_r - p_r$ (массив p). Кроме того, необходимо задать шаг g , число итераций при постоянном шаге t и число изменений шага t_0 .

```

begin integer  $m, n, n1, n2, l, t0, t$ ;
      read ( $m, n, n1, n2, l, t0, t$ );
begin real  $s, c0, g$ ; integer  $f0, f, i, j, k$ ;
      real array  $b, y[1 : n, 1 : m], a[1 : n1, 1 : m]$ ,
       $d[1 : n2, 1 : m]$ ;
       $u, z[1 : n, 1 : n2], c, r, x[1 : n1], p[1 : n]$ ,
       $h[1 : n2], v[1 : m]$ ;
      read  $g, b, a, d, c, p, h$ ;
      procedure simplex; comment описание этой процедуры из-за
      его громоздкости и общеизвестности здесь не приводится. Она
      вычисляет  $y[k, j]$ , сообщающие  $\max \sum_{j=1}^m v[j] \times y[k, j]$ , и соответствую-
      ющие двойственные переменные  $z[k, l]$  при условиях  $\sum_{j=1}^m d[l, j] \times$ 
       $y[k, j] \leq u[k, l], y[k, j] \geq 0, k -$  фиксировано. Для каждой из за-
      дач, начиная со второй итерации, в качестве начального опорного
      плана берется оптимальное решение соответствующей задачи на
      предыдущей итерации;
      fo := 1;
      for  $i := 1$  step 1 until  $n1$  do  $x[i] := 0$ ;
q1 :    $f := 1$ ;
q2 :   for  $i := 1$  step 1 until  $n1$  do  $r[i] := 0$ ;
       $s := 0$ ;
      for  $k := 1$  step 1 until  $n$  do
begin for  $j := 1$  step 1 until  $m$  do
begin  $v[j] := 0$ ;

```

```

    for  $i := 1$  step 1 until  $n1$  do
       $v[j] := v[j] + x[i] \times a[i, j]$ ;
       $v[j] := b[k, j] - v[j]$ ;
end;
  j
  for  $l := 1$  step 1 until  $n2$  do
     $u[k, l] := h[l] \times p[k]$ ;
    procedure simplex;
    print  $z[k, 1 : n2]$ ;
    for  $j := 1$  step 1 until  $m$  do
       $s := s + b[k, j] \times y[k, j]$ ;
    for  $i := 1$  step 1 until  $n1$  do
      for  $j := 1$  step 1 until  $m$  do
         $r[i] := r[i] + a[i, j] \times y[k, j]$ ;
      end;
    end;
    k
    for  $i := 1$  step 1 until  $n1$  do
       $r[i] := r[i] - c[i]$ ;
       $c0 := 0$ ;
      for  $i := 1$  step 1 until  $n1$  do
         $c0 := c0 + r[i] \uparrow 2$ ;
      end;
       $c0 := \text{sqrt } c0$ ;
      for  $i := 1$  step 1 until  $n1$  do
begin
         $x[i] := x[i] + g \times r[i] / c0$ ;
        if  $x[i] \leq 0$  then  $x[i] := 0$ 
end;
      end;
      print  $(s, x)$ ;
       $f := f + 1$ ;
      if  $f \leq t$  then go to q2 else  $g := g/2$ ;
       $f0 := f0 + 1$ ;
      if  $f0 \leq t0$  then go to q1;
end
end;
```

Литература

1. Данциг Дж. Линейное программирование, его применения и обобщения. – М.: "Прогресс", 1966.
2. Гольштейн Е.Г., Юдин Д.Б. Новые направления в линейном программировании. – М.: "Советское радио", 1966.
3. Шор Н.З. Применение обобщенного градиентного спуска в блочном программировании // Кибернетика. – 1967. – № 3. – С. 53-55.

Метод минимизации, использующий операцию растяжения пространства в направлении разности двух последовательных градиентов

Н. З. Шор, Н. Г. Журбенко
Кибернетика. – 1971. – № 3. – С. 51–59.

В работах [1], [2] описан класс алгоритмов оптимизации градиентного типа с использованием операции растяжения пространства в направлении градиента (или его аналога) (ОГСРП-алгоритмы). Серьезной проблемой при построении такого рода алгоритмов был вопрос выбора на каждой итерации величины шага спуска, который сравнительно просто разрешался в том случае, когда было заранее известно значение функции в точке минимума. В более общем случае приходилось усложнять алгоритмы [2]. В данной статье предлагается новый класс алгоритмов, основанный на использовании операции растяжения пространства в направлении разности двух последовательных градиентов. В отличие от вышеупомянутых алгоритмов типа ОГСРП этот класс позволяет применить способ выбора шага, близкий к применяемому в методе наискорейшего спуска. При этом получаются алгоритмы, обеспечивающие монотонность или, в некотором смысле, "почти" монотонность спуска.

Дадим описание общей структуры алгоритмов с растяжением пространства в направлении разности двух последовательных градиентов. Алгоритмы этого класса будем сокращенно называть r -алгоритмами. Затем для одного предельного варианта r -алгоритмов докажем, при определенных условиях, квадратичную скорость сходимости. Будут рассмотрены вопросы сходимости других вариантов r -алгоритмов и дан анализ проведенных численных экспериментов.

1. Общая структура r -алгоритмов

Пусть $f(x)$ – непрерывно дифференцируемая функция, определенная в n -мерном евклидовом пространстве E_n , и выполняется условие

$$\lim_{\|x\| \rightarrow \infty} f(x) = +\infty, \quad (1)$$

$g_f(x)$ – градиент функции $f(x)$ в точке x . Рассмотрим следующий итеративный процесс минимизации $f(x)$.

Первый шаг задается начальным значением $x_0 \in E_n$; вычисляем $g_f(x_0)$, выбираем $h_1 > 0$ и находим

$$x_1 = x_0 - h_1 g_f(x_0).$$

Принимаем $g_f(x_0) = \tilde{g}_1$; $B = E$ (единичная матрица). Переходим к второму шагу.

Пусть в результате вычислений после k шагов процесса ($k = 1, 2, \dots$) 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. $g_f(x_k)$ – градиент функции $f(x)$ в точке x_k .
2. $B_k^* g_f(x_k) = g_k^*$.

g_k^* – градиент функции $\varphi_k(y) = f(B_k y)$ в точке $y = A_k x_k$, где $A_k = B_k^{-1}$.

$$3. \quad r_k = g_k^* - \tilde{g}_k. \quad (2)$$

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 $k = 1, 2, \dots$:

$$r_k = B_k^* [g_f(x_k) - g_f(x_{k-1})], \quad (3)$$

однако использование формулы (2) дает некоторую экономию вычислений в целом.

$$4. \quad \xi_{k+1} = \frac{r_k}{\|r_k\|}.$$

Нормировка вектора r_k нужна для подготовки очередной операции растяжения пространства в направлении, определяемом вектором r_k .

5. β_{k+1} – величина, обратная коэффициенту растяжения пространства на $(k + 1)$ -ом шаге.

$$6. \quad B_{k+1} = B_k \cdot R_{\beta_{k+1}}(\xi_{k+1}).$$

$B_{k+1} = A_{k+1}$, где $A_{k+1} = R_{\frac{1}{\beta_{k+1}}}(\xi_{k+1})A_k$ – матрица преобразования пространства после $(k + 1)$ -го шага; здесь $R_\alpha(\xi)$ – оператор растяжения пространства в направлении ξ с коэффициентом α ($\alpha \geq 0$, $\|\xi\| = 1$).

Подробно свойства этого оператора описаны в [1]. Отметим основную формулу:

$$R_\alpha(\xi)x = x + (\alpha - 1)(x, \xi)\xi.$$

7. h_{k+1} – шаговый множитель в формуле спуска в заданном направлении.

$$\mathbf{8.} \quad \tilde{g}_{k+1} = R_{\beta_{k+1}}(\xi_{k+1}) g_k^*$$

$$\begin{aligned} \tilde{g}_{k+1} = R_{\beta_{k+1}}(\xi_{k+1}) g_k^* &= R_{\beta_{k+1}}(\xi_{k+1}) B_k^* g_f(x_k) = \\ &= \left(B_k R_{\beta_{k+1}}(\xi_{k+1}) \right)^* g_f(x_k) = B_{k+1}^* g_f(x_k). \end{aligned}$$

Отсюда \tilde{g}_{k+1} – значение градиента функции $\varphi_{k+1}(y) = f(B_{k+1}y)$ в точке $\tilde{y}_{k+1} = A_{k+1}x_k$.

$$\mathbf{9.} \quad x_{k+1} = x_k - h_{k+1} B_{k+1} \tilde{g}_{k+1}. \tag{4}$$

Применим к обеим частям формулы (4) оператор A_{k+1} . Тогда

$$y_{k+1} = A_{k+1}x_{k+1} = A_{k+1}x_k - h_{k+1}\tilde{g}_{k+1} = \tilde{y}_{k+1} - h_{k+1}\tilde{g}_{k+1}.$$

Таким образом, формула (9) 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 мы будем называть r -алгоритмами. Имеется еще одна возможность модификации r -алгоритмов – использование так называемой операции восстановления, когда периодически после заданного числа итераций происходит "восстановление" матрицы B_k , т. е. она заменяется единичной матрицей. Вопрос о сходимости алгоритмов с монотонным изменением значения минимизирующей функции и с 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 [3]. В частности, если $f(x)$ непрерывно дифференцируема и h_{k+1} в формуле

(4) выбирается из условия минимума по направлению, то при использовании r -алгоритмов с восстановлением множество предельных точек последовательности $\{x\}$ 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.

Мы рассмотрим более подробно один предельный вариант r -алгоритма с восстановлением, в котором β_k , $k = 1, 2, \dots$, выбирается равным нулю, а h_{k+1} выбирается из условия минимума выражения

$$f(x_k - h_{k+1} B_{k+1} \tilde{g}_{k+1}).$$

При этом, как легко заметить, либо после некоторого числа $k^* < n$ шагов g_{k^*+1} станет равным 0, либо $B_{n+1} = 0$. В самом деле, оператор $R_0(\eta)$ означает проектирование на подпространство, ортогональное вектору η .

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 η_i , $i = 1, 2, \dots$

Запишем цепочку равенств

$$r_1 = g_f(x_1) - g_f(x_0); \dots$$

$$r_k = R_0\left(\frac{r_{k-1}}{\|r_{k-1}\|}\right) \cdot R_0\left(\frac{r_{k-2}}{\|r_{k-2}\|}\right) \cdots R_0\left(\frac{r_1}{\|r_1\|}\right) \cdot [g_f(x_k) - g_f(x_{k-1})];$$

$$k > 1.$$

Мы видим, что векторы r_1, \dots, r_k получаются в результате процесса последовательной ортогонализации, примененной к векторам

$$g_f(x_1) - g_f(x_0), \quad g_f(x_2) - g_f(x_1), \dots, g_f(x_k) - g_f(x_{k-1}),$$

$k = 1, 2, \dots$. Если r_1, \dots, r_n отличны от нуля, то оператор $B_{n+1} = 0$, так как он означает проекцию на подпространство, ортогонально дополняющее всё пространство. Если для некоторого $k \leq n$, $r_k = 0$, то это означает, что

$$B_{k-1}^* [g_f(x_k) - g_f(x_{k-1})] = 0,$$

а это возможно лишь при условии

$$B_k^* g_f(x_{k-1}) = \tilde{g}_k = 0 \text{ и } x_k = x_{k-1}.$$

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. Отметим, что доказательство аналогичной теоремы для других модификаций метода сопряженных градиентов встречает серьезные трудности [4].

Теорема. 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 Липшица

$$\|H(x) - H(x')\| \leq L \|x - x'\|; \quad x, x' \in S. \quad (5)$$

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 работы приведенного выше алгоритма (если для некоторого $\bar{k} < n \tilde{g}_{\bar{k}}$, 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 и минимальное – символом m_0 . В силу положительной определенности $H(0)$ $m_0 > 0$. Исходя из непрерывности гессиана $H(x)$, получаем, что найдется такая окрестность нуля $S_\delta = \{x : \|x\| < \delta\}$ и такие положительные числа M и m , что при $x \in S_\delta$ максимальное собственное число оператора $H(x)$ не будет превышать M , а минимальное будет не меньше m .

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 (5), получаем

$$\|G(\mu e)\| \leq \mu L,$$

откуда

$$g_f(x_0) = \varepsilon \left[H(0)e + r(x_0) \right], \quad (6)$$

причем

$$\|r(x_0)\| \leq \frac{1}{\varepsilon} \int_0^\varepsilon \mu L d\mu = \frac{L\varepsilon}{2}. \quad (7)$$

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 ε

$$h^* \leq \frac{\|g_f(x_0)\|}{m} \leq \frac{M\varepsilon}{m}.$$

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_δ

$$\frac{d^2\varphi(h)}{dh^2} \geq m,$$

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 $\left\| G_1\left(\mu \frac{g_f(x_0)}{\|g_f(x_0)\|}\right) \right\| \leq L\mu$. Отсюда

$$\begin{aligned}
 g_f(x_1) &= g_f\left(x_0 - h^* \frac{g_f(x_0)}{\|g_f(x_0)\|}\right) = \\
 &= g_f(x_0) - h^* H(x_0) \frac{g_f(x_0)}{\|g_f(x_0)\|} + h^* r_1(x_0), \quad (8)
 \end{aligned}$$

причем

$$\|r_1(x_0)\| \leq \frac{Lh^*}{2} < \frac{LM_\varepsilon}{2m}. \quad (9)$$

Используя (6)–(9), получаем

$$\begin{aligned}
 &\left(g_f(x_0), g_f(x_1) \right) = \\
 &= \varepsilon^2 \left[\left(H(0)e + r(x_0), H(0)e + r(x_0) \right) - \frac{h^*}{\|g_f(x_0)\|} \times \right. \\
 &\quad \times \left. \left(\left(H(0) + G(x_0) \right) \left(H(0)e + r(x_0) \right), H(0)e + r(x_0) \right) \right] + \\
 &\quad + \varepsilon h^* \left(H(0)e + r(x_0), r_1(x_0) \right) = \\
 &= \varepsilon^2 \left[\left(H(0)e, H(0)e \right) - \frac{h^*}{\|g_f(x_0)\|} \left(H^2(0)e, H(0)e \right) \right] + r_2(x_0),
 \end{aligned}$$

причем при достаточно малом ε , как легко видеть, $|r_2(x_0)| \leq d\varepsilon^3$, где d – некоторое положительное число. Приравняв $\left(g_f(x_0), g_f(x_1) \right)$ нулю,

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 ε^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

$$s = \frac{|(g_f(x_0) - g_f(x_1), x_1)|}{\|g_f(x_0) - g_f(x_1)\|},$$

$$x_1 = x_0 - h^* \frac{g_f(x_0)}{\|g_f(x_0)\|} = x_0 - \frac{\|H(0) e\|^2 g_f(x_0)}{(H^2(0) e, H(0) e)} - \frac{\nu q g_f(x_0)}{\|g_f(x_0)\|}, \quad (10)$$

где

$$-\frac{d\varepsilon}{\|H(0) e\|^2} \leq \nu \leq \frac{d\varepsilon}{\|H(0) e\|^2}. \quad (11)$$

Далее, используя (8)–(11), получаем

$$\begin{aligned} & (g_f(x_0) - g_f(x_1), x_1) = \\ & = \frac{h^* \varepsilon^2}{\|g_f(x_0)\|} \left((H(0) + G(x_0)) (H(0) e + r(x_0)), \right. \\ & \quad \left. e - \frac{\|H(0) e\|^2}{(H^2(0) e, H(0) e)} (H(0) e + r(x_0)) \right) - \end{aligned}$$

$$\begin{aligned}
& - \frac{1}{\|g_f(x_0)\|} \left(\nu q g_f(x_0), g_f(x_0) - g_f(x_1) \right) - \left(h^* r_1(x_0), x_1 \right) = \\
& = \varepsilon^2 \frac{h^*}{\|g_f(x_0)\|} \left(H^2(0)e, e \right) - \left(H(0)e, H(0)e \right) + r_3(x_0) = r_3(x_0),
\end{aligned}$$

где $|r_3(x_0)| \leq c_1 \varepsilon^3$, а c_1 – положительное число, которое легко вычисляется с использованием оценок (7), (9), (11). Так как

$$\left(g_f(x_0), g_f(x_1) \right) = 0$$

то

$$\|g_f(x_1) - g_f(x_0)\| \geq \|g_f(x_0)\| \geq m\varepsilon.$$

Отсюда

$$s \leq \frac{c_1}{m} \varepsilon^2 = c_2 \varepsilon^2, \quad \text{где } c_2 = \frac{c_1}{m}. \quad (12)$$

Рассмотрим функцию $\psi(x)$, определенную в p -мерной гиперплоскости

$$L(x_0) = \left\{ x : \left(g_f(x_1) - g_f(x_0), x - x_1 \right) = 0 \right\}$$

и равную $f(x)$ для точек этой гиперплоскости.

Покажем, что $\min \psi(x) = \min_{x \in L(x_0)} f(x)$ достигается в точке такой, что

$$\|x_1^*\| \leq c_3 \varepsilon^2.$$

где c_3 – положительная величина, не зависящая от ε при достаточно малом ε .

Рассмотрим точку \bar{x}_1 , лежащую в $L(x_0)$ на кратчайшем расстоянии от начала координат. Из (12) получаем

$$s = \|\bar{x}_1\| \leq c_2 \varepsilon^2,$$

откуда

$$\min_{x \in L(x_0)} f(x) \leq f(\bar{x}_1) \leq \frac{M\varepsilon^4 c_2^2}{2} = c_4 \varepsilon^4,$$

С другой стороны,

$$f(x_1^*) \geq \frac{m \|x_1^*\|}{2}.$$

Получаем

$$\frac{m \|x_1^*\|^2}{2} \leq c_4 \varepsilon^4, \quad \|x_1^*\| \leq \sqrt{\frac{2c_4}{m}} \varepsilon^2 = c_3 \varepsilon^2,$$

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

$$\|x_n\| = \|x_{p+1}\| \leq \|x_1^*\| + \|x_{p+1} - x_1^*\| \leq c_3 \varepsilon^2 + c_p \|x_1 - x_1^*\|, \quad (13)$$

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

$$\|x_1 - x_1^*\|^2 \leq \frac{M}{m} \varepsilon^2.$$

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.

2. Варианты алгоритмов без восстановления. Численные эксперименты

Ниже приводится вариант r -алгоритма без восстановления. Хотя строгие результаты о сходимости этих алгоритмов и 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 [4], [6]. 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 программой для вычислительной машины БЭСМ-6. Опишем кратко программу. Так как общая схема алгоритма описана, то остановимся лишь на алгоритме одномерного поиска минимума функции $f(x)$ в направлении η , с помощью которого осуществляется выбор шага h_{k+1} . Для реализации наискорейшего спуска с высокой точностью приходится производить большое количество вычислений значений функции. В разработанном алгоритме поиск минимума по направлению производится таким образом, что точность поиска минимума увеличивается в процессе счета.

Во всех приведенных числовых примерах одномерный поиск минимума производится следующим образом.

Выбираем числа $\mu \geq 1$, $0 < \gamma < 1$ и целое число L . Выбираем число h – начальный пробный шаг.

На $(k + 1)$ -ой итерации определены:

x_k – точка, полученная на k -ой итерации;

η – вектор, в направлении которого необходимо осуществить спуск из точки x_k ;

h – пробный шаг.

Из точки x_k делаем шаг длины h в направлении η : $z_1 = x_k + h \eta$. Вычисляем $f(z_1)$. Если $f(z_1) < f(x_k)$, то из точки z_1 снова делаем шаг длины h в направлении η : $z_2 = z_1 + h \eta$. Вычисляем $f(z_2)$. Если $f(z_2) < f(z_1)$, то из точки z_2 опять делаем шаг по η и т. д. 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

$\left(\mu \frac{l}{L} h\right)$. Если оказалось, что функция возрастает на первом шаге, то за новый шаг принимаем число (γh) , во всех остальных случаях пробный шаг h не изменяем.

При счете всех приведенных задач γ , μ и начальный шаг h принимались равными соответственно 0.1, 5/4, 0.1. Вышеописанная 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 k итераций вычисляется величина $d = \max_i |b_{ii}|$ и если $d < \varepsilon$, то матрицу B умножаем на число ρ . При счете $k = 10$, $\varepsilon = 1$, $\rho = 10^3$.

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 $R_\alpha(\eta)$.

Благодаря тому, что $R_\alpha(\eta)$ – матрица специального вида, операция $B \cdot R_\alpha(\eta)$ 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:

$f = f(x_1, \dots, x_n)$ – минимизируемая функция;

x^* – истинная точка минимума $f(x)$;

f^* – истинный минимум $f(x)$;

x_0 – начальная точка – точка на нулевой итерации;

\tilde{x}_N – точка, полученная в результате счета;

k – счетчик итерации;

f – значение функции;

α – коэффициент растяжения;

d – величина градиента в точке \tilde{x} : $d = \|\nabla f(\tilde{x})\|$;

$\delta = \max_i |\tilde{x}_i - x_i^*|$;

l – среднее количество вычислений значений функции на одной итерации.

Условием прекращения счета во всех примерах являлось выполнение

одного из следующих неравенств:

$$\|\Delta f(x_k) < 10^{-6}\|, \quad \|x_k - x_{k-1}\| < 10^{-7}, \quad \|B^* \nabla f(x_k)\| < 10^{-18}.$$

N – номер итерации, при которой произошел останов.

Пример 1 [5], [6].

$$\begin{aligned} f(x) &= (x_1^2 - x_2)^2 + (x_1 - 1)^2; \\ x^* &= (1; 1); \quad f^* = 0; \quad x_0 = (-1, 2; 1); \quad \text{при } \alpha, \text{ равном 2 и 3,} \\ \tilde{x}_N &= (1, 000000; 1, 000000). \end{aligned}$$

Пример 2 [4].

$$\begin{aligned} f(x) &= \sum_{i=1}^{10} (e^{-0,2i} + 2e^{-0,4i} - x_1 e^{-0,2x_2 i} - x_3 e^{-0,4x_4 i})^2; \\ x^* &= (1; 1; 2; 1); \quad f^* = 0; \quad x_0 = (0; 0; 0; 0); \\ \alpha &= 2; \quad \tilde{x}_{83} = (0, 999999; 1, 000000; 2, 000000; 1, 000000); \\ \alpha &= 3; \quad \tilde{x}_{90} = (0, 999999; 0, 999999; 2, 000001; 1, 000000). \end{aligned}$$

Пример 3 [4].

$$\begin{aligned} f(x) &= 10^{-3} \sum_{i=1}^{10} (10^3 e^{-0,2i} + 210^3 e^{-0,4i} - x_1 e^{-0,2x_2 i} - x_3 e^{-0,2x_4 i})^2; \\ x^* &= (1000; 1; 2000; 2); \quad f^* = 0; \quad x_0 = (500; 0; 2500; 3); \\ \alpha &= 2; \quad \tilde{x}_{100} = (1000, 000; 1, 000000; 2000, 000; 2, 000000); \\ \alpha &= 3; \quad \tilde{x}_{72} = (1000, 001; 1, 000000; 1999, 999; 2, 000000). \end{aligned}$$

Пример 4 [6].

$$\begin{aligned} f(x) &= 100 (x_1^2 - x_2) + (x_1 - 1)^2 + 90 (x_3^2 - x_4)^2 + (x_3 - 1)^2 + \\ &\quad + 10,1 \left((x_2 - 1)^2 + (x_4 - 1)^2 \right) + 19,8 (x_2 - 1) (x_4 - 1); \\ x^* &= (1; 1; 1; 1); \quad f^* = 0; \quad x_0 = (-3; -1; -3; -1); \\ \alpha &= 2; \quad \tilde{x}_{99} = (0, 999999; 0, 999999; 1, 000000; 1, 000000); \\ \alpha &= 3; \quad \tilde{x}_{76} = (1, 000000; 1, 000000; 1, 000000; 1, 000000). \end{aligned}$$

Пример 5 [6].

$$\begin{aligned} f(x) &= (e^{x_1} - x_2)^4 + 100 (x_2 - x_3)^6 + th^4 (x_3 - x_4) + x_1^8 + (x_4 - 1)^2; \\ x^* &= (0; 1; 1; 1); \quad f^* = 0; \quad x_0 = (1; 2; 2; 2); \end{aligned}$$

$$\alpha = 2; \quad \tilde{x}_{36} = (0, 0006; 1, 0005; 1, 0000; 1, 0000);$$

$$\alpha = 3; \quad \tilde{x}_{32} = (-0, 0004; 0, 9996; 1, 0000; 1, 0000).$$

Пример 6 [6].

$$f(x) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^2 + 10(x_1 - x_4)^4;$$

$$x^* = (0; 0; 0; 0); \quad f^* = 0; \quad x_0 = (10; 10; 10; -10);$$

$$\alpha = 2; \quad \tilde{x}_{50} = \left(4_{10^{-7}}; -2_{10^{-7}}; 6_{10^{-7}}; 7_{10^{-7}}\right);$$

$$\alpha = 3; \quad \tilde{x}_{46} = \left(3, 8_{10^{-5}}; 0, 4_{10^{-5}}; -4, 6_{10^{-5}}; -4, 6_{10^{-5}}\right).$$

Более подробная информация о результатах численных экспериментов содержится в таблице.

Приведенные выше данные о решении с помощью r -алгоритма ряда задач минимизации с ярко выраженными овражными особенностями говорят о том, что этот алгоритм является эффективным средством решения такого рода задач. По числу итераций, требуемых для достижения определенной точности, он близок к наиболее удачным модификациям алгоритмов типа Давидона–Флетчера–Пауэлла и метода сопряженных градиентов [6], [4]. Алгоритм устойчив по отношению к неточностям определения минимума по направлению (в алгоритме был использован заведомо грубый способ локализации минимума по направлению). Как показал опыт вычислений, коэффициент растяжения пространства α целесообразно выбирать между 2 и 3 (мы приводим лишь результаты экспериментов для α , равных 2 и 3, но опыты проводились и для множества других значений коэффициента растяжения), так как в большинстве случаев при дальнейшем увеличении α количество итераций если и уменьшается, то незначительно, но 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).

Были проведены также эксперименты, в которых счет производился 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.

α	I		II		III		IV		V		VI	
	$f(\alpha_k)$											
0	2.4 ₁₀ +1	2.4 ₁₀ +1	1.0 ₁₀ +1	5.4 ₁₀ +2	5.4 ₁₀ +2	1.9 ₁₀ +4	1.9 ₁₀ +4	2.3 ₁₀ +0	2.3 ₁₀ +0	1.6 ₁₀ +6	1.6 ₁₀ +6	1.6 ₁₀ +6
10	1.5 ₁₀ +0	2.7 ₁₀ -4	1.5 ₁₀ -3	6.7 ₁₀ -2	4.8 ₁₀ -2	4.0 ₁₀ +0	3.8 ₁₀ +0	1.5 ₁₀ -4	1.5 ₁₀ -4	6.3 ₁₀ +0	6.3 ₁₀ +0	1.5 ₁₀ +1
20	2.5 ₁₀ -1	7.7 ₁₀ -5	2.7 ₁₀ -4	1.9 ₁₀ -2	1.3 ₁₀ -2	2.4 ₁₀ +0	2.2 ₁₀ +0	1.3 ₁₀ -9	4.2 ₁₀ -9	2.2 ₁₀ -4	2.2 ₁₀ -4	4.8 ₁₀ -3
30	1.6 ₁₀ -3	3.5 ₁₀ -5	1.0 ₁₀ -4	1.5 ₁₀ -2	3.6 ₁₀ -3	1.3 ₁₀ +0	5.5 ₁₀ -1	4.0 ₁₀ -14	2.2 ₁₀ -13	2.3 ₁₀ -8	2.3 ₁₀ -8	1.6 ₁₀ -7
40	2.8 ₁₀ -7	8.2 ₁₀ -5	4.2 ₁₀ -5	5.6 ₁₀ -3	1.6 ₁₀ -4	4.6 ₁₀ -1	1.8 ₁₀ -1	-	2.2 ₁₀ -13	5.1 ₁₀ -11	5.1 ₁₀ -11	2.4 ₁₀ -11
50	1.9 ₁₀ -10	2.1 ₁₀ -7	1.5 ₁₀ -5	1.1 ₁₀ -3	1.3 ₁₀ -7	4.3 ₁₀ -3	4.5 ₁₀ -2	-	-	-	-	4.6 ₁₀ -13
60	3.4 ₁₀ -13	5.7 ₁₀ -11	2.0 ₁₀ -7	1.9 ₁₀ -5	2.7 ₁₀ -10	2.2 ₁₀ -4	4.0 ₁₀ -8	-	-	-	-	-
70	-	1.2 ₁₀ -12	7.1 ₁₀ -10	2.6 ₁₀ -8	5.3 ₁₀ -14	3.3 ₁₀ -6	3.2 ₁₀ -11	-	-	-	-	-
80	-	6.0 ₁₀ -15	3.9 ₁₀ -12	1.1 ₁₀ -10	-	3.0 ₁₀ -7	-	-	-	-	-	-
90	-	2.4 ₁₀ -17	2.4 ₁₀ -17	1.9 ₁₀ -13	-	3.2 ₁₀ -10	-	-	-	-	-	-
N	63	83	90	100	72	99	76	36	32	46	46	50
$f(\alpha_N)$	1.5 ₁₀ -14	8.5 ₁₀ -17	2.4 ₁₀ -15	9.0 ₁₀ -16	3.7 ₁₀ -14	5.7 ₁₀ -12	3.7 ₁₀ -12	5.4 ₁₀ -16	1.2 ₁₀ -14	3.4 ₁₀ -14	3.4 ₁₀ -14	4.6 ₁₀ -13
δ	10 ⁻⁷	10 ⁻⁷	10 ⁻⁶	10 ⁻⁴	10 ⁻³	10 ⁻⁶	10 ⁻⁶	10 ⁻³	10 ⁻³	10 ⁻⁶	10 ⁻⁶	10 ⁻⁴
d	5.4 ₁₀ -6	2.8 ₁₀ -8	3.5 ₁₀ -8	8.6 ₁₀ -7	4.5 ₁₀ -7	2.1 ₁₀ -5	4.0 ₁₀ -5	4.8 ₁₀ -8	2.5 ₁₀ -8	3.0 ₁₀ -6	3.0 ₁₀ -6	1.1 ₁₀ -5
l	4.3	4.8	4.7	4.3	5.7	4.4	6.9	4.1	3.8	10	10	13.9

^oРимскими цифрами в первой строке указаны номера примеров

3. Возможность применения r -алгоритмов к задачам минимизации негладких функций

Пусть мы имеем кусочно-гладкую выпуклую функцию вида

$$f(x) = \max_{1 \leq i \leq l} \varphi_i(x),$$

где $\varphi_i(x)$ – выпуклые гладкие функции, и пусть мы находимся в точке x_0 такой, что $\varphi_i(x_0) = f(x_0)$ при $i \in I^*$, где I^* – некоторое подмножество целочисленного интервала $[1, l]$, и в некоторой окрестности этой точки существует гладкая гиперповерхность

$$S^* = \left\{ x : f(x) = \varphi_i(x), \quad i \in I^* \right\}.$$

Эту гиперповерхность можно рассматривать, как "русло оврага" в том случае, когда

$$(g_{\varphi_i}(x), g_{\varphi_j}(x)) < 0, \quad i, j \in I^*, \quad i \neq j.$$

Из определения гиперповерхности S^* следует, что для любого касательного к этой гиперплоскости в точке x вектора $r(x)$ справедливо соотношение

$$\left(r(x), g_{\varphi_i}(x) - g_{\varphi_j}(x) \right) = 0, \quad i, j \in I^*.$$

Таким образом, растяжение пространства в направлении $g_{\varphi_i}(x) - g_{\varphi_j}(x)$ уменьшает составляющие обобщенных градиентов, направленные ортогонально "руслу оврага", и оставляет без изменения составляющие, направленные вдоль "русла оврага".

Эти эвристические соображения показывают, что последовательное применение операций растяжения пространства в направлении разности двух последовательных обобщенных градиентов при определенном способе выбора шага (когда мы идем до минимума в выбранном направлении и переходим его с пробным шагом, чтобы вычислить градиент по другую сторону от "русла оврага") приближенно соответствует операции проектирования градиента на "русло оврага".

Были проведены численные эксперименты, которые показывают, что применение операции растяжения пространства в направлении разности двух последовательных градиентов значительно ускоряет сходимость обобщенного градиентного спуска. По материалам исследования применимости r -алгоритмов к задачам минимизации негладких функций готовится отдельная статья.

Литература

1. ШОР Н.З. Использование операций растяжения пространства в задачах минимизации выпуклых функций // Кибернетика. – 1970. – № 1. – С. 6–12.
2. Н. З. ШОР, О скорости сходимости метода обобщенного градиентного спуска с растяжением пространства // Кибернетика. – 1970. – № 2. – С. 80–85.
3. ЛЮБИЧ Ю.И. Общая теория релаксационных процессов для выпуклых функционалов // Успехи математических наук. – 1970. – № 1.
4. ПОЛЯК Б.Т. Метод сопряженных градиентов, Труды II Зимней школы по математическому программированию и смежным вопросам (Дрогобыч), вып. 1, М., 1969.
5. РОЗЕНБРОК Х., СТОРИ С. Вычислительные методы для инженеров–химиков. – М.: "Мир", 1968.
6. HUANG H.I., LEVY A.V. Experiments on Quadratically Convergent Algorithms for Function Minimization // Journal of Optimization Theory and Applications. – 1970. – 3, v. 6.

Метод отсечения с растяжением пространства для решения задач выпуклого программирования

Н. З. Шор

Кибернетика. – 1977. – № 1. – С. 94–95.

В настоящей работе описан простой алгоритм решения задачи выпуклого программирования, гарантирующий уменьшение объема области, в которой локализуется оптимум, со скоростью геометрической прогрессии, причем знаменатель прогрессии зависит лишь от размерности задачи. Этот алгоритм относится к классу алгоритмов обобщенного градиентного спуска с растяжением пространства в направлении градиента (ОГСРП) [1]-[2]. С другой стороны, его можно рассматривать как метод отсечения, в котором операция растяжения пространства используется для симметризации области, в которой локализован оптимум.

Пусть мы имеем задачу выпуклого программирования:

$$\min f_0(x) \quad (1)$$

при ограничениях

$$f_i(x) \leq 0, \quad i = 1, \dots, m, \quad x \in E_n. \quad (2)$$

$f_\nu(x)$ – выпуклые функции, определенные на E_n ; $g_\nu(x)$ – соответствующие субградиенты, причем имеется априорная информация, что оптимальная точка x^* существует (она не обязательно единственная) и находится в шаре радиуса R с центром в точке x_0 (формально к системе ограничений (2) можно добавить ограничение $\|x - x_0\| \leq R$).

Рассмотрим следующий итеративный алгоритм (при $n > 1$).

Перед первым шагом имеем $x_0 \in E_n$, $B_0 = I$ – единичная матрица, $h_0 = \frac{R}{n+1}$.

Пусть проделано k шагов и мы получили $x_k \in E_n$; B_k – матрица $n \times n$, $h_k > 0$.

$(k+1)$ -й шаг. Вычисляем:

1)

$$g(x_k) = \begin{cases} g_0(x_k), & \text{если } \max_{1 \leq i \leq m} f_i(x_k) \leq 0, \\ g_{i^*}(x_k), & \text{если } \max_{1 \leq i \leq m} f_i(x_k) = f_{i^*}(x_k) > 0, \end{cases} \quad (3)$$

если $g(x_k) = 0$, то x_k – оптимальная точка; заметим, что $(g(x_k), x_k - x^*) \geq 0$;

2)

$$\xi_k = \frac{B_k^* g(x_k)}{\|B_k^* g(x_k)\|}; \quad (4)$$

3)

$$x_{k+1} = x_k - h_k \cdot B_k \cdot \xi_k; \quad (5)$$

4)

$$B_{k+1} = B_k \cdot R_\beta(\xi_k), \quad \beta = \sqrt{\frac{n-1}{n+1}}; \quad (6)$$

$R_\beta(\xi_k)$ – оператор растяжения пространства в направлении ξ_k с коэффициентом β [3];

5)

$$h_{k+1} = h_k \cdot r, \quad r = \frac{n}{\sqrt{n^2 - 1}}. \quad (7)$$

Лемма. Последовательность $\{x_k\}_{k=0}^\infty$ генерируемая алгоритмом (3)–(7), удовлетворяет неравенству

$$\|A_k(x_k - x^*)\| \leq h_k \cdot (n+1), \quad A_k = B_k^{-1}, \quad k = 0, 1, 2, \dots \quad (8)$$

Доказательство. Доказательство проведем индукцией по k . Для $k = 0$ неравенство (8) переходит в $\|x_0 - x^*\| \leq R$ и выполняется по предположению.

Пусть (8) выполняется для $k = \bar{k}$. Докажем его выполнение для $k = \bar{k} + 1$.

Пусть для краткости $A_k(x_k - x^*) = z_k$, $\alpha = \frac{1}{\beta}$,

$$\begin{aligned} \|z_{\bar{k}+1}\|^2 &= \|R_\alpha(\xi_{\bar{k}})(z_{\bar{k}} - h_{\bar{k}} \xi_{\bar{k}})\|^2 = \\ &= \left\| z_{\bar{k}} - h_{\bar{k}} \xi_{\bar{k}} + (\alpha - 1) \left[(z_{\bar{k}} - h_{\bar{k}} \xi_{\bar{k}}, \xi_{\bar{k}}) \xi_{\bar{k}} \right] \right\|^2 = \\ &= \left\| z_{\bar{k}} + [-\alpha h_{\bar{k}} + (\alpha - 1)(z_{\bar{k}}, \xi_{\bar{k}})] \xi_{\bar{k}} \right\|^2 = \\ &= \|z_{\bar{k}}\|^2 - 2h_{\bar{k}}(z_{\bar{k}}, \xi_{\bar{k}}) \alpha^2 + (\alpha^2 - 1)(z_{\bar{k}}, \xi_{\bar{k}})^2 + \alpha^2 h_{\bar{k}}^2 = \end{aligned}$$

$$= \|z_{\bar{k}}\|^2 - 2h_{\bar{k}}(z_{\bar{k}}, \xi_{\bar{k}}) \cdot \frac{n+1}{n-1} + \frac{2}{n-1}(z_{\bar{k}}, \xi_{\bar{k}})^2 + \frac{n+1}{n-1}h_{\bar{k}}^2. \quad (9)$$

Заметим, что

$$\begin{aligned} (z_{\bar{k}}, \xi_{\bar{k}}) &= \frac{1}{\|\tilde{g}(x_{\bar{k}})\|} \left(A_k(x_{\bar{k}} - x^*), B_{\bar{k}}^*g(x_k) \right) = \\ &= \frac{1}{\|\tilde{g}(x_{\bar{k}})\|} \left((x_{\bar{k}} - x^*), g(x_{\bar{k}}) \right) \geq 0. \end{aligned}$$

Здесь $\tilde{g}(x_{\bar{k}}) = B_{\bar{k}}^*g(x_{\bar{k}})$. Отсюда

$$\frac{2}{n-1}(z_{\bar{k}}, \xi_{\bar{k}})^2 - 2h_{\bar{k}} \frac{n+1}{n-1}(z_{\bar{k}}, \xi_{\bar{k}}) \leq \frac{2}{n-1} \left[(z_{\bar{k}}, \xi_{\bar{k}})^2 - (z_{\bar{k}}, \xi_{\bar{k}})^2 \right] = 0.$$

Таким образом,

$$\begin{aligned} \|z_{\bar{k}+1}\|^2 &\leq \|z_{\bar{k}}\|^2 + \frac{n+1}{n-1}h_{\bar{k}}^2 \leq (n+1)^2h_{\bar{k}}^2 + \frac{n+1}{n-1}h_{\bar{k}}^2 = \\ &= (n+1)^2 \cdot \frac{n^2}{n^2-1}h_{\bar{k}}^2 = (n+1)^2 \cdot h_{\bar{k}}^2, \end{aligned}$$

откуда следует справедливость (8).

Лемма доказана.

Множество точек x , удовлетворяющих неравенству

$$\|A_k(x_k - x)\| \leq (n+1)h_k = R \cdot \left(\frac{n}{\sqrt{n^2-1}} \right)^k,$$

представляет собой эллипсоид Φ_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

$$\frac{v(\Phi_{k+1})}{v(\Phi_k)} = \frac{\left(\frac{n}{\sqrt{n^2-1}} \right)^n \cdot \det A_k}{\det A_{k+1}} = \frac{\left(\frac{n}{\sqrt{n^2-1}} \right)^n \cdot \det A_k}{\det R_\alpha(\xi_k) \cdot \det A_k} =$$

$$= \frac{1}{\alpha} \left(\frac{n}{\sqrt{n^2 - 1}} \right)^n = \sqrt{\frac{n-1}{n+1}} \left(\frac{n}{\sqrt{n^2 - 1}} \right)^n = q_n < 1.$$

Таким образом, объем эллипсоида, в котором локализуется оптимальная точка x^* в соответствии с неравенством (8), убывает со скоростью геометрической прогрессии со знаменателем q_n .

К алгоритму (3)–(7) можно прийти, исходя из схемы последовательных отсечений¹. Рассмотрим, шар $S_0 = \{x : \|x - x_0\| \leq R\}$. Проведем гиперплоскость $(g(x_0), x - x_0)$. Область локализации x^* сужается до полусферы $\bar{S} = S_0 \cap \Pi_0$, где $\Pi_0 = \{x : (g(x_0), x - x_0) \leq 0\}$. Опишем вокруг \bar{S} эллипсоид минимального объема. Элементарные геометрические построения показывают, что центр этого эллипсоида будет равен

$$x_0 - \frac{R}{n+1} \cdot \frac{g(x_0)}{\|g(x_0)\|} = x_0 - \frac{R}{n+1} \xi_0,$$

т.е. будет совпадать с x_1 . Если теперь произвести растяжение пространства в направлении $\xi_0 = \frac{g(x_0)}{\|g(x_0)\|}$, то в растянутом пространстве образом указанного эллипсоида будет шар с радиусом $R \cdot \frac{n}{\sqrt{n^2 - 1}}$, с центром в точке $R_\alpha(\xi_0) x_1$. Теперь можно повторить процесс отсечения в растянутом пространстве и т. д. Легко увидеть, что при этом получается алгоритм вида (3)–(7).

Следует отметить, что геометрическая скорость убывания объема области, в которой локализован оптимум, для общего класса задач выпуклого программирования была показана, насколько известно автору, впервые для интересных в теоретическом отношении алгоритмов центрированных сечений и их модификаций, предложенных А. Левиным в [4].

Однако эти алгоритмы требуют для размерности $n > 2$ весьма громоздких процедур на каждом шаге, что делает их практически мало пригодными, по-видимому, уже для $n = 4$.

¹Такая схема была рассмотрена в [3].

Литература

1. Шор Н.З. Использование операции растяжения пространства в задачах минимизации выпуклых функций // Кибернетика. – № 1. – С. 6–12.
2. Шор Н.З. Обобщенные градиентные методы минимизации негладких функций и их применение к задачам математического программирования // Экономика и математические методы. – 1976. – № 12. – Вып. 2. – С. 337–356.
3. Юдин Д.Б., Немировский А.С. Информационная сложность и эффективные методы решения выпуклых экстремальных задач // Экономика и математические методы. – 1976. – № 12. – Вып. 2. – С. 357–369.
4. Левин А.Ю. Об одном алгоритме минимизации выпуклых функции // ДАН СССР. – 1965. – Т. 160. – № 6. – С. 1244–1247.

О методе решения одного класса динамических распределительных задач

Л. В. Беляева, Н. Г. Журбенко, Н. З. Шор

Экономика и матем. методы. – 1978. – Т. 14. – Вып. 1. – С. 137–146.

Обобщенные градиентные методы в сочетании с идеями декомпозиции являются эффективными при решении задач математического программирования большой размерности, особенно транспортно-производственного типа [1–5]. Разработанные в последние годы модификации обобщенных градиентных алгоритмов, использующие для ускорения сходимости операцию растяжения пространства ([6–8]), позволили существенно расширить их приложения. В данной статье показывается, как методы градиентного типа с растяжением пространства могут применяться для решения довольно сложных динамических распределительных задач. Самостоятельный интерес для специалистов в области исследования операций может представить описанная здесь экономико-математическая модель.

1. Рассматриваемая модель возникла из следующей практической задачи.

Имеются m поставщиков (предприятий) и n потребителей (строек). Планируемый период разбит на T интервалов (будем называть их кварталами).

Заданы: a_{it} – объем продукции, производимой предприятием i за квартал t ,

$t = 1, \dots, T$; a_i – объем продукции, выпускаемой предприятием i за планируемый период, $a_i = \sum_{t=1}^T a_{it}$; b_{jt} – объем спроса на продукцию потребителя j за квартал t ,

$t = 1, \dots, T$; b_j – потребность стройки j за планируемый период, $b_j =$

$\sum_{t=1}^T b_{jt}$; c_{ij} – стоимость перевозки единицы продукции от предприятия i к потребителю j ; 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 по кварталам, вообще говоря, не имеет места

$$\sum_{i=1}^m a_{it} \neq \sum_{j=1}^n b_{jt}.$$

На интуитивном уровне задача заключается в определении такого распределения продукции между потребителями, чтобы при этом выполнялись условия:

1) для большинства предприятий или, по крайней мере, для предприятий с большим приоритетом удовлетворена их потребность по кварталам;

2) объем преждевременно поставленной продукции должен быть по возможности меньшим (это условие можно интерпретировать как минимизацию затрат на хранение преждевременно произведенной продукции);

3) транспортные издержки минимальны по сравнению с другими распределениями, удовлетворяющими первым двум условиям;

4) потребители в течение всего планируемого периода должны по возможности прикрепляться к одному поставщику.

Цель нашей работы – в построении модели, отражающей интуитивное содержание задачи, и разработке метода ее решения.

2. Введем обозначения: β_{jt} – объем продукции, являющейся спросом потребителя j за первые t кварталов; $\beta_{jt} = \sum_{\tau=1}^t b_{j\tau}$, $t = 1, \dots, T$; α_{it} – объем продукции, производимой предприятием i за первые t кварталов; $\alpha_{it} = \sum_{\tau=1}^t a_{i\tau}$, $t = 1, \dots, T$; x_{ijt} – объем продукции, поставляемой предприятием i потребителю j за первые t кварталов.

Предположим, что все $\beta_{jt} \neq 0$, но это ограничение, как будет видно из дальнейшего, не является существенным и введено нами лишь для простоты изложения. Окончательные результаты справедливы для $\beta_{jt} \geq 0$.

Величину x_{ijt} / β_{jt} назовем степенью обеспеченности потребителя j предприятием i за первые t кварталов.

При построении модели дополним постановку задачи, не изменяя существенно ее экономического содержания, следующими условиями:

1) преждевременно произведенная продукция предприятия хранится на его складах, т. е. потребители не получают излишка продукции

$$\sum_{i=1}^m x_{ijt} \leq \beta_{jt}, \quad t = 1, \dots, T-1;$$

2) степень обеспеченности потребителя не убывает со временем

$$\frac{x_{ijt}}{\beta_{jt}} \geq \frac{x_{ijt-1}}{\beta_{jt-1}}, \quad t = 2, \dots, T. \quad (1)$$

Введем штрафные коэффициенты: R_{jt} – штраф за недопоставку единицы продукции потребителю j в конце квартала t ; e_{it} – штраф за наличие единицы нераспределенной продукции на предприятии i в конце квартала t , а ω_{it} – её объем.

Модель описывается следующей задачей линейного программирования

$$\min \sum_{j,i=1}^{n,m} c_{ij} x_{ijT} + \sum_{j,t=1}^{n,T-1} R_{jt} \left(\beta_{jt} - \sum_{i=1}^m x_{ijt} \right) + \sum_{i,t=1}^{m,T-1} e_{it} \omega_{it}, \quad (2)$$

$$\sum_{i=1}^m x_{ijT} = b_j, \quad j = 1, \dots, n, \quad (3)$$

$$\sum_{j=1}^n x_{ijT} = a_i, \quad i = 1, \dots, m, \quad (4)$$

$$\beta_{jt-1} x_{ijt} \geq \beta_{jt} x_{ijt-1}, \quad i = 1, \dots, m, \quad j = 1, \dots, n, \quad t = 2, \dots, T, \quad (5)$$

$$\sum_{j=1}^n x_{ijT} + \omega_{it} = \alpha_{it}, \quad i = 1, \dots, m, \quad t = 2, \dots, T-1, \quad (6)$$

$$\sum_{i=1}^m x_{ijt} \leq \beta_{jt}, \quad j = 1, \dots, n, \quad t = 1, \dots, T-1, \quad (7)$$

$$x_{ijt} \geq 0, \quad t = 1, \dots, T, \quad i = 1, \dots, m, \quad j = 1, \dots, n, \quad (8)$$

$$\omega_{it} \geq 0, \quad t = 1, \dots, T-1, \quad i = 1, \dots, m. \quad (9)$$

Ограничения (5), имеющие смысл и для $\beta_{jt} = 0$, соответствуют (1). Хотя неравенство $x_{ijt+1} \geq x_{ijt}$ не введено формально в список ограничений задачи, однако оно является следствием (5), (7) и (8), поскольку $\beta_{jt+1} \geq \beta_{jt}$.

Условие (5) запрещает прерывать обслуживание потребителя данным поставщиком, что "способствует" выполнению требования 4). Более точная формулировка 4) приводит к задаче целочисленного программирования, получение удовлетворительного решения которой было бы весьма проблематичным в случае большой размерности задачи.

Переходим к описанию метода решения задачи (2)–(9). Оно будет производиться в два этапа: на первом 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 задачи¹.

3. Прежде чем перейти к двойственной задаче, запишем (2)–(9) в эквивалентной форме

$$\min \sum_{j,i=1}^{n,m} c_{ij} x_{ijT} + \sum_{j,\tau=1}^{n,T-1} \left(\sum_{t=1}^{\tau} \beta_{jt}, R_{jt} \right) \left(\sum_{i=1}^m y_{ij\tau} \right) + \sum_{i,t=1}^{m,T-1} e_{it} \omega_{it}, \quad (10)$$

$$\sum_{i=1}^m x_{ijT} = b_j, \quad j = 1, \dots, n, \quad (11)$$

$$\sum_{j=1}^n x_{ijT} = a_i, \quad i = 1, \dots, m, \quad (12)$$

$$\sum_{j=1}^n \beta_{jt} \sum_{\tau=0}^{t-1} y_{ij\tau} + \omega_{it} = \alpha_{it}, \quad i = 1, \dots, m, \quad t = 1, \dots, T-1, \quad (13)$$

$$-\frac{1}{b_j} x_{ijT} + \sum_{\tau=0}^{T-1} y_{ij\tau} = 0, \quad i = 1, \dots, m, \quad j = 1, \dots, n, \quad (14)$$

$$x_{ijT} \geq 0, \quad i = 1, \dots, m, \quad j = 1, \dots, n, \quad (15)$$

¹Под схемой оптимального прикрепления мы понимаем выделение пар индексов (i, j) , для которых $x_{ijT} > 0$.

$$y_{ij\tau} \geq 0, \quad i = 1, \dots, m, \quad j = 1, \dots, n, \quad \tau = 0, \dots, T-1, \quad (16)$$

$$\omega_{it} \geq 0, \quad i = 1, \dots, m, \quad t = 1, \dots, T-1. \quad (17)$$

Нетрудно показать, что задача (10)–(17) эквивалентна (2)–(9), причем решение x_{ijt} для (2)–(9) определяется как

$$x_{ijt} = \beta_{jt} \sum_{\tau=0}^{t-1} y_{ij\tau}, \quad i = 1, \dots, m, \quad j = 1, \dots, n, \quad t = 1, \dots, T, \quad (18)$$

где $y_{ij\tau}$ – решение (10)–(17).

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 $r_{j\tau} = \sum_{t=1}^{\tau} \beta_{jt} R_{jt}$, $\tau = 1, \dots, T-1$, $j = 1, \dots, m$, и учитывая (18), получаем

$$\begin{aligned} & \sum_{\tau=1}^{T-1} \left(\sum_{t=1}^{\tau} \beta_{jt} R_{jt} \right) \left(\sum_{i=1}^m y_{ij\tau} \right) = \\ &= \sum_{\tau=1}^{T-1} r_{j\tau} \sum_{i=1}^m \left(\frac{x_{ij\tau+1}}{\beta_{j\tau+1}} - \frac{x_{ij\tau}}{\beta_{j\tau}} \right) = \\ &= \sum_{\tau=1}^{T-1} r_{j\tau} \left(\frac{\tilde{x}_{j\tau+1} - \beta_{j\tau+1}}{\beta_{j\tau+1}} - \frac{\tilde{x}_{j\tau} - \beta_{j\tau}}{\beta_{j\tau}} \right) = \\ &= \sum_{\tau=1}^{T-1} r_{j\tau} \left(\frac{\beta_{j\tau} - \tilde{x}_{j\tau}}{\beta_{j\tau}} \right) - \sum_{\tau=1}^{T-1} r_{j\tau} \left(\frac{\beta_{j\tau+1} - \tilde{x}_{j\tau+1}}{\beta_{j\tau+1}} \right) = \\ &= \sum_{\tau=1}^{T-1} r_{j\tau} \left(\frac{\beta_{j\tau} - \tilde{x}_{j\tau}}{\beta_{j\tau}} \right) - \sum_{\tau=1}^{T-1} (r_{j\tau+1} - \beta_{j\tau+1} R_{j\tau+1}) \left(\frac{\beta_{j\tau+1} - \tilde{x}_{j\tau+1}}{\beta_{j\tau+1}} \right) = \end{aligned}$$

$$= r_{j1} \frac{\beta_{j1} - \tilde{x}_{j1}}{\beta_{j1}} + \sum_{\tau=2}^{T-1} R_{j\tau} (\beta_{j\tau} - \tilde{x}_{j\tau}) = \sum_{\tau=1}^{T-1} R_{j\tau} (\beta_{j\tau} - \tilde{x}_{j\tau}).$$

Обозначим через v_j , u_{iT} , u_{it} , z_{ij} двойственные переменные, соответствующие ограничениям (11)–(14). Тогда задача, двойственная к (10)–(17), будет

$$\max \sum_{j=1}^m b_j v_j - \sum_{i,t=1}^{m,T} \alpha_{it} u_{it}, \quad (19)$$

$$v_j \leq c_{ij} + \frac{1}{b_j} z_{ij} + u_{iT}, \quad i = 1, \dots, m, \quad j = 1, \dots, n, \quad (20)$$

$$z_{ij} \leq \sum_{t=1}^{T-1} \beta_{jt} u_{it}, \quad i = 1, \dots, m, \quad j = 1, \dots, n, \quad (21)$$

$$z_{ij} \leq \sum_{t=1}^{\tau} \beta_{jt} R_{jt} + \sum_{t=\tau+1}^{T-1} \beta_{jt} u_{it}, \quad i = 1, \dots, m, \quad j = 1, \dots, n, \\ \tau = 1, \dots, T-2, \quad (22)$$

$$z_{ij} \leq \sum_{t=1}^{T-1} \beta_{jt} R_{jt}, \quad i = 1, \dots, m, \quad j = 1, \dots, n, \quad (23)$$

$$u_{it} \geq -e_{it}, \quad i = 1, \dots, m, \quad t = 1, \dots, T-1. \quad (24)$$

Пусть v_j^* , u_{iT}^* , z_{ij}^* – решение задачи (19)–(24). Тогда, очевидно, выполняется соотношение

$$v_j^* = \min_{i=1, \dots, m} \left[c_{ij} + \frac{1}{b_j} z_{ij}^* + u_{iT}^* \right]. \quad (25)$$

Обозначим через $I^*(j)$ множество индексов i , при которых достигается минимум в правой части формулы (25).

Поскольку переменные задачи (10)–(17) x_{ijT} являются двойственными для системы ограничений (20), то $x_{ijT} > 0$ лишь для $i \in I^*(j)$. Поэтому $I^*(j)$ – множество номеров предприятий, к которым прикрепляется потребитель j . Таким образом, для определения схемы оптимального прикрепления потребителей к пунктам производства достаточно решить двойственную задачу (19)–(24). Отметим, что в случае

$R_{jt} = 0$, $e_{it} = 0$, $i = 1, \dots, m$, $j = 1, \dots, n$, $t = 1, \dots, T - 1$, задаче (19)–(24) соответствует двойственная к классической транспортной задаче, определяемой параметрами c_{ij} , a_i , b_j .

4. Для решения задачи (19)–(24) применяем схему декомпозиции по множеству переменных. Зафиксировав переменные u_{it} , приходим к следующему

$$\max \sum_{j=1}^n b_j v_j, \quad (26)$$

$$v_j \leq c_{ij} + \frac{1}{b_j} z_{ij} + u_{i\tau}, \quad i = 1, \dots, m, \quad j = 1, \dots, n,$$

$$z_{ij} \leq r_{ij\tau}(u), \quad i = 1, \dots, m, \quad j = 1, \dots, n, \quad \tau = 0, \dots, T - 1,$$

где

$$r_{ij0}(u) = \sum_{t=1}^{T-1} \beta_{jt} u_{it},$$

$$r_{ij\tau}(u) = \sum_{t=1}^{\tau} \beta_{jt} R_{jt} + \sum_{t=\tau+1}^{T-1} \beta_{jt} u_{it}, \quad \tau = 1, \dots, T - 2,$$

$$r_{ij\tau-1}(u) = \sum_{t=1}^{T-1} \beta_{jt} R_{jt}.$$

Решение (26) определяется, очевидно, формулами

$$z_{ij}(u) = \min_{\tau=0, \dots, T-1} r_{ij\tau}(u), \quad (27)$$

$$v_j(u) = \min_{i=1, \dots, m} \left[c_{ij} + \frac{1}{b_j} z_{ij}(u) + u_{iT} \right]. \quad (28)$$

Таким образом, (19)–(24) сводится к задаче максимизации кусочно-линейной вогнутой функции $\tilde{F}(u)$

$$\tilde{F}(u) = \sum_{j=1}^n b_j v_j(u) - \sum_{i,t=1}^{m,T} \alpha_{it} u_{it} \quad (29)$$

при ограничениях

$$u_{it} \geq -e_{it}, \quad i = 1, \dots, m, \quad t = 1, \dots, T-1. \quad (30)$$

Рассмотрим задачу безусловной максимизации также кусочно-линейной вогнутой функции $F(u)$

$$\max_u F(u), \quad (31)$$

где

$$F(u) = \sum_{j=1}^n b_j v_j(u) - \sum_{i,t=1}^{m,T-1} \alpha_{it} \max\{u_{it}; -e_{it}\} - \sum_{i=1}^m \alpha_{iT} u_{iT},$$

а $v_j(u)$ – по прежнему определяется соотношениями (27)–(28).

Пусть \tilde{u}^* – решение (31). Тогда нетрудно видеть, что решение u^* задачи (29)–(30) находится как

$$u_{it}^* = \max\{\tilde{u}_{it}^*; -e_{it}\}, \quad t = 1, \dots, T-1, \quad i = 1, \dots, m,$$

$$u_{iT}^* = \tilde{u}_{iT}^*.$$

Итак, выявление схемы прикрепления потребителей к пунктам производства сводится к задаче безусловной максимизации (31) кусочно-линейной вогнутой функцией, для решения которой можно применить одну из модификаций метода оптимизации с растяжением пространства – r -алгоритм [8], опробованный на задачах минимизации кусочно-гладких функций [5,9].

5. Определение поставок x_{ijt} сводится к решению прямой задачи (10)–(17) по найденному решению двойственной – (19)–(24).

Опишем один общий метод получения решения прямой задачи, а затем покажем его применение к (2)–(9).

Пусть задана разрешимая задача выпуклого программирования, удовлетворяющая условиям теоремы Куна–Таккера

$$\begin{aligned} & \min_{x \in \Omega} f(x), \\ & g_i(x) \leq 0, \quad i = 1, \dots, m. \end{aligned} \quad (32)$$

Запишем двойственную задачу

$$\max_{\lambda_i \geq 0} \min_{x \in \Omega} L(x, \lambda), \quad (33)$$

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 ограниченности Ω .

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

$$g(x(\lambda)) = \left(g_1(x(\lambda)), \dots, g_m(x(\lambda)) \right)$$

является обобщенным градиентом $\Psi'(\lambda)$ функции $\Psi(\lambda)$ в точке λ .

Пусть λ^* – решение двойственной задачи, а x^* – прямой, соответствующее λ^* , т.е. выполняются соотношения $\lambda_i^* g_i(x^*) = 0$, $i = 1, \dots, m$. Из теоремы двойственности следует, что $\Psi(\lambda^*) = f(x^*)$.

Множество

$$M(\lambda^*) = \left\{ x(\lambda^*) \mid L(x(\lambda^*), \lambda^*) = f(x^*) \right\}. \quad (34)$$

Очевидно, $M(\lambda^*)$ – выпуклое замкнутое множество, и 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 $\min L(x, \lambda^*)$.

Рассмотрим случай, когда (32) 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 $\max \psi(\lambda)$ применим метод обобщенного градиентного спуска

$$\lambda^{k+1} = \lambda^k + h_k \Psi'(\lambda^k), \quad k \geq 1 \quad (35)$$

с регулировкой шага, обеспечивающей сходимость $\{\lambda^k\}$ к решению λ^* .

$$h_k \geq 0, \quad h_k \rightarrow 0, \quad \sum_{k=1}^{\infty} h_k = \infty. \quad (36)$$

Определим последовательность $\{x_N\}$

$$x_N = \sum_{k=1}^N \mu_k^N x^k. \quad (37)$$

Здесь и далее $x^k = x(\lambda^k)$, $\mu_k^N = h_k / \sum_{i=1}^N h_i$. Так как $\sum_{k=1}^N \mu_k^N = 1$, $\mu_k^N \geq 0$, а Ω – выпуклое множество, то $x_N \in \Omega$.

Покажем, что любая предельная точка \tilde{x} последовательности $\{x_N\}$ является решением прямой задачи (32).

Из замкнутости множества Ω следует, что $\tilde{x} \in \Omega$.

Установим, что $g(\tilde{x}) = 0$.

Поскольку $g_i(x)$ линейно, то

$$g(x_N) = \sum_{k=1}^N \mu_k^N g(x^k) = \sum_{k=1}^N \mu_k^N \Psi'(\lambda^k).$$

Из (35) вытекает

$$\sum_{k=1}^N \mu_k^N \Psi'(\lambda^k) = (\lambda^{N+1} - \lambda^1) / \sum_{i=1}^N h_i,$$

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 получаем требуемый результат.

Докажем, что $f(\tilde{x}) = f(x^*)$. Для задачи линейного программирования функция $\Psi(\lambda)$ является кусочно-линейной вогнутой, следовательно, существует такая окрестность $D(\lambda^*)$ точки λ^* , что для $\forall \lambda \in D(\lambda^*)$ будет выполняться $G(\lambda) \subset G(\lambda^*)$, где $G(\lambda)$ – множество обобщенных градиентов функции $\Psi(\lambda)$ в точке λ . 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 $\{\lambda^k\} \subset D(\lambda^{k*})$.

Выше мы указали, что $\Psi'(\lambda) = g(x(\lambda))$. Оказывается верно, в некотором смысле, и 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 (29). Отсюда, поскольку $g(x^k) = \Psi'(\lambda^k) \subset D(\lambda^*)$, следует, что $x^k \in M(\lambda^*)$. Но $M(\lambda^*)$ – выпуклое множество, поэтому $x_N \in M(\lambda^*)$, а из замкнутости $M(\lambda^*)$ получаем $\tilde{x} \in M(\lambda^*)$. Следовательно, $f(\tilde{x}) = f(x^*)$. Итак, $\tilde{x} \in \Omega$, $g(\tilde{x}) = 0$, $f(\tilde{x}) = f(x^*)$, т. е. x является решением (32), что и требовалось доказать.

Опишем схему применения изложенного метода получения решения прямой задачи применительно к паре (10)–(17) и (19)–(24).

Прежде всего, выделим множество потребителей, для которых x_{ijt} находится непосредственно из решения v^* , u^* , z^* задачи (19)–(24).

Пусть $I^*(j)$ – множество индексов i , для которых достигается минимум в (28) при $u = u^*$; $T_i^*(j)$ – множество индексов τ , для которых достигается минимум в (27) при $u = u^*$; I_1 – множество потребителей, для которых $I^*(j)$ состоит из единственного элемента, $i^*(j)$; W – множество пар индексов (i, j) , $i \in I^*(j)$, для которых $T_i^*(j)$ состоит лишь из одного элемента τ_{ij}^* .

Определим множество

$$I_2 = \left\{ j \mid j \in I_1, (i^*(j), j) \in W \right\}.$$

Нетрудно видеть, что для $j \in I_2$

$$x_{ijt} = \begin{cases} 0, & t \neq i^*(j), \\ 0, & t \leq \tau_{ij}^*, \quad i = i^*(j), \\ \beta_{jt}, & t > \tau_{ij}^*. \end{cases}$$

Таким образом, план поставок x_{ijt} для потребителей из I_2 определяется непосредственно из решения двойственной задачи (19)–(24). Для расчета определения плана поставок остальных потребителей необходимо решить задачу (10)–(17), в которой предварительно исключены переменные, соответствующие потребителям j из I_2 , т. е. величины a_i , α_{it} заменены на $\left(a_i - \sum_{j \in I_2} x_{ijT} \right)$ и $\left(\alpha_{it} - \sum_{j \in I_2} x_{ijt} \right)$, а индекс j принимает значения из $\tilde{I} = (1, \dots, n) \setminus I_2$.

Нетрудно получить следующую оценку для $\tilde{n} = \left| \tilde{I} \right|$: $\tilde{n} \leq mT$. Найдя схему оптимального прикрепления потребителей, приходим к задаче типа (10)–(17) сравнительно небольшого объема ($n = \tilde{n}$), для решения которой применим описанный выше метод получения решения прямой задачи.

Функцию Лагранжа для (10)–(17) будем строить относительно ограничений (12), (13), обозначая, как и раньше, соответствующие множители через u_{iT} и u_{it} . Для того чтобы множество Ω было ограничено, введем дополнительное неравенство $\omega_{it} \leq \alpha_{it}$, являющееся следствием системы всех ограничений задачи (10)–(17). Тогда двойственная задача будет

$$\max \Psi(u), \tag{38}$$

где

$$\Psi(u) = \min_{u \in \Omega} L(x, u).$$

Здесь для краткости через x обозначено множество переменных x_{ijT} , y_{ijT} , v_{it} ; Ω задается системой ограничений

$$\sum_{i=1}^m x_{ijT} = b_j, \quad -\frac{1}{b_j} x_{ijT} + \sum_{\tau=0}^{T-1} y_{ij\tau} = 0,$$

$$\omega_{it} \leq \alpha_{it}, \quad x_{ijT} \geq 0, \quad y_{ijT} \geq 0, \quad v_{it} \geq 0,$$

а функция Лагранжа $L(x, u)$ определена стандартным способом.

Нетрудно видеть, что решение $x(u)$ задачи $\min_{x \in \Omega} L(x, u)$ находится следующим образом

$$\omega_{it} = \begin{cases} \alpha_{it}, & u_{it} + e_{it} \leq 0, \\ 0, & u_{it} + e_{it} > 0, \end{cases}$$

$$y_{ij\tau} = \begin{cases} 1, & i = i^*(j), \quad \tau = \tau^*(j), \\ 0, & i \neq i^*(j) \quad \text{или} \quad \tau \neq \tau^*(j), \end{cases}$$

где $i^*(j)$, $\tau^*(j)$ — любая пара индексов (i, τ) , для которых достигается минимум в

$$\min_{(i, \tau)} r_{ij\tau}(u).$$

Здесь $r_{ij\tau}(u)$ определено так же, как в (27)

$$x_{ijT} = \begin{cases} b_j, & i = i^*(j), \\ 0, & i \neq i^*(j). \end{cases}$$

Таким образом, вычисление значений функции $\Psi(u)$, как и ее обобщенного градиента, не представляет трудностей. Поэтому для решения задачи (38) можно применить алгоритм обобщенного градиентного спуска и по (37) получить решение исходной задачи — (10)–(17).

Изложенная модель разработана для решения задачи планирования поставок металлоконструкций на строительные объекты Министерства энергетики и электрификации СССР ($n \approx 2000$, $m \approx 50$, $T = 4$).

В настоящее время производится экспериментальное исследование этого алгоритма. Работа выполнена в отделе экономической кибернетики Института кибернетики АН УССР. В постановке задачи принимали участие сотрудники отдела АСУ Одесского филиала НИИ "Оргэнергострой" Э. Д. Беккер и В. Я. Браверман.

В заключение отметим, что аналогичный метод решения может быть применен для динамических распределительных задач с взаимозаменяемыми продуктами и при отсутствии требования об обязательном удовлетворении спроса потребителей за весь период T .

Литература

1. ШОР Н.З. Применение метода градиентного спуска для решения сетевой транспортной задачи. В сб. Материалы научного семинара по теоретическим и прикладным вопросам кибернетики и исследования операций. Вып. 1. Киев, 1962 (НС по кибернетике АН УССР).
2. БАКАЕВ А.А., МИХАЛЕВИЧ В.С., БРАНОВИЦКАЯ С.В., ШОР Н.З. Методика и опыт решения сетевых транспортных задач большого объема на ЭЦВМ. В сб. Математические методы и проблемы производства, М., 1963.
3. ШОР Н.З., ГОРБАЧ Г.И. Решение задач распределительного типа методом обобщенного градиентного спуска. Тр. Семинара НС АН УССР по кибернетике "Теория оптимальных решений". Вып. 1, 1963.
4. ШОР Н.З., РОСИНА Н.И. Схема разложения задач линейного и выпуклого программирования и ее применение для решения задач планирования перевозок. Сб. докл. I Всес. конф. по оптимизации и моделированию транспортных сетей. Киев, 1967.
5. ШОР Н.З., ГАЛУСТОВА Л.А., СТРУТИНСКАЯ С.П., МОМОТ А.И. Оптимальное проектирование единой газоснабжающей системы. В сб. Применение математических методов в экономических исследованиях и планировании. Киев, 1972 (НС по кибернетике АН УССР).
6. ШОР Н.З. Использование операций растяжения пространства в задачах минимизации выпуклых функций // Кибернетика. – 1970. – № 1. – С. 6–12.

7. ШОР Н.З. О скорости сходимости метода обобщенного градиентного спуска с растяжением пространства // Кибернетика. – 1970. – № 2. – С. 80–85.
8. ШОР Н.З., ЖУРБЕНКО Н.Г. Метод минимизации, использующий операцию растяжения пространства в направлении двух последовательных градиентов // Кибернетика. – 1971. – № 3. – С. 51–59.
9. ШОР Н.З., ШАБАШОВА Л.П. О решении минимаксных задач методом обобщенного градиентного спуска с растяжением пространства // Кибернетика. – 1972. – № 1. – С. 82–88.

Использование методов негладкой оптимизации в задачах стохастического программирования¹

*Н. З. Шор, Т. А. Бардадым, Н. Г. Журбенко,
А. П. Лиховид, П. И. Стецюк*

Кибернетика и системный анализ. – 1999. – № . – С. 33–47.

Введение

В настоящей статье представлены результаты работ, выполненных по проекту "Методы решения и аппроксимации в стохастическом программировании" в рамках сотрудничества между Институтом исследования операций университета Цюриха и Институтом кибернетики им. В. М. Глушкова НАНУ. Основной целью проекта было создание специализированных программных модулей для решения задач линейного стохастического программирования [1] для системы моделирования SLP–IOR [2] – [5], разработанной в Институте исследования операций университета Цюриха.

В Институте кибернетики им. В.М.Глушкова накоплен значительный опыт решения различных классов задач стохастического программирования [6] – [17] и эффективного применения методов негладкой оптимизации [18] – [20], который использован при разработке новых модулей для системы SLP–IOR. Реализованные в этих модулях быстроходящиеся субградиентные методы можно применять не только для линейных, но и для нелинейных задач стохастического программирования.

Предложенный подход был реализован в виде двух программных модулей для решения задач стохастического программирования разных типов, описанию которых и посвящена данная статья. В разд. 1 приведены сведения о системе моделирования SLP–IOR и описан класс

¹Работа выполнена при поддержке Федерального управления иностранных дел и Национального научного фонда Швейцарии в рамках программы научного сотрудничества между государствами Западной Европы и новыми независимыми государствами.

моделей, на которые рассчитана данная версия системы; разд. 2 посвящен описанию метода декомпозиции по переменным и краткой схемы r -алгоритма; в разд. 3 описана схема решения простейших двухэтапных задач, реализованная в новом модуле; разд. 4 содержит описание подхода, использованного в блоке решения задач с фиксированной рекурсией; в разд. 5 представлены результаты тестирования.

1. Система моделирования SLP–IOR

Возникновение систем моделирования как отдельных программных комплексов было вызвано необходимостью упростить и унифицировать подготовку данных и отчетов при моделировании, что облегчает процесс программирования и делает систему более понятной для специалистов других областей – экономистов, социологов и т.п.

Система SLP–IOR разработана как система моделирования для специалистов в области исследования операций. Для работы с системой пользователь должен быть знаком с основами исследования операций, включая стохастическое программирование и, кроме того, иметь представление о методах, реализованных в том или ином блоке решения задач, чтобы знать, на какие классы задач они рассчитаны и где наиболее эффективны. Выбор конкретного программного модуля зависит не только от типа модели, но и от ее размерности и доступных ресурсов компьютера.

Подробный перечень методов стохастического программирования, использованных при создании системы SLP–IOR, можно найти в [2]; полезными окажутся также работы [1, 6, 21, 22].

В системах моделирования обычно можно выделить часть, отвечающую за подготовку алгебраического представления модели и часть, связанную с решением задач конкретного типа. Авторы системы SLP–IOR в качестве средства алгебраического представления модели использовали пакет GAMS 2.25 (Общая система алгебраического моделирования) [23], уделив основное внимание разработке средств решения линейных задач стохастического программирования.

Подробную информацию о системе моделирования SLP–IOR можно найти в [2] – [5].

Модели SLP–IOR. Данная версия системы SLP–IOR предназначена для решения двухэтапных задач линейного стохастического программирования и задач с ограничениями вероятностного характера. Двух-

этапные модели имеют вид:

$$\min_x \{c^T x + E_\omega Q(x, \omega)\} \quad (1)$$

при следующих ограничениях: первый этап –

$$Ax \quad .\mathbf{R.} \quad b, \quad (2)$$

$$x \in [l, u], \quad (3)$$

второй этап –

$$Q(x, \omega) = \min q^T(\omega)y \quad (4)$$

при ограничениях

$$W(\omega)y \quad .\mathbf{R.} \quad h(\omega) - T(\omega)x, \quad (5)$$

$$y \geq 0. \quad (6)$$

Здесь $.\mathbf{R.}$ – любое из соотношений $\leq, \geq, =$; x, y – переменные задач первого и второго этапа соответственно; E – символ математического ожидания; q – вектор цен второго этапа (коэффициенты рекурсии); c – вектор цен первого этапа; $W(\omega)$ – матрица рекурсии; $A, T(\omega)$ – матрицы правых частей задач первого и второго этапов; b – правая часть ограничений первого этапа; $h(\omega)$ – случайная правая часть ограничений второго этапа; $\omega \in \Omega$ – элемент вероятностного пространства Ω .

Если $W(\omega) = W$, задача (1)–(6) в терминологии [1] называется задачей с фиксированной рекурсией. Если, кроме того, для $\forall z$

$$\{y | Wy \quad .\mathbf{R.} \quad z, \quad y \geq 0\} \neq \emptyset,$$

задача (1)–(6) называется задачей с полной рекурсией, т. е. ограничения (5) допустимы для любых x и ω . В частном случае, когда матрица рекурсии имеет вид $W = (I, -I)$, I – единичная матрица, а $.\mathbf{R.}$ в (5) означает только знак равенства, задача (1)–(6) называется задачей с простой рекурсией.

Модели с общим вероятностным ограничением имеют вид

$$\min_x c^T x$$

при ограничении

$$P(\{\omega | Tx \geq (\leq) h(\omega)\}) \geq \alpha$$

$$Ax \quad .\mathbf{R.} \quad b,$$

$$x \in [l, u],$$

где $0 \leq \alpha \leq 1$ – заданный уровень вероятности. Введенные здесь обозначения указывают, что любое из неравенств \leq, \geq может использоваться построчно.

Модели с разделенными вероятностными ограничениями имеют вид

$$\min_x c^T x$$

при ограничениях

$$P(\{\omega | t_i^T T x \geq (\leq) h_i(\omega)\}) \geq \alpha_i \quad \forall i,$$

$$Ax \leq b,$$

$$x \in [l, u],$$

где $0 \leq \alpha_i \leq 1$ – заданные уровни вероятности для отдельных ограничений, а $t_i(\omega)$ – i -ая строка матрицы $T(\omega) \forall i$.

Данная версия системы SLP–IOR содержит средства решения описанных выше типов задач линейного стохастического программирования.

2. Схема декомпозиции по переменным и r-алгоритм

При разработке программных модулей для системы SLP–IOR были использованы методы декомпозиции, обычно применяемые для решения задач линейного и выпуклого программирования большой размерности. В математическом программировании главным образом рассматриваются две схемы декомпозиции: по ограничениям и по переменным. Другие, более сложные схемы декомпозиции как правило можно интерпретировать как суперпозицию этих основных схем.

Приведем, согласно [19], схему декомпозиции по переменным, которая применялась при создании программных модулей к системе SLP–IOR.

1. Рассмотрим задачу выпуклого программирования, переменные в которой разбиты на два подмножества. Требуется найти

$$\inf_{x,y} f_0(x, y), \tag{7}$$

при ограничениях

$$f_i(x, y) \leq 0, \quad i = 1, \dots, n, \tag{8}$$

где x и y - векторы переменных:

$$x = (x^{(1)}, \dots, x^{(l)}), \quad y = (y^{(1)}, \dots, y^{(m)}),$$

а f_0 и f_i , $i = 1, \dots, n$, - выпуклые функции $(l + m)$ -мерного вектора $z = (x, y) \in E^l \times E^m$. Зафиксируем $x = \bar{x}$ и рассмотрим следующую задачу: найти

$$\inf_y f_0(\bar{x}, y) \quad (9)$$

при ограничениях

$$f_i(\bar{x}, y) \leq 0, \quad i = 1, \dots, n. \quad (10)$$

Задача (9), (10) является задачей выпуклого программирования. Для значений \bar{x} , для которых решение задачи (9), (10) существует, определим функцию Φ следующим образом:

$$\Phi(\bar{x}) = \inf_{y \in D(\bar{x})} f_0(\bar{x}, y), \quad (11)$$

где $D(\bar{x})$ - множество всех y , удовлетворяющих (10). Доказательство следующей теоремы можно найти, например, в [19].

Теорема. Если f_i , $i = 0, 1, \dots, n$ - выпуклые функции и задача (7), (8) имеет оптимальное решение $z^* = (x^*, y^*)$, то функция Φ , определенная в (11), является выпуклой (конечно определенной на некотором выпуклом подмножестве W в E^l). Если для некоторого $\bar{x} \in \text{int}W$ для (10) выполняется условие Слейтера, то субградиент Φ в точке $x = \bar{x}$ можно вычислить по формуле

$$g_\Phi(\bar{x}) = g_{L_u^x}(\bar{x}, y(\bar{x})), \quad (12)$$

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 задачи (9)-(10), $u = \{u_i\}$ - множители Лагранжа для (9), (10), полученные по теореме Куна-Таккера, а $g_{L_u^x}(\bar{x}, y(\bar{x}))$ - это проекция субградиента функции $L_u(z) = L_u(x, y)$ на подпространство $E_{(x)}^l$, имеющее нулевую проекцию на подпространство $E_{(y)}^m$ (субградиент берется в точке $z = (\bar{x}, y(\bar{x}))$).

Следствие. Если в этой теореме дополнительно предположить, что $f_i(x, y)$, $i = 0, 1, \dots, n$, непрерывно дифференцируемы по y , то формулу (12) можно упростить:

$$g_\Phi(\bar{x}) = g_{f_0}^x(\bar{x}, y(\bar{x})) + \sum_{i=1}^n u_i(\bar{x}) g_{f_i}^x(\bar{x}, y(\bar{x})), \quad (13)$$

где $g_{f_i}^x(\bar{x}, y(\bar{x}))$ – проекции произвольных субградиентов f в $\bar{z} = (\bar{x}, y(\bar{x}))$ на $E_{(x)}^l$, $i = 0, 1, \dots, n$.

Формула (13) позволяет строить алгоритмы для решения задач выпуклого программирования (в частности, линейного программирования) с помощью декомпозиции по переменным. До сих пор предполагалось, что задачу (9), (10) можно решить за конечное число шагов, получить множители Лагранжа $u = \{u_i\}$, и вычислить "частные" субградиенты $g_{f_i}^x$, $i = 0, \dots, n$. Это предположение выполняется, например, если (9), (10) – задача линейного или квадратичного программирования. Дополнительные трудности возникают, когда задача (9), (10) не имеет решений при некоторых \bar{x} . Обычный способ преодолеть эту трудность заключается в применении штрафных функций, при этом задача (7)–(8) заменяется следующей: найти

$$\min_{x, y, v} f_0(x, y) + M \sum_{i=1}^n v_i$$

при ограничениях

$$f_i(x, y) - v_i \leq 0, \quad v_i \geq 0,$$

где M – достаточно большое положительное число. Задача (9), (10) сводится к задаче нахождения

$$\min_{y, v} \left[f_0(\bar{x}, y) + M \sum_{i=1}^n v_i \right] \quad (14)$$

при ограничениях

$$f_i(\bar{x}, y) - v_i \leq 0, \quad v_i \geq 0, \quad i = 1, \dots, n. \quad (15)$$

Очевидно, что для любого \bar{x} задача (14), (15) имеет непустое допустимое множество, которое удовлетворяет условию Слейтера.

Таким образом, мы можем полагать, что задача (7), (8) приведена к виду, в котором функция $\Phi(\bar{x})$ определена для всех \bar{x} , и можно вычислить субградиент $g_\Phi(\bar{x})$ в любой точке \bar{x} , т. е. для минимизации Φ можно использовать субградиентный алгоритм. Таким образом, получена схема решения задачи (7), (8), где на k -й итерации выполняются такие шаги:

- 1) для ранее вычисленного $\bar{x} = \bar{x}_k$ решается задача (9)–(10) и находятся множители Лагранжа $u(\bar{x}_k) = \{u_i(\bar{x}_k)\}$,

- 2) вычисляется вектор $g_{\Phi}(\bar{x}_k)$ согласно (12), или, если это возможно, согласно (13),
- 3) вычисляется \bar{x}_{k+1} с помощью обычного субградиентного процесса или другого субградиентного метода (например, r -алгоритма).

Опишем более подробно вычисление субградиента в частном случае, когда (7)–(8) является задачей линейного программирования. Предположим, нужно найти

$$\max [(c, x) + (d, y)]$$

при ограничениях

$$Ax + By \leq e, \quad x \geq 0, \quad y \geq 0,$$

где x и y – векторные переменные размерности l и m , соответственно, A – матрица размера $n \times l$, B – матрица размера $n \times m$; c, d и e – фиксированные векторы размерности l, m и n , соответственно. Зафиксируем $x = \bar{x}$ и рассмотрим задачу: найти

$$\max_y [(c, \bar{x}) + (d, y)]$$

при ограничениях

$$By \leq e - A\bar{x}, \quad y \geq 0,$$

и двойственную ей: найти

$$\min_v (e - A\bar{x}, v), \tag{16}$$

при ограничениях

$$B^T v \geq d, \tag{17}$$

$$v \geq 0, \tag{18}$$

где v – вектор переменных размерности n , B^T – транспонированная матрица B .

Пусть $\bar{v}(\bar{x})$ – решение (16)–(18), а α_j – j -я строка матрицы A ($j = 1, \dots, l$).

Тогда j -й элемент вектора $g(\bar{x})$ размерности l определяется следующим образом:

$$g_j(\bar{x}) = (-\alpha_j, \bar{v}(\bar{x})) + c_j, \quad g(\bar{x}) = (g_1(\bar{x}), \dots, g_l(\bar{x})),$$

где c_j – j -й элемент вектора c .

2. r -алгоритм. Для минимизации функции $\Phi(\bar{x})$ в (11) требуется эффективный по числу вычислений функции и ее градиента метод оптимизации негладких функций. Такому требованию удовлетворяет r -алгоритм [18]–[20]. Приведем краткое описание r -алгоритма, применительно к решению задачи минимизации выпуклой функции $f(x)$, определенной на n -мерном пространстве E^n и обладающей свойством $f(x) \rightarrow +\infty$ при $\|x\| \rightarrow \infty$.

При описании алгоритма будем использовать оператор $R_\alpha(\xi)$ растяжения (сжатия) пространства с коэффициентом $\alpha \geq 1$ ($0 < \alpha < 1$) в направлении ξ , $\|\xi\| = 1$, т.е.

$$R_\alpha(\xi)x = x + (\alpha - 1)(x, \xi)\xi.$$

Обозначим $g_f(x)$ субградиент функции f в точке x .

Рассмотрим общую схему r -алгоритма.

0-й шаг. Заданы начальная точка $x_0 \in E^n$ и невырожденная матрица B_0 размера $n \times n$ (например, $B_0 = I_n$ (единичная матрица)).

1-й шаг. Вычисляем:

$$g_0 = g_f(x_0); \tilde{g}_0 = B_0^T g_0; \xi_0 = \frac{\tilde{g}_0}{\|\tilde{g}_0\|};$$

$$x_1 = x_0 - h_0 B_0 \xi_0 \quad (h_0 > 0 - \text{шаговый множитель}).$$

Запоминаем x_1 , $g_f(x_0)$, B_0 . После k шагов имеем: $x_k \in E^n$; $g_f(x_{k-1}) \in E^n$; матрицу B_{k-1} размера $n \times n$.

($k+1$)-й шаг. Вычисляем:

$$1) \quad g_f(x_k); d_k = g_f(x_k) - g_f(x_{k-1});$$

$$2) \quad r_k = B_k^T d_k; \xi_k = \frac{r_k}{\|r_k\|};$$

$$3) \quad B_k = B_{k-1} R_{\beta_k}(\xi_k), \quad 0 < \beta_k < 1;$$

$$4) \quad x_{k+1} = x_k - h_k B_k p_k,$$

где $p_k = \frac{\tilde{g}_k}{\|\tilde{g}_k\|}$; $\tilde{g}_k = B_k^T g_f(x_k)$, $h_k > 0$ – k -й шаговый множитель.

Запоминаем x_{k+1} , $g_f(x_k)$, B_k . Если выполнен критерий останова – процедура завершается; в противном случае переходим к следующему шагу.

Комментарий. Пусть A_{k-1} – получившаяся матрица преобразования пространства после k шагов r -алгоритма: $y = A_{k-1}x$, или $x = B_{k-1}y$, где $B_{k-1} = A_{k-1}^{-1}$, и $\varphi_k(y) = f(B_{k-1}y)$. Поскольку $g_{\varphi_k}(y) = B_{k-1}^T g_f(x)$, r_k является разностью двух субградиентов функции φ_k , взятых в точках $y_k = A_{k-1}x_k$ и $\tilde{y}_{k-1} = A_{k-1}x_{k-1}$. Таким образом, ξ_k – нормированное направление разности двух последовательных субградиентов преобразованной функции $\varphi_k(y)$. В этом направлении производится растяжение преобразованного пространства и получается результирующая матрица $A_k = R_{\alpha_k}(\xi_k)A_{k-1}$. Обратная матрица $B_k = B_{k-1}R_{\beta_k}(\xi_k)$, $\beta_k = 1/\alpha_k$. Рассмотрим $\varphi_{k+1}(y) = f(B_k y)$ и выполним шаг по субградиенту для функции $\varphi_{k+1}(y)$ из $\tilde{y}_k = B_k x_k$:

$$y_{k+1} = \tilde{y}_k - h_k \frac{B_k^T g_f(x_k)}{\|B_k^T g_f(x_k)\|}.$$

В исходном пространстве точке y_{k+1} соответствует точка $x_{k+1} = B_k y_{k+1}$, поэтому

$$x_{k+1} = x_k - h_k B_k \frac{B_k^T g_f(x_k)}{\|B_k^T g_f(x_k)\|}.$$

Таким образом, вычислены x_{k+1} , $g_f(x_k)$, B_k и можно переходить к следующей итерации.

Семейство r -алгоритмов имеет две последовательности параметров: $\{\beta_k\}_{k=1}^{\infty}$ и $\{h_k\}_{k=1}^{\infty}$. Чтобы получить “хорошую” сходимость к точке оптимума $f(x)$, при минимизации негладких выпуклых функций рекомендуется использовать постоянный коэффициент растяжения пространства $\alpha_k = \alpha$, $\alpha \in [2, 4]$. Выбор значений шагового множителя h_k определяется специальной адаптивной процедурой, обеспечивающей сравнительно небольшое количество вычислений значений функции при одномерном спуске (см. [18] – [20]).

3. Программы для решения двухэтапных задач с простой рекурсией

Напомним, что в задачах с простой рекурсией в постановке (1)–(6) матрица рекурсии имеет вид $W = (I, -I)$, где I – единичная матрица, а **.R.** в (2) означает только знак равенства.

Предполагается, что случайные величины в задаче являются независимыми и характеризуются конечным дискретным эмпирическим распределением, причем стохастическая только правая часть.

Рассматривается двухэтапная задача стохастического программирования: найти

$$\min \left\{ F(x) + c^T x + \sum_{w \in \Omega} p(w) [q^{+T} y^+(x, w) + q^{-T} y^-(x, w)] \right\} \quad (19)$$

при следующих ограничениях: первый этап –

$$f^I(x) + Ax \leq b, \quad (20)$$

$$l \leq x \leq u, \quad (21)$$

второй этап –

$$y^+ - y^- = h(w) - (f^{II}(x) + Tx), \quad (22)$$

$$y^+, y^- \geq 0, \quad (23)$$

где $F(x)$ – нелинейная выпуклая функция по переменным x ; $f^I(x)$, $f^{II}(x)$ – векторы нелинейных выпуклых функций по переменным x для ограничений первого и второго этапов соответственно; q^+ , q^- – векторы коэффициентов затрат на коррекцию; y^+ , y^- – векторы коррекции; $p(w)$ – вероятность случайного события w .

Метод решения данной задачи основывается на применении схемы декомпозиции по переменным (разд. 2). Для учета ограничений первого этапа используется метод негладких штрафных функций.

Преобразуем исходную задачу (19)–(23) в следующую:
найти

$$\begin{aligned} \min_{x, y^+, y^-} \{ & F(x) + c^T x + \rho_1 \max\{0, f^I(x) + Ax - b\} + \\ & + \rho_2 \max\{0, x - u, l - x\} + \sum_{w \in \Omega} p(w) [q^{+T} y^+(x, w) + q^{-T} y^-(x, w)] \} \end{aligned} \quad (24)$$

при ограничениях

$$y^+ - y^- = h(w) - (f^{II}(x) + Tx), y \geq 0, w \in \Omega.$$

Здесь ρ_1, ρ_2 – значения штрафных коэффициентов.

Следуя схеме декомпозиции по переменным зафиксируем значения переменных первого этапа $x = \bar{x}$ и рассмотрим следующую задачу:

найти

$$\min_{y^+, y^-} \left\{ \sum_{w \in \Omega} p(w) [q^{+T} y^+(\bar{x}, w) + q^{-T} y^-(\bar{x}, w)] \right\} \quad (25)$$

при ограничениях:

$$y^+(\bar{x}, w) - y^-(\bar{x}, w) = h(w) - (f^{II}(\bar{x}) + T\bar{x}), \quad (26)$$

$$y^+(\bar{x}, w), y^-(\bar{x}, w) \geq 0, w \in \Omega. \quad (27)$$

Решение задачи (25)–(27) может быть получено путем следующих простых вычислений.

Пусть k – индекс ограничения (26). Определим

$$\Delta_k = h_k(w) - (f_k^{II}(\bar{x}) + t_k\bar{x}), \quad (28)$$

где t_k – k -я строка матрицы T . Тогда переменные коррекции определяются соотношениями:

$$y_k^+(\bar{x}, w) = \begin{cases} \Delta_k & , \text{ если } \Delta_k \geq 0, \\ 0 & , \text{ если } \Delta_k < 0; \end{cases} \quad (29)$$

$$y_k^-(\bar{x}, w) = \begin{cases} -\Delta_k & , \text{ если } \Delta_k < 0, \\ 0 & , \text{ если } \Delta_k \geq 0. \end{cases} \quad (30)$$

Двойственные переменные определяются соотношениями:

$$u_k(\bar{x}, w) = \begin{cases} -p(w)q_k^+ & , \text{ если } \Delta_k \geq 0, \\ p(w)q_k^- & , \text{ если } \Delta_k < 0. \end{cases} \quad (31)$$

Введем обозначения:

$$\begin{aligned} \Psi(\bar{x}, y) &= F(\bar{x}) + c^T\bar{x} + \rho_1 \max\{0, f^I(\bar{x}) + A\bar{x} - b\} + \\ &+ \rho_2 \max\{0, \bar{x} - u, l - \bar{x}\} + \\ &+ \sum_{w \in \Omega} p(w)[q^{+T} y^+(\bar{x}, w) + q^{-T} y^-(\bar{x}, w)], \end{aligned} \quad (32)$$

$$\begin{aligned} \psi_k(\bar{x}, y(\bar{x}), w) &= \\ &= (f_k^{II}(\bar{x}) + t_k\bar{x}) + y_k^+(\bar{x}, w) - y_k^-(\bar{x}, w) - h_k(w). \end{aligned} \quad (33)$$

Следуя схеме декомпозиции по переменным, определим функцию координирующей задачи Φ следующим образом:

$$\Phi(\bar{x}) = \inf_{y \in D(\bar{x})} \Psi(\bar{x}, y),$$

где $D(\bar{x})$ – множество всех y , удовлетворяющих (26)–(27).

Субградиент функции Φ в точке $x = \bar{x}$ может быть вычислен по формуле:

$$g_{\Phi}(\bar{x}) = g_{\Psi}^x(\bar{x}, y(\bar{x})) + \sum_k \sum_{w \in \Omega} (u_k(\bar{x}, w) g_{\psi_k}^x(\bar{x}, y(\bar{x}))), \quad (34)$$

где $g_f^x(\cdot)$ — "частные" субградиенты по x :

$$g_{\Psi}^x(\bar{x}, y(\bar{x})) = g_F(\bar{x}) + c + \rho_1 \max\{0, [g_{f_{m^*}}^x(\bar{x}) + a_{m^*}]\}_-^+ \rho_2 \max\{0, e_{j^*}\},$$

$$g_{\psi_k}^x(\bar{x}, y(\bar{x}), w) = g_{f_k}^x(\bar{x}) + t_k.$$

Здесь m^* — индекс ограничения первого этапа с максимальной невязкой; a_{m^*} — m^* -я строка матрицы A ; e_{j^*} — орт пространства переменных первого этапа: $e_{j^*} = \{0, 0, \dots, 1, \dots, 0\}$ (единица на j^* -м месте); j^* — индекс переменной первого этапа с максимальной невязкой в ограничениях (21) (знаки "+", "-" соответствуют правой или левой части неравенства).

В данном случае для вычисления $g_{\Phi}(\bar{x})$ не нужно определять величины $y^+(\bar{x}), y^-(\bar{x})$. Необходимо только для каждого ограничения второго этапа k найти знаки двойственных переменных $u_k(\bar{x}, w)$. В алгоритме предусмотрено упорядочение случайных реализаций правой части для каждого ограничения второго этапа, что дает возможность существенно уменьшить общее время нахождения решения.

Алгоритм решения задачи (19)–(23) сводится к следующему:

s-я итерация:

- 1) для полученных на предыдущем шаге значений $\bar{x} = \bar{x}_s$ находим знаки множителей Лагранжа $u_k(\bar{x}_s, w)$ для каждого ограничения второго этапа;
- 2) вычисляем вектор $g_{\Phi}(\bar{x}_s)$ по формуле (34);
- 3) вычисляем \bar{x}_{s+1} , выполняя шаг алгоритма субградиентного типа.

Программная реализация этого подхода была написана на языке C++ для среды MS DOS и IBM PC/AT - совместимых компьютеров. В качестве алгоритма решения координирующей задачи использовался r -алгоритм [18]–[20]. Проведенные численные эксперименты (разд. 5) показали практическую эффективность предлагаемого подхода.

4. Программный модуль для решения линейных двухэтапных задач с фиксированной матрицей W

Разработанный программный модуль предназначен для решения двухэтапных линейных задач стохастического программирования с фиксированной матрицей W (в терминологии [1] – задач с фиксированной рекурсией) с помощью схемы декомпозиции, в которой используется r -алгоритм [18]–[20].

Данная версия программы применима для решения двухэтапных задач стохастического программирования следующего класса: матрица W фиксирована; случайными величинами являются значения правых частей задачи второго этапа; случайные величины имеют конечное дискретное распределение; задачи второго этапа являются задачами линейного программирования.

Математическая модель записывается следующим образом:

$$\min_x \{c^T x + F(x)\}$$

при ограничениях: первый этап –

$$Ax \leq b, \quad (35)$$

$$x \geq 0, \quad (36)$$

где

$$F(x) = EQ(x, \omega) = \sum_j p_j Q_j(x);$$

второй этап –

$$Q_j(x) = \min q^T y$$

при ограничениях

$$Wy \leq h(\omega_j) - Tx, \quad y \geq 0,$$

где p_j – вероятность реализации $w = w_j$.

Исходная задача заменяется следующей:

$$\min_z f(z), \quad (37)$$

где

$$f(z) = f_1(z) + f_2(z), \quad (38)$$

$$f_1(z) = \min_x [c^T x + R_x^+ x^+ + R_x^- x^-], \quad (39)$$

$$Ax \quad .\mathbf{R}. \quad b, \quad (40)$$

$$Tx + x^+ - x^- = z, \quad (41)$$

$$x \geq 0, x^+ \geq 0, x^- \geq 0, \quad (42)$$

$$f_2(z) = \sum_j p_j \min [q^T y + R_y^+ y^+ + R_y^- y^-],$$

$$Wy + y^+ - y^- \quad .\mathbf{R}. \quad h(\omega_j) - z, \quad (43)$$

$$y \geq 0, y^+ \geq 0, y^- \geq 0. \quad (44)$$

Дополнительные переменные x^+, x^- (пара для каждой строки (41)) и две переменных y^+, y^- введены в задачу (37)–(44), для обеспечения совместности ограничений (41), (42) и (43), (44). Предполагается, что исходные ограничения (35), (36) являются совместными. Значения штрафных коэффициентов (положительные числа) обозначены $R_x^+, R_x^-, R_y^+, R_y^-$. Очевидно, что если оптимальные значения переменных x^+, x^-, y^+, y^- в решении модифицированной задачи (37)–(44) равны нулю, то оно является решением исходной задачи. Для ее решения используется схема декомпозиции по переменным z с использованием r -алгоритма. Решение подзадач линейного программирования по схеме декомпозиции проводится специализированными программами симплекс-алгоритма.

5. Результаты тестирования

В данном разделе приведены результаты тестирования разработанных программных модулей как на тестовых примерах, из библиотеки тестовых задач пакета SLP-IOR, так и рассмотренных авторами статьи.

Тип и параметры моделей указаны в табл. 1. Случайными являются только правые части ограничений второго этапа с дискретной функцией распределения.

Расчеты проводились на персональном компьютере IBM PC Pentium 200МГц. Используемые обозначения: $m1, m2$ – число ограничений первого и второго этапа соответственно; $n1, n2$ – число переменных первого этапа и второго этапа соответственно; $NRvars$ – число случайных переменных; $NRealiz$ – число случайных реализаций (состояний).

В табл. 2 приведены результаты тестирования модуля Shor1 для решения линейных двухэтапных задач с простой рекурсией.

Результаты сравнивались с результатами работы специализированного программного модуля для решения линейных двухэтапных задач с простой рекурсией, входящего в систему SLP–IOR под названием Sraproх, разработанного Каллом и Майером в 1992 и основанного на методе, описанном в работе Калла и Стояна [24].

Как видно из табл. 2, время достижения оптимальной точки для модуля Shor1 незначительно отличалось от времени для специализированного модуля Sraproх, однако на всех примерах оно оказалось больше.

В табл. 3 приведены результаты тестирования модуля Shor2 для решения линейных двухэтапных задач с полной рекурсией.

Результаты сравнивались с результатами работы программы Qdесom, входящей в пакет SLP–IOR.

Как видно из табл. 3, в моделях Ser1, Pgr, Pgr2 время решения для модуля Qdесom было меньшим, чем для модуля Shor2, но в примере Test10 более быстрым оказался модуль Shor2. Более того, программе Qdесom вообще не удалось найти верное значение целевой функции.

Этот простой тест предназначен для проверки программы на численную устойчивость: первый этап –

$$\min_x EF(x, y(x, \omega), \omega), \quad x \in R^1$$

при ограничении

$$0 \leq x \leq 10^6,$$

второй этап –

$$\min C^+ y^+ + C^- y^- = h(\omega) - x$$

при ограничении

$$y^+ - y^- = h(\omega) - x, \quad y^+, y^- \geq 0.$$

Здесь $C^+ = 10^6$, $C^- = 10^{-3}$, $h(\omega)$ – дискретная случайная величина, $h_i = q^{i-1}$; $q = 1 - 0.001$; вероятность реализации h_i равна

$$p_i = Nq^{-(i-1)}, \quad N = \frac{(q^{-1} - 1)}{q^{-NRealiz} - 1}.$$

Задачи Test10 и Test11 отличаются только количеством реализаций (см. табл. 1).

Квадратичные модели. В табл. 4, 5 приведены результаты тестирования квадратичных моделей с простой рекурсией, полученных путем внесения в линейные модели (целевую функцию и ограничения) квадратичных добавок по переменным первого этапа в виде

Таблица 1. Характеристики моделей

Модель	Тип рекурсии	$NRealiz$	$m1$	$n1$	$m2$	$n2$	$NRvars$
CEP1	Полная	216	9	8	7	15	3
PGP	–	3	2	4	7	12	1
PGP2	–	576	2	4	7	16	3
TEST10	–	2000	1	1	1	2	1
AIRBNDS	Простая	750	4	17	5	10	5
AIRCRAF2	–	3125	4	17	5	10	5
AIRCRAF3	–	$1 \cdot 10^9$	4	17	5	10	5
AIRCRAFT	–	750	4	17	5	10	5
TEST11	–	1000	1	2	1	2	1

$$\min\{\varepsilon x^T x + c^T x + \sum_{w \in \Omega} p(w)[q^{+T} y^+(x, w) + q^{-T} y^-(x, w)]\} \quad (45)$$

при следующих ограничениях: первый этап –

$$\varepsilon x^T x + Ax \leq b, \quad (46)$$

$$l \leq x \leq u; \quad (47)$$

второй этап –

$$y^+ - y^- = h(w) - (\varepsilon x^T x + Tx), \quad (48)$$

$$y^+, y^- \geq 0, \quad (49)$$

где $\varepsilon > 0$.

Результаты расчетов представлены для различных значений ε . При $\varepsilon = 1.e - 5$ решение было близким к решению для соответствующей линейной модели. Во всех случаях время решения квадратичных задач было сопоставимо со временем решения соответствующих линейных.

Заключение

В данной статье приведены результаты работ по использованию методов недифференцируемой оптимизации для решения задач стохастического линейного программирования, выполненных в рамках проекта

Таблица 2. Результаты для моделей с простой рекурсией

Модель	Время решения(с)		Значение целевой функции	
	SHOR1	SRAPPROX	SHOR1	SRAPPROX
AIRBNDS	0.44	0.11	2358.285002	2358.284975
AIRCRAF2	0.44	0.11	1635.172495	1635.172413
AIRCRAF3	0.49	0.11	2193.042275	2193.042189
AIRCRAFT	0.44	0.05	1567.042237	1567.042189
TEST11	0.27	0.22	0.000418	0.000417

Таблица 3. Результаты для моделей с полной рекурсией

Модель	Время решения(с)		Значение целевой функции	
	QDECOM	SHOR_2	QDECOM	SHOR_2
CEP1	3.57	66.76	355159.955734	355159.998172
PGP	0.159	1.65	381.853333	381.853335
PGP2	13.46	153.41	447.324345	447.324345
TEST10	6.48	3.96	499.966885	0.000687

Таблица 4. Время решения для квадратичных моделей с простой рекурсией(сек)

Модель	SHOR1			
	$\varepsilon = 1.e - 5$	$\varepsilon = 1.e - 1$	$\varepsilon = 1.0$	$\varepsilon = 10.0$
QAIRBNDS	0.6	0.6	0.44	0.44
QAIRCRAF2	0.49	0.6	0.38	0.44
QAIRCRAF3	0.55	0.6	0.44	0.49
QAIRCRAFT	0.49	0.66	0.44	0.44

Таблица 5. Значение целевой функции при останове для квадратичных моделей с простой рекурсией

Модель	SHOR1			
	$\varepsilon = 1.e - 5$	$\varepsilon = 1.e - 1$	$\varepsilon = 1.0$	$\varepsilon = 10.0$
QAIRBNDS	2358.143732	1014.954216	1187.0	4500.008
QAIRCRAF2	1635.001399	969.081108	1184.75	4489.08
QAIRCRAF3	2192.849326	1160.733464	1184.75	4489.08
QAIRCRAFT	1566.849234	923.149841	1184.75	4489.08

"Методы решения и аппроксимации в стохастическом программировании".

Были разработаны и включены в систему SLP–IOR блоки решения задач линейного стохастического программирования двух типов: задач с простой рекурсией и задач с фиксированной рекурсией.

Тестирование созданных программных модулей показало, что время решения тестовых задач сравнимо с результатами для специализированных модулей. Существуют примеры, где применение новых блоков решения задач оказалось более эффективным.

Разработанные программные модули характеризуются численной устойчивостью. Эти модули можно использовать также при решении нелинейных задач стохастического программирования.

Авторы благодарят коллег из Института исследования операций университета Цюриха профессора Питера Калла, доктора Яноша Майера и профессора Дитхарда Клатте за плодотворное сотрудничество и выражают глубокую признательность Национальному научному фонду Швейцарии, благодаря которому такое сотрудничество стало возможным.

Литература

1. KALL P., WALLACE S. W. Stochastic programming. – Chichester; New York; Brisbane: John Wiley and Sons, 1994. – 303 p.
2. KALL P., MAYER J. SLP–IOR: An interactive model management system for stochastic linear programs // Math. Programming. – 1996. – 75, № 2. – P. 221–240.
3. KALL P., MAYER J. SLP–IOR: A model management system for stochastic linear programming – system design // Optimization-based computer-aided modelling and design /Ed. by A.J.M. Beulens and H.J. Sebastian. – Berlin: Springer-Verlag, 1992. – P. 139–157.
4. KALL P., MAYER J. A model management system for stochastic linear programming // System Modelling and Optimization / Ed. by P. Kall. – Berlin: Springer-Verlag, 1992. – P. 580–587.
5. KALL P., MAYER J. SLP–IOR: On the design of a workbench for testing SLP codes. // Zurich, 1992. – (Prepr. / IOR, Univ. of Zurich).
6. ЕРМОЛЬЕВ Ю.М. Методы стохастического программирования. – М: Наука, 1976. – 232 с.

7. Михалевич В.С., Трубин В.А., Шор Н.З. Оптимизационные задачи производственно-транспортного планирования. – М.: Наука, – 1986. – 264 с.
8. ПАКЕТ прикладных программ для решения задач производственно-транспортного планирования большой размерности / Сергиенко И.В., Шор Н.З., Трубин В.А. и др. // Кибернетика. – 1983. – № 3. – С. 57–71.
9. ЕРМОЛЬЕВ Ю.М., ШОР Н.З. Метод случайного поиска для двухэтапной задачи стохастического программирования и его обобщения // Там же. – 1968. – № 1. – С. 90–92.
10. ЕРМОЛЬЕВ Ю.М., НУРМИНСКИЙ Е.А. Предельные экстремальные задачи // Там же. – 1973. – № 4. – С. 130–132.
11. НУРМИНСКИЙ Е.А. Условие сходимости алгоритмов стохастического программирования // Там же. – № 3. – С. 84–87.
12. ГОДОНОГА А.Ф. Применение r -алгоритма для решения двухэтапных транспортных стохастических задач // Теория оптимальных решений. – Киев: Ин-т кибернетики АН УССР, 1982. – С. 73–78.
13. БЕЛЯЕВА Л.В., ЖУРБЕНКО Н.Г., ШОР Н.З. Об одной нелинейной транспортной задаче. // Там же. – 1976. – С. 83–89.
14. ERMOLIEV Y., GAIVORONSKII A., NEDEVA S. Stochastic optimization problems with incomplete information on distribution functions // SIAM J. Control and Opt. – 1985. – № 5. – P. 697–716.
15. ЕРМОЛЬЕВ Ю.М., ГАЙВОРОНСКИЙ А. Стохастические методы решения минимаксных задач // Кибернетика. – 1983. – № 4. – С. 92–97.
16. ГАЙВОРОНСКИЙ А. Аппроксимационные методы решения задач стохастического программирования // Там же. – 1982. – № 2. – С. 85–90.
17. ЛИХОВИД А.П. Алгоритм решения специального класса задач стохастического программирования // Теория и приложения методов оптимизации. – Киев: Ин-т кибернетики НАН Украины. – 1998. – С. 8–12.
18. ШОР Н.З. Методы минимизации недифференцируемых функций и их приложения. – Киев: Наук. думка, 1979. – 200 с.

19. SHOR N.Z. Nondifferentiable Optimization and Polynomial Problems. – Boston; Dordrecht; London: Kluwer Acad. Publ., 1998. – 412 p.
20. ШОР Н.З., СТЕЦЕНКО С.И. Квадратичные экстремальные задачи и недифференцируемая оптимизация – Киев: Наук. думка, 1989. – 208 с.
21. HIGLE J. L., SEN S. Stochastic decomposition. A statistical method for large scale stochastic linear programming. – Boston; Dordrecht; London: Kluwer Acad. Publ., 1996. – 198 p.
22. INFANGER G. Planning under uncertainty: Solving large-scale stochastic linear programs. – Danvers: Boyd and Fraser, 1994. – 215 p.
23. BROOK A., KENDRICK D., MEERAUS A. GAMS: User's Guide, Release 2.25. – Danvers: Boyd and Fraser, 1992. – 289 p.
24. KALL P., STOYAN D. Solving stochastic programming problems with recourse including error bounds // Math. Oper. Statist. – 1982. – **13**, № 3. – P. 431–447.

Ч А С Т Ь I I

Вторая часть содержит четыре статьи, связанные с использованием методов недифференцируемой оптимизации при решении матричных, полиномиальных и дискретных задач оптимизации.

Первой помещена обзорная статья [190] по негладким матричным задачам: определение максимального (минимального) собственного числа симметричной матрицы, суммы (взвешенной суммы) k наибольших собственных чисел симметричной матрицы, оптимальных функций Ляпунова в теории устойчивости динамических систем, оптимальных по объему вписанных во многогранник и описанных вокруг многогранника эллипсоидов.

Три следующие статьи [195, 198, 207] связаны главным образом со свойствами лагранжевых (двойственных) оценок в невыпуклых квадратичных моделях. Техника лагранжевых оценок позволяет выделить подклассы задач, разрешимых за полиномиальное время. В сочетании с методом ветвей и границ это позволяет создавать эффективные алгоритмы для решения многоэкстремальных задач.

В статье [195] рассмотрены свойства лагранжевых двойственных оценок для полиномиальных задач, алгоритмы нахождения этих оценок на основе методов негладкой оптимизации.

В статье [198] описано использование функционально избыточных ограничений для улучшения лагранжевых оценок во многоэкстремальных квадратичных задачах. Методика введения избыточных ограничений показана на моделях двух экстремальных задач на графах: задачи нахождения максимального независимого подмножества вершин графа и задачи о максимальном разрезе графа.

В англоязычной статье [207] дан подробный обзор техники лагранжевых оценок и даны примеры ее использования как для задач полиномиального типа и так и ряда экстремальных задач на графах.

Задачи минимизации матричных функций и недифференцируемая оптимизация

Н. З. Шор

Обзорные прикладной и промышленной математики.
— М.: научн. изд.-во "ТВП", 1995. — Т. 2. — С. 113–138.

1. Введение

Пусть $P_{m,n}$ — класс вещественных прямоугольных матриц размерности $m \times n$. Элементы матрицы $A \in P_{m,n}$ будем обозначать a_{ij} , $i = 1, \dots, m$, $j = 1, \dots, n$. Рассмотрим k -мерный вектор параметров $u = \{u_1, \dots, u_k\}$, семейство функций a_{ij} , $i = 1, \dots, m$, $j = 1, \dots, n$ и семейство матричных функций $\Phi_0(A)$, $\Phi_1(A)$, \dots , $\Phi_r(A)$, зависящих от вектора элементов $\{a_{ik}\}$ матрицы A . Пусть $A(u) = \{a_{ij}(u)\}$, $i = 1, \dots, m$, $j = 1, \dots, n$.

Определение. Под задачей оптимизации (минимизации) матричных функций будем понимать следующее: найти

$$\inf_{u \in U} \Phi_0(A(u)), \quad U \subseteq R^k, \quad A \in P_{m,n} \quad (1)$$

при ограничении

$$\Phi_\rho(A(u)) \leq 0, \quad \rho = 1, \dots, r. \quad (2)$$

Формально этот класс задач весьма широк, ведь в форме матрицы можно представить любой вектор, так что к виду (1)–(2) можно свести произвольную конечномерную задачу математического программирования. В более узком смысле мы будем понимать под задачей минимизации матричных функций (или матричной оптимизации) такие задачи, которые *удобно* формулировать в терминах матричного исчисления (оптимизация определителей, спектров (собственных чисел), областей локализации корней характеристических уравнений при ограничениях, которые легко формулируются как условия принадлежности определенному классу матриц и т.п.). Такие задачи в большом числе возникают в математической статистике, планировании экспериментов, теории кодирования, экстремальных задачах на графах и в комбинаторике, при

качественном анализе решений, в вопросах устойчивости, идентификации параметров систем, описываемых дифференциальными уравнениями, а также в задачах выбора оптимального управления и локализации фазовых состояний динамических систем.

Уже этот неполный перечень показывает важность разработки достаточно общих методов матричной оптимизации. С этой тематикой непосредственно связаны матричные неравенства, по которым имеется обширная литература ([5], [55]).

Многие важные характеристики матриц (например, собственные числа симметричных квадратных матриц) представляют собой негладкие функции от элементов матриц. Как следствие этого обстоятельства, задачи матричной оптимизации зачастую оказываются негладкими. В последнее время количество публикаций по матричной оптимизации растет очень быстро. Изучается структура субдифференциалов матричных функций, необходимые и достаточные условия экстремума, вопросы двойственности, при этом интенсивно используются современные методы нелинейного и выпуклого анализа.

В то же время алгоритмические и вычислительные аспекты, за редкими исключениями, разработаны гораздо слабее. При этом проявляется определенная тенденция, связанная с "недоверием" к эффективности, и, в определенной мере, с недооценкой методов негладкой оптимизации. Например, предлагаются гибридные процедуры, заключающиеся в разработке алгоритмов, которые учитывают тонкую структуру субдифференциала в точке минимума и обладают асимптотически высокой скоростью сходимости (сверхлинейной или даже квадратичной). Однако хорошие качества этого, как правило, достаточно сложного алгоритма фактически начинают проявляться лишь в небольшой окрестности минимума, в которой субградиент "хорошо" аппроксимируется элементами субдифференциала G^* в точке минимума [12], [36]. Для попадания же в эту окрестность рекомендуют использовать произвольные методы негладкой оптимизации, в том числе практически и не очень быстрые, например, субградиентный метод или метод эллипсоидов.

Мы исповедуем несколько иной подход. В настоящее время разработан ряд практически эффективных методов негладкой оптимизации, обладающих для выпуклых задач достаточно высокой скоростью сходимости по функционалу и не только в асимптотическом (локальном) смысле, но и в глобальном смысле. Примером могут служить методы субградиентного типа с растяжением пространства в направлении разности последовательных субградиентов (r -алгоритмы). В последнее время отработаны модификации этих алгоритмов, в которых удается

с помощью сравнительно несложных вычислений уточнять в процессе счета нижнюю оценку минимума, контролировать точность по функционалу. Эти средства позволяют практически эффективно решать с высокой точностью различные выпуклые задачи матричной оптимизации, не прибегая к "гибридизации" со специфическими алгоритмами, сходящимися хорошо лишь локально. Для того, чтобы пояснить основные идеи, рассмотрим весьма интересный класс задач, имеющий многочисленные приложения, а именно, класс задач максимизации суммы k наибольших собственных чисел симметричной $(n \times n)$ -матрицы, элементы которой аффинным образом зависят от параметров. При изложении постановки и свойств этой задачи существенно используются работы [8], [36], [38], [39], [40]. Сокращенно указанный класс задач будем называть $MS(n, k)$ -задачами.

Другой, более распространенный, класс матричных задач с ограничениями на неотрицательную определенность будет рассмотрен в §3.

2. Свойства $MS(n, k)$ -задач

Задачи этого типа возникают при анализе целого ряда экстремальных задач на графах.

1. Оценка Ловаса для числа элементов максимального внутренне устойчивого множества графа. Пусть $G(V, E)$ — неориентированный граф без петель, где V — множество его вершин, $V = \{1, 2, \dots, n\}$, $E \subseteq \{\{i, j\}\}$ — множество его ребер, $\{i, j\}$ — ребро, соединяющее вершины i и j . Мы считаем, что $\{i, j\} = \{j, i\}$. Подмножество $\bar{V} \subseteq V$ попарно несмежных вершин, т.е. таких вершин, что если $i_1, i_2 \in \bar{V}$, то $\{i_1, i_2\} \notin E$, называется *внутренне устойчивым* (независимым) *множеством* (ВУМ) графа G . Мощность максимального ВУМ в G обозначается $\alpha(G)$ и называется числом внутренней устойчивости графа G .

Венгерский математик Ловас, исследуя так называемую информационную (по Шеннону) емкость $\theta(G)$ графа [56], предложил следующую оценку для $\alpha(G)$ (она же является оценкой сверху для $\theta(G)$).

Сопоставим графу $G(V, E)$ класс симметричных матриц $A(G)$ вида $\{a_{ij}\}_{i,j=1}^n$, у которых на диагонали стоят единицы и, кроме того, $a_{ij} = 1$, если $\{i, j\} \notin E$. Остальные элементы принимают произвольные значения. Положим $a_{ij} = x_{ij}$ для всех $\{i, j\} \in E$ и допустим, что $x_{ij} = x_{ji}$. Мы получаем параметрическое семейство симметричных матриц $A(x)$, $x = \{x_{ij}\} = \{x_{ji}\}$ для всех $\{i, j\} \in E$. Здесь и далее символом $\lambda_k(M)$

будем обозначать k -е собственное число матрицы M , где индекс k соответствует расположению собственных чисел в порядке убывания с учетом кратности. В частности, $\lambda_1(M)$ — максимальное собственное число. Тогда оценка Ловаса $\nu(G)$ ($\geq \theta(G) \geq \alpha(G)$) выражается следующим образом [32]:

$$\nu(G) = \min_x \lambda_1(A(x)). \quad (3)$$

Мы видим, что задача (3) принадлежит классу задач $MS(n, 1)$. Из того, что $\lambda_1(M) = \max_{\|y\|=1} (My, y)$, вытекает, что $\lambda_1(A(x))$ ($A(x) \in \mathcal{A}(G)$) есть выпуклая функция от x , определенная на $|E|$ -мерном евклидовом пространстве, (i, j) -компонента субградиента этой функции в точке \bar{x} вычисляется по формуле $g_{ij} = \bar{y}_i \bar{y}_j$, $\{i, j\} \in E$, где \bar{y}_s есть s -я компонента произвольного нормированного вектора, соответствующего собственному числу $\lambda_1(A(x))$. Мы видим, что если в точке \bar{x} собственное число $\lambda_1(A(x))$ имеет кратность $r > 1$, то субградиент функции $\varphi_1(x) = \lambda_1(A(x))$ в точке \bar{x} определяется неоднозначно, т.е. $\varphi_1(x)$ может оказаться негладкой функцией [54]. Отметим, что оценка $\nu(G)$ может быть получена как результат решения ряда других задач матричной оптимизации, в которых используется условие неотрицательной определенности соответствующих матриц [22]. Об этом речь будет идти в §3.

Нарасимхан и Манбер сравнительно недавно получили обобщение оценки Ловаса [35]. Пусть $\alpha_k(G)$ обозначает размер максимального по числу вершин k -дольного индуцированного подграфа графа $G(V, E)$, т.е. максимальное число вершин, которое можно разместить в k внутренне устойчивых множествах графа G . Соответствующая оценка сверху $\nu_k(G)$ имеет вид:

$$\nu_k(G) = \min_x \sum_{i=1}^k \lambda_i(A(x)), \quad A(x) \in \mathcal{A}(G), \quad x \in \{x_{ij}\}, \quad \{i, j\} \in E. \quad (4)$$

При $k = 1$ эта оценка совпадает с оценкой Ловаса (3). Пусть $\varphi_k(x) = \sum_{i=1}^k \lambda_i(A(x))$. Оказывается, что $\varphi_k(x)$ является выпуклой функцией вектора x , не обязательно дифференцируемой. При этом точка \bar{x} является точкой разрыва градиента, если среди первых собственных чисел матрицы $A(\bar{x})$, упорядоченных в порядке неубывания, есть хотя бы одно с кратностью, превышающей 1. Это следует из вариационного определения, данного Фань Цзы [5], [55] для функции $\psi_k(A) = \sum_{i=1}^k \lambda_i(A)$, где A — из класса симметричных вещественных матриц S_n размерности

$n \times n$:

$$\psi_k(A) = \max_{Y \in M_n^k} \left\{ \text{tr}(AY Y^T) \right\}, \quad k \leq n, \quad (5)$$

где $\text{tr}Z$ означает след (сумму диагональных элементов) матрицы Z , M_n^k — класс таких прямоугольных матриц Y размера $n \times k$, что $Y^T Y = I_k$ (I_k — единичная матрица размера $k \times k$), T — знак транспонирования. Так как $\psi_k(A)$ представляет собой результат поточечной максимизации параметрического семейства линейных функций от элементов $\{a_{ij}\}_{i,j=1}^n$ матрицы A по компактному множеству параметров $Y = \{y_{ij}\}$, $i = 1, \dots, m$, $j = 1, \dots, n$, то $\psi_k(A)$ — выпуклая функция параметров матрицы A . Отсюда следует, что $\psi_k(x) = \sum_{i=1}^k \lambda_i(A(x))$ также является выпуклой по x , так как элементы $A(x)$ выражаются аффинным образом через координаты x .

Замечание. Как следствие предыдущих рассуждений, мы получаем следующее утверждение. Пусть $w_1 \geq w_2 \geq \dots \geq w_k > 0$ суть компоненты весового вектора $\{w = w_1, \dots, w_k\}$. Тогда функция

$$\psi_k^w(A) = \sum_{i=1}^k w_i \lambda_i(A)$$

является выпуклой функцией элементов матрицы A . В самом деле, $\psi_k^w(A)$ представима в виде линейной комбинации выпуклых функций с неотрицательными коэффициентами:

$$\psi_k^w(A) = w_k \psi_k(A) + (w_{k-1} - w_k) \psi_{k-1}(A) + \dots + (w_1 - w_2) \psi_1(A).$$

Отметим, что максимальное значение в формуле Фань Цзы (5) достигается на тех матрицах Y^* , столбцы которых представляют собой ортогональную систему собственных векторов, соответствующих максимальным собственным числам матрицы A . Это проверяется непосредственно:

$$\text{tr}(AY^*(Y^*)^T) = \sum_{i=1}^k (\lambda_i(A) y_i^*, y_i^*) = \sum_{i=1}^k \lambda_i(A)$$

(здесь y_i^* есть i -й столбец матрицы Y^*).

Если кратность k максимальных собственных чисел матрицы \bar{A} равна 1, то $\psi_k(A)$ дифференцируема при $A = \bar{A}$ и градиент $g_\psi(\bar{A}) = \{\partial \psi_k / \partial a_{ij}\}_{i,j=1}^n$ вычисляется по формуле

$$\frac{\partial \psi_k}{\partial a_{ij}}(\bar{A}) = \sum_{r=1}^k \bar{y}_i^r \bar{y}_j^r, \quad (6)$$

где \bar{y}_s^r ($s = 1, \dots, n$) — компоненты r -го ортонормированного собственного вектора матрицы \bar{A} . Если кратность собственного числа матрицы \bar{A} хотя бы для одного r , $r = 1, \dots, k$, превышает 1, то соответствующие собственные векторы определяются (с точностью до знака) неоднозначно, градиент $g_\psi(A)$ в точке \bar{A} терпит разрыв, но существуют субградиенты, вычисляемые по формуле (6) для произвольных ортонормированных представителей собственных векторов. Таким образом, используя алгоритмы нахождения собственных чисел и собственных векторов симметричных матриц, мы легко можем вычислить субградиент функции $\psi_k(\cdot)$ в произвольной точке \bar{A} . Функции вида $\varphi_k(x) = \psi_k(A(x))$, как и функции вида $\varphi_k^w(x) = \psi_k^w(A(x))$, нужно рассматривать как сложные функции и соответственным образом вычислять частные производные и субградиенты.

Однако вычисление собственных чисел и собственных векторов симметричных $(n \times n)$ -матриц при больших n требует большого времени. Можно ли обойтись без столь громоздких вычислений?

Легко видеть, что если в формуле (5) мы заменим условие $Y \in M_n^k$ на условие $YU^T \in \text{co}\{ZZ^T\}$, $Z \in M_n^k$, то формула останется верной (со — символ выпуклого замыкания множества). Определим класс матриц $M_{k,n}^{C_0} = \text{co}\{ZZ^T : Z \in M_n^k\}$.

Теперь можно воспользоваться следующим результатом [36].

Теорема. *Справедливо соотношение*

$$M_{k,n}^{C_0} = \{C : \text{tr}C = k, 0 \leq \lambda_n(C) \leq \lambda_1(C) \leq 1\}. \quad (7)$$

Другими словами, $M_{k,n}^{C_0}$ — это класс симметричных $(n \times n)$ матриц, спектр которых расположен на интервале $[0, 1]$, а сумма диагональных элементов равна k . Для $\psi_k(A)$ получаем следующее выражение:

$$\psi_k(A) = \max_{C \in M_{k,n}^{C_0}} (A, C), \quad (8)$$

где $(A, C) = \sum_{i,j=1}^n a_{ij}c_{ij}$. Так как $I_k - C$, $C \in M_{k,n}^{C_0}$, является неотрицательно определенной матрицей [47], то $c_{ii} \leq 1$, $i = 1, \dots, n$ (свойство неотрицательной определенности матрицы A будем записывать как $A \succeq 0$). Отсюда $|c_{ij}| \leq 1$ для любых пар (i, j) , т.е. множество $M_{k,n}^{C_0}$ является ограниченным, а, значит, и компактным. Функция $\psi_k(A(x))$ — выпуклая функция параметра, и ставится задача ее минимизации. Используя (8), получаем

$$f^* = \min_x \psi_k(A(x)) = \min_x \max_{C \in M_{k,n}^{C_0}} (A(x), C) = \min_x \max_{C \in M_{k,n}^{C_0}} \sum_{i,j=1}^n a_{ij}(x)c_{ij}.$$

Из условия $C \in M_{k,n}^{C_0}$ следует, что $|c_{ij}| \leq 1$ для всех (i, j) и

$$\sum_{i=1}^n c_{ii} = k, \quad c_{ii} \geq 0, \quad i = 1, \dots, n. \quad (9)$$

Пусть C_0 принадлежит относительной внутренней $M_{k,n}^{C_0}$. Будем строить максимизирующую последовательность C_1, \dots, C_p, \dots таким образом, чтобы $\{C_p\}_{p=1}^{\infty}$ оставались внутренними точками. Для этого можно воспользоваться внутренними штрафными функциями $S_\varepsilon(C) = \varepsilon (\ln \det C + \ln \det (I_k - C))$. Эти функции вогнуты на множестве $C \succ 0$ и при подходе к границе ограничений по спектру ($\lambda_n(C) \geq 0, \lambda_1(C) \leq 1$) стремятся к $-\infty$. Условие (9) можно удовлетворить путем исключения c_{nn} :

$$c_{nn} = k - \sum_{i=1}^{n-1} c_{ii}. \quad (10)$$

Таким образом, для приближенного решения задачи (8) добавим к $\varphi_k(A(x))$ штрафную функцию с параметром ε и получим следующую задачу: найти

$$f_\varepsilon^* = \min_x \psi_k^\varepsilon(A(x)) = \min_x \max_{C \in K} \{(A(x), C) + \varepsilon (\ln \det C + \ln \det (I_k - C))\},$$

где

$$K = \left\{ C : |c_{ij}| \leq 1 \text{ для всех } (i, j), c_{nn} = k - \sum_{i=1}^{n-1} c_{ii} \right\}. \quad (11)$$

Как правило, рассматривая конкретную задачу, можно найти априорные двусторонние ограничения на область вариации параметров x_{ij} вида

$$\Pi = \{a_{ij} \leq x_{ij} \leq \beta_{ij} \text{ для всех } (i, j)\}. \quad (12)$$

Путем перестановки операций взятия \max и \min (что в данном случае правомерно, так как речь идет о седловой точке выпукло-вогнутой функции [26]), получаем

$$f_\varepsilon^* = \max_{C \in K} \left\{ \left[\min_{x \in \Pi} (C, A(x)) \right] + \varepsilon (\ln \det C + \ln \det (I_k - C)) \right\}. \quad (13)$$

При фиксированном C внутренняя задача минимизации решается тривиально при аффинном вхождении параметров x в элементы матрицы A .

2. Задача о разбиении графа Пусть задан неориентированный граф $G(V, E)$. Рассмотрим задачу о разбиении его вершин V на заданное число непересекающихся подмножеств с заданным ограничением на число вершин в каждом подмножестве так, чтобы число ребер, у которых концы лежат в разных подмножествах, было минимально. Такая задача возникает особенно часто при проектировании электронных схем, устройств, разработке сложных программ и т.п., когда применяется блочный принцип разработки или конструирования.

Донат и Хофман [9], [10] предложили для приближенного решения этой сложной комбинаторной задачи использовать следующую оценку:

$$\rho^* = -\frac{1}{2} \min_x \left\{ \sum_{i=1}^k m_i \lambda_i(A_0 + D(x)) : \text{tr}(D(x)) = 0 \right\},$$

где $m_1 \geq m_2 \geq \dots \geq m_k \geq 0$, $\sum_{i=1}^k m_i = |V|$, $V = \bigcup_{i=1}^k V_i$, а $\{V_i\}_{i=1}^k$ — неизвестные подмножества разбиения множества вершин, $|V_i| = m_i$, $i = 1, \dots, k$. Недиagonальные элементы матрицы $A_0 = \{a_{ij}^0\}$ равны 1, если вершины i и j связаны ребром, и 0 — в противоположном случае, а a_{ii}^0 определяются по формуле

$$a_{ii}^0 = - \sum_{j=1, j \neq i}^{|V|} a_{ij}^0, \quad i = 1, \dots, |V|.$$

Пусть $D(x)$ — диагональная матрица, у которой на i -м диагональном месте стоит x_i , т.е. i -я компонента вектора $\{x_1, \dots, x_{|V|}\}$. Мы видим, что задача нахождения ρ^* относится к классу $SM_{n,k}^w$ -задач. Кулум, Донат и Вулф [8] предложили сходящийся алгоритм решения этой задачи для случая равных m_i , $i = 1, \dots, k$. Их алгоритм требует вычисления собственных чисел и собственных векторов на каждой итерации и в полной мере учитывает специфику задачи.

Замечание. Статья [8] замечательна еще тем, что она была опубликована в выпуске, впервые целиком посвященном недифференцируемой оптимизации, в котором наряду с интересными публикациями западных авторов достаточно полно была представлена библиография и советских работ этого направления, в то время малоизвестных на Западе.

В работе [36] на основе обобщения результатов статьи [8] предложен усовершенствованный алгоритм для решения общей задачи минимизации взвешенной суммы k максимальных собственных чисел типа $MS^w(n, k)$, который эффективно работает, если начальное приближение достаточно близко к точке минимума.

Итак, мы видим, что разнообразные прикладные задачи сводятся к классу $MS^w(n, k)$ -задач. Используя теорию двойственности, эти задачи можно свести к задаче максиминного (минимаксного) типа, имеющий вид (13). Используя штрафные функции вида $\varepsilon \ln \det(\cdot)$, можно приближенно учитывать ограничения на спектр матрицы C . Для получения значений прямых переменных (параметров матричной функции) целесообразно вводить квадратичные сепарабельные добавки вида $\delta_i x_i^2$, где $\delta_i > 0$, а x_i — параметры. При этом получаем

$$f_{\varepsilon, \delta}^* = \max_{C \in K} \left\{ \min_{x \in \Pi} \left[(C, A(x)) + \varepsilon (\ln \det C + \ln \det(I_k - C)) + \sum_i \delta_i x_i^2 \right] \right\}.$$

В этой сглаженной задаче обеспечивается существование единственной седловой точки. Внутренняя квадратичная задача для x при фиксированном C решается тривиально, а для нахождения приближенного значения C^* можно, как показывают эксперименты, с успехом применять r -алгоритм. Полученной матрице C_ε^* соответствует вектор параметров x_ε^* . Величина $(C_\varepsilon^*, A(x_\varepsilon^*))$ дает оценку снизу для $\psi_k^w(A(x^*))$. Решив для полученного x_ε^* задачу на собственные значения для матрицы $A(x_\varepsilon^*)$, мы можем оценить точность решения и, если она недостаточна, работать с меньшими множителями при штрафных и сглаживающих добавках.

3. Задачи минимизации матричных параметрических функций с ограничениями на неотрицательную определенность

Рассмотрим множество S_n вещественных симметричных матриц размера $n \times n$. Каждая матрица $A = \{a_{ij}\}_{i,j=1}^n \in S_n$ задается $n(n+1)/2$ своими элементами: n из них лежат на главной диагонали, остальные a_{ij} ($i \neq j$) равны a_{ji} , так что достаточно указать еще $n(n-1)/2$ попарно несимметричных элементов. Таким образом, в S_n можно естественным образом ввести операцию умножения матрицы на вещественное число и операцию сложения, сводящиеся к поэлементному выполнению этих операций, и в результате получим векторное пространство S_n^L размерности $n(n+1)/2$. Пусть $\{y_1, \dots, y_n\}$ — переменные элементы евклидова пространства E^n . Каждой матрице $A \in S_n$ мы можем сопоставить квад-

ратичную форму

$$K_A(y) = (Ay, y) = \sum_{i,j=1}^n a_{ij}y_iy_j.$$

Матрица A называется неотрицательно определенной ($A \succeq 0$), если $(Ax, x) \geq 0$ для всех x . Если $(Ax, x) > 0$ для всех $x \neq 0$, то A называется положительно определенной матрицей ($A \succ 0$). Важность этих классов матриц обусловлена тем, что функция $K_A(y)$ является выпуклой (строго выпуклой) тогда и только тогда, когда $A \succeq 0$ ($A \succ 0$).

Нетрудно убедиться, что неотрицательно (положительно) определенные матрицы образуют в пространстве S_n^L выпуклый конус \overline{K}_n^+ (K_n^+), причем \overline{K}_n^+ является замыканием K_n^+ в естественной топологии [54]. Как же проверить, что $A \succeq 0$ ($A \succ 0$)? Существует целый ряд критериев для решения этой задачи [47]:

- (а) возможность представления $K_A(x)$ в виде суммы квадратов линейных форм (суммы квадратов n линейно независимых форм);
- (б) критерий Сильвестра: определители всех главных миноров неотрицательны (положительны), при этом достаточно рассмотреть систему из n вложенных последовательно друг в друга миноров;
- (в) неотрицательность (положительность) минимального собственного числа матрицы A : $\lambda_n(A) \geq 0$ ($\lambda_n(A) > 0$).

С критериями (а) и (б) удобно работать, когда речь идет о фиксированной матрице. Если же мы рассматриваем множество матриц $A(x) \in S_n$, зависящих от векторного параметра $x = \{x_1, \dots, x_m\}$, то с критериями (а) и (б) работать сложно, даже если элементы $a_{ij}(x)$ представляют собой линейные функции от параметров. Критерий Сильвестра, например, сводится тогда к системе, вообще говоря, невыпуклых неравенств, некоторым из них может соответствовать даже несвязная область в пространстве параметров. В случае параметрической оптимизации с аффинным вхождением параметров в элементы матрицы наибольший интерес представляет критерий (в), так как $\lambda_n(A)$ представляет собой вогнутую функцию от элементов матрицы: соответственно, $\varphi(x) = \lambda_n(A(x))$ будет вогнутой функцией от x , т.е. неравенство

$$-\varphi(x) \leq 0 \tag{14}$$

будет выпуклым.

Путем тех же рассуждений, которые мы проводили в §2, легко показать, что функция вида $\varphi(x)$, как правило, будет негладкой, разрыв градиента может произойти в тех точках \bar{x} , для которых кратность $\lambda_n(A(\bar{x}))$ превышает 1. Из этого следует не очень приятный факт, что большинство задач параметрической матричной оптимизации с ограничениями на неотрицательную определенность являются негладкими, что затрудняет их решение классическими методами, причем этот факт, как ни удивительно, был обнаружен сравнительно недавно. Вот как пишет об этом в [18] известный математик Флетчер, один из авторов знаменитого квазиньютоновского метода ДФП (Давидона - Флетчера - Пауэлла) безусловной гладкой оптимизации: "Мой интерес к задачам с матричными ограничениями полуопределенности был инициирован некоторое время тому назад [14] изучением задачи "образовательного" тестирования: дана положительно определенная матрица S ; сколько можно "извлечь" из ее диагонали, чтобы в результате осталась неотрицательно определенная матрица. Это приводит к задаче: найти

$$\max \sum_{i=1}^n \theta_i, \quad \theta = \{\theta_i\}_{i=1}^n \in R^n, \quad (15)$$

при ограничениях

$$S - \text{diag}\theta \succeq 0, \quad \theta \geq 0. \quad (16)$$

Первые попытки решить эту задачу ... (идут ссылки на [14]–[16]) ... были не очень успешны. ... в них ограничение на неотрицательную определенность сводилось к ограничению на собственное число, и применялась стандартная методика нелинейного программирования. В моём случае это соответствовало предположению, что ограничение на собственное число будет гладким в точке решения, за исключением редких случаев. Это оказалось неправильным, в действительности подавляющее большинство таких задач, хотя и не все, негладки в точке решения."

Далее в этой работе упоминается ещё одна подобная задача (модификация матрицы, которая возникает в Ньютоно-подобных методах безусловной минимизации). В этом случае G – симметричная неопределенная матрица. Как по возможности меньше "добавить" к диагонали, чтобы в результате получить неотрицательно определенную матрицу? Поставленный вопрос приводит к задаче такого же типа, что и предыдущая: найти

$$\min_{\theta \geq 0} \sum_{i=1}^n \theta_i, \quad \theta \in R^n, \quad G + \text{diag} \theta \succeq 0. \quad (17)$$

Далее в [18] изучается структура субградиентов и штрафных функций, связанных с матричными ограничениями в форме неотрицательной определенности. Это, по-видимому, одна из первых работ, в которой подробно изучались свойства субдифференциалов негладких функций, возникающих в связи с ограничениями неотрицательной определенности, и строились практически эффективные алгоритмы, когда выбираемые параметры входят лишь в диагональные элементы матрицы (например, для задачи "образовательного тестирования").

Эти методы [18] можно рассматривать как модификации *SQP*-метода (последовательное квадратичное программирование). В отличие от подхода Флетчера наш подход полностью основан на алгоритмах негладкой оптимизации. Мы столкнулись с экстремальными матричными задачами, в которых основные ограничения связаны с неотрицательной определенностью, при рассмотрении двойственных оценок для нахождения максимального взвешенного внутренне устойчивого множества графа $G(V, E)$ [32]. Подобная задача уже рассматривалась в §2. Мы предложим формулировку этой задачи в виде экстремальной задачи квадратичного типа. Пусть $V = \{1, \dots, n\}$. Сопоставим каждой вершине $i \in V$ переменную $x_i \in \{0, 1\}$, положив $x_i = 1$, если вершина i принадлежит искомому ВУМ, и $x_i = 0$, если не принадлежит. Тогда мы получим следующую задачу нелинейного программирования [54]: найти

$$\nu_w(G) = \max \sum_{i=1}^n w_i x_i \quad (18)$$

при условиях несовместимости

$$x_i x_j = 0 \quad \text{для всех} \quad \{i, j\} \in E \quad (19)$$

и булевости переменных

$$x_k^2 - x_k = 0 \quad \text{для всех} \quad k \in V, \quad (20)$$

где $w = \{w_1, \dots, w_n\}$ — вектор положительных "весов". Введем множители Лагранжа $u_{ij} = u_{ji}$ для ограничений (19) и u_k для ограничений (20). Вектор, составленный из указанных множителей, будем обозначать u . Построим функцию Лагранжа для задачи (18)–(20):

$$L(x, u) = \sum_{i=1}^n w_i x_i - \sum_{\{i, j\} \in E} u_{ij} x_i x_j - \sum_{k \in V} u_k (x_k^2 - x_k).$$

Пусть $\psi(u) = \sup_x L(x, u)$, $\psi^* = \inf_u \psi(u)$ ($\psi(u) \geq \alpha_w(G)$ для всех u и $\psi^* \geq \alpha_w(G)$). Легко убедиться, что $\psi(u)$ — вогнутая функция, $\text{dom } \psi$

содержит все u , для которых матрица $-M(u)$ квадратичной по x формы функции Лагранжа $L(x, u)$ является отрицательно определенной, где $M(u) = \{m_{ij}\}_{i,j=1}^n$ и

$$m_{ij} \begin{cases} u_{ij}, & \{i, j\} \in E, \\ u_i, & i = j, \\ 0 & \text{в остальных случаях.} \end{cases}$$

Пусть $\Omega^+ = \{u : M(u) \succ 0\}$, $\bar{\Omega}^+ = \{u : M(u) \succeq 0\}$. Как показано в [54], двойственная оценка ψ^* для $\alpha_w(G)$ совпадает с $\nu_w(G)$, где ψ^* — оптимальное значение следующей параметрической матричной задачи: найти

$$\psi^* = \inf_u \sup_x \left[\sum_{i=1}^n (w_i + u_i)x_i - (M(u), x) \right], \quad M(u) \succeq 0 \text{ (или } u \in \Omega^+ \text{)}. \quad (21)$$

Таким образом мы получим задачу матричной оптимизации с ограничениями типа неотрицательной определенности.

Другая формулировка задачи, связанной с оценкой Ловаса, также содержит подобные ограничения. Пусть $G(V, E)$ — граф. Рассмотрим класс \mathcal{B} симметричных матриц B с элементами b_{ij} ($i = 1, \dots, n$, $j = 1, \dots, n$), для которых $b_{ij} = 0$, если $\{i, j\} \in E$, и $\sum_{k=1}^n b_{kk} = 1$, $B \succeq 0$. Как показано в [32], [22],

$$\nu_w(G) = \max_{B \in \mathcal{B}} \sum_{i,j=1}^n \sqrt{w_i w_j} b_{ij}. \quad (22)$$

Задача (22) также является задачей выпуклого программирования с ограничением в форме неотрицательной определенности.

Рассмотрим естественные обобщения [54].

Задачи вида (18)–(20) являются примерами задачи квадратичного типа. Используя функцию Лагранжа, мы получили для нее двойственную оценку (21). Подобная оценка может быть получена для произвольной задачи квадратичного типа, имеющий вид: найти

$$f^* = \inf K_0(x), \quad x \in E^n, \quad (23)$$

$$K_i(x) \leq 0, \quad i \in \mathcal{I}, \quad (24)$$

$$K_j(x) = 0, \quad j \in \mathcal{J}, \quad \mathcal{J} \cap \mathcal{I} = \emptyset, \quad (25)$$

где K_0 и K_i , $i \in \mathcal{I}$, K_j , $j \in \mathcal{J}$, — квадратичные или линейные функции. Формируем функцию Лагранжа ($u = \{u_i, i \in \mathcal{I}; u_j, j \in \mathcal{J}\}$ — вектор множителей Лагранжа)

$$L(x, u) = K_0(x) + \sum_{i \in \mathcal{I}} u_i K_i(x) + \sum_{j \in \mathcal{J}} u_j K_j(x)$$

и маргинальную функцию

$$\psi(u) = \inf_x L(x, u).$$

Если $\text{dom } \psi$ не является пустым множеством и $\psi(u)$ — собственная выпуклая функция, то можно получить нетривиальную оценку снизу для f^* :

$$f^* \geq \varphi^* = \sup_{\{u: u_i \geq 0, i \in \mathcal{I}\}} \psi(u).$$

Пусть $M_L(u)$ — матрица квадратичной формы по x функции Лагранжа $L(x, u)$. Заметим, что элементы этой матрицы зависят от u линейно, а $\text{dom } \psi$ содержит $\Omega^+ = \{u : M_L(u) \succ 0\}$ и не содержит u , для которых $M_L(u)$ имеет отрицательное минимальное собственное число. Пусть $\bar{\Omega}^+ = \{u : M_L(u) \succeq 0\}$. Тогда оценка определяется в результате решения матричной экстремальной задачи:

$$\varphi^* = \sup_{\{u: u_i \geq 0, i \in \mathcal{I}\}} \inf_x [(M_L(u)x, x) + l(x, u)], \quad u \in \bar{\Omega}^+, \quad (26)$$

где $l(x, u)$ — это линейная по x часть функции Лагранжа. (Более подробное изложение такого подхода к задачам квадратичного типа имеется в монографии [54].)

К оптимизационным задачам на классе неотрицательно определенных матриц приводят некоторые задачи теории устойчивости динамических систем. Показателями функционирования сложных объектов, описываемых системами дифференциальных уравнений, являются различные характеристики: время переходного процесса, расхождение траекторий из-за возмущений, область устойчивости и диссипации и т.п. Прямой метод Ляпунова исследования устойчивости теоретически дает возможность получать достаточно точные оценки таких характеристик. Однако не существует общего конструктивного подхода к построению функций Ляпунова. Поэтому оптимизация по определенным критериям функций Ляпунова представляет существенный интерес. Ниже мы будем опираться на результаты работы [51].

Рассмотрим систему дифференциальных уравнений с постоянными коэффициентами:

$$\frac{dx}{dt} = Ax, \quad x \in \mathcal{R}^n. \quad (27)$$

Построим функцию Ляпунова в виде квадратичной формы $v(x) = (H(x), x)$, полная производная которой (в силу (27)) равна заранее заданной функции $w(x) = -(Cx, x)$. Положительно определенные симметричные матрицы H и C связаны уравнением Ляпунова (ср. [47]):

$$A^T H + H A = -C.$$

Функция Ляпунова дает возможность получить следующую оценку для системы (27):

$$\|x(x_0, t)\| \leq \sqrt{\frac{\lambda_1(H)}{\lambda_n(H)}} \|x_0\| \exp \left\{ -\frac{\lambda_n(C)}{2\lambda_1(H)} t \right\}, \quad (28)$$

где $\|\cdot\|$ — одна из норм в \mathcal{R}^n , $x(x_0; t)$ — решение системы (27) при $x = x_0$ для $t = 0$.

Используя оценку (28), можно получить интегральную оценку

$$J = \int_0^\infty \|x(x_0, t)\| dt \leq 2\|x_0\| \left(\sqrt{\frac{\lambda_1(H)}{\lambda_n(H)}} \right)^3 \quad (29)$$

нормы фазовой координаты $\|x(x_0, t)\|$.

В работе [51] доказано, что наиболее точная оценка интегрального критерия качества вида (29) определяется функцией Ляпунова, для которой

$$H^* = \arg \min_{H \in \mathcal{G}} \frac{(\lambda_1(H))^{3/2}}{\sqrt{\lambda_n(H)} [\lambda_n(-(A^T H + H A))]}, \quad (30)$$

где \mathcal{G} — множество таких симметричных положительно определенных матриц H , что матрица $C = -(A^T H + H A)$ положительно определена. Функция Ляпунова $v^*(x) = (H^*(x), x)$ называется оптимальной в интегральном смысле.

Функция

$$\varphi(H) = \frac{(\sqrt{\lambda_1(H)})^{3/2}}{\sqrt{\lambda_n(H)} \lambda_n[-(A^T H + H A)]},$$

непрерывна, почти дифференцируема на открытом выпуклом конусе \mathcal{G} . Перейдем от минимизации функции $\varphi(H)$ к минимизации функции $\psi(H) = 2 \ln \varphi(H)$:

$$\psi(H) = 3 \ln \lambda_1(H) - \ln \lambda_n(H) - 2 \ln \lambda_n [-(A^T H + H A)].$$

Так как $\varphi(H) = \varphi(\alpha H)$ при произвольном положительном α , то можно считать, что $\lambda_1(H) = 1$. При этом условии получаем задачу нелинейного программирования следующего вида: найти

$$\max_H \psi_1(H) = \max_H [\ln \lambda_n(H) + 2 \ln \lambda_n [-(A^T H + H A)]] . \quad (31)$$

при условиях

$$\lambda_1(H) = 1, \quad H \succ 0; \quad C = -(A^T H + H A) \succ 0. \quad (32)$$

Приведем одно утверждение из [54], которое будет использовано в дальнейшем.

Лемма. Пусть $f(x)$ — собственная вогнутая функция, $x \in \mathcal{R}^n$, $\varphi(t)$ — монотонно неубывающая вогнутая функция, определенная для $t > 0$. Тогда сложная функция $\psi(x) = \varphi[f(x)]$ является собственной вогнутой функцией, причем

$$\text{dom } \psi = \text{dom } f \cap \{x : f(x) > 0\} \quad (33)$$

при условии, что правая часть (33) непуста.

Легко проверить, что составляющие функции $\psi(H)$ удовлетворяют условию леммы. Если $H \in \mathcal{G}$, то $\lambda_n(H)$ и $\lambda_n(C)$ принимают положительные значения. Отсюда следует, что функция $\psi(H)$ является вогнутой при $H \in \mathcal{G}$.

Далее, ограничение $\lambda_1(H) = 1$ в данном случае можно заменить выпуклым ограничением $\lambda_1(H) \leq 1$ (легко убедиться, что максимум достигается на границе этой области).

Итак, получаем задачу выпуклого программирования: найти

$$\min [-\ln \lambda_n(H) - 2 \ln \lambda_n [-(A^T H + H A)]] , \quad (34)$$

$$\lambda_1(H) \leq 1, \quad (35)$$

$$H \succ 0; \quad -(A^T H + H A) \succ 0. \quad (36)$$

Если двигаться из допустимой области $G(H)$, то при приближении к границе этой области $\lambda_n(H)$ и (или) $\lambda_n(-(A^T H + H A))$ стремится

к 0, т.е. минимизируемая функция (34) стремится к $+\infty$. (Это значит, что минимум достигается во внутренней точке области $G(H)$). Условие $\lambda_1(H) \leq 1$ можно учесть приближенно, введя штрафную добавку вида $\varepsilon \ln \det(\mathcal{J}_n - H)$.

Учитывая все выше сказанное, мы можем применить для решения задачи получения оптимальной функции Ляпунова (в интегральном смысле) модификацию r -алгоритма с контролем шагового множителя, препятствующего выходу из областей положительной определенности матриц H и $C(H) = -(A^T H + H A)$.

Важный класс задач матричной оптимизации с ограничением на положительную определенность составляют задачи нахождения оптимальных по объему вписанных и описанных эллипсоидов. Эти задачи имеют широкий круг приложений: эллипсоидальные аппроксимации компактных множеств (внешние и изнутри), например, в теории оценивания областей достижимости управляемых динамических систем, задачи идентификации параметров, оптимальное планирование экспериментов, аппроксимация областей локализации оптимальных параметров в задачах математического программирования — вот далеко не полный перечень возможных приложений.

Пусть W_n — множество выпуклых компактных тел n -мерного евклидова пространства E_n , содержащих внутренние точки. Пусть $M \in W_n$; $O_1(M)$ (соответственно $O_2(M)$) — множество описанных вокруг M (вписанных в M) эллипсоидов $E(K, y_0)$ вида $\{x : (K(x - y_0), x - y_0) \leq 1\}$ (здесь K — положительно определенная симметричная матрица, y_0 — центр эллипсоида). Как известно, n -мерный объем эллипсоида $E(K, y_0)$ выражается через его матрицу K следующим образом:

$$v[E(K, y_0)] = \omega_n (\det K)^{-1/2},$$

где ω_n есть объем n -мерной сферы единичного радиуса.

Заметим, что задачи о вписанных и описанных эллипсоидах тесно связаны между собой через так называемые полярные преобразования Минковского [22], [44]: $\text{po}M = \{y : (x, y) \leq 1; \forall x \in M\}$.

Пусть $M \in W_n$ и $0 \in \text{int}M$. Тогда

$$\text{po}(\text{po}M) = M,$$

$$\text{po} \left(\bigcap_{i=1}^m M_i \right) = \text{co} \left(\bigcup_{i=1}^m \text{po}M_i \right); \quad M_i \in W_n; \quad 0 \in \text{int}M_i, \quad i = 1, \dots, m.$$

Если $M_1 \subset M_2$, то $\text{po}M_1 \supset \text{po}M_2$; если M — эллипсоид, то $\text{po}M$ — также эллипсоид; если $M = E(K, 0)$, то $\text{po}M = E(K^{-1}, 0)$ и

$$v(M)v(\text{po}M) = \omega_n^2.$$

Следовательно, если E^* — эллипсоид минимального объема, описанный вокруг M (центр эллипсоида фиксируется как начало координат), то $\text{po}E^*$ — эллипсоид максимального объема, вписанный в $\text{po}M$, при условии, что центр его остается в 0. Джон [44] доказал существование и единственность оптимальных вписанных эллипсоидов, а также существование для каждого $M \in W_n$ гомотетичных эллипсоидов с общим центром (вписанного и описанного) с коэффициентом гомотетии не менее $1/n$ (для центрально симметричных тел коэффициент гомотетии не менее $1/\sqrt{n}$).

Пусть M — многогранник, задаваемый как выпуклая оболочка точек a_1, \dots, a_k , т.е. $M = \text{co}\{a_1, \dots, a_k\}$. Рассмотрим задачу построения эллипсоида минимального объема, включающего M , с центром в начале координат. Эта задача эквивалентна следующей задаче математического программирования: найти

$$\min_{K \succ 0} (-\ln \det K) \quad (37)$$

при ограничениях

$$(K a_i, a_i) \leq 1; \quad i = 1, \dots, k. \quad (38)$$

Как показано в [44], эта задача сводится к задаче безусловной минимизации негладкой выпуклой функции:

$$\varphi(K) = -\ln \det K + n \max\{0, \max_{1 \leq i \leq k} \{(K a_i, a_i) - 1\}\}, \quad K \succ 0. \quad (39)$$

Если $\text{co}\{M \cup 0\}$ имеет непустую внутренность, то $\det K^* > 0$. Так как $\ln \det K \rightarrow +\infty$ при $\det K \rightarrow 0$, то при монотонном движении по внутренним точкам допустимой области $K \succ 0$ целевая функция будет играть роль барьера и ограничение $K \succ 0$ будет автоматически выполняться. Матрица $K = \{k_{ij}\}_{i,j=1}^n$ имеет $n(n+1)/2$ неизвестных параметров. Если $k \ll n(n+1)/2$, то имеет смысл решать двойственную задачу [44]:

$$\max_u \left\{ \ln \det(AU A^T) : U = \text{diag } u; \quad \sum_{i=1}^k u_i = n; \quad u \geq 0 \right\}, \quad (40)$$

где $u = \{u_1, \dots, u_k\}$ суть множителя Лагранжа, относящиеся к ограничениям (38); A есть $(n \times k)$ -матрица $(n \times k)$, столбцы которой состоят из координат точек a_i , $i = 1, \dots, k$.

Так как $\ln \det(AUA^T)$ — вогнутая функция от u , то задача (38) представляет собою задачу выпуклого программирования, для решения которой можно использовать метод негладких штрафов в сочетании с r -алгоритмом.

В [44] предложен и обоснован "простой" алгоритм решения прямой задачи (37)–(38), основанный на субградиентном процессе с переменной метрикой, требующий на каждой итерации $O(nk)$ арифметических операций. При этом последовательность эллипсоидов в трансформированном пространстве представляется в виде шаров. Если центр описанного эллипсоида не фиксируется, а является переменным, то мы получаем, на первый взгляд, более сложную задачу. Однако она сводится (например, [37] и [44]) к задаче с фиксированным центром, но для точек, представленных в $(n + 1)$ -мерном пространстве путем присоединения дополнительной координаты с равным значением 1 для всех точек.

В работе [37] предложен полиномиальный алгоритм решения задачи (37)–(38) (сложности $O(k^{3.5} \ln(1/\varepsilon))$, где ε — относительная погрешность по функционалу), основанный на применении одной из модификаций метода внутренних точек (используется барьерная штрафная функция, к ней применяется квазиньютоновская процедура выбора направления спуска, а параметры штрафной функции согласованно пересчитываются).

Если $M = \text{co}\{a_1, \dots, a_k\}$ содержит 0 в качестве внутренней точки, то при полярном преобразовании $\text{ro}M = \{y : (a_i, y) \leq 1; i = 1, \dots, k\}$ минимальный по объему описанный эллипсоид с центром в точке 0 переходит в максимальный по объему вписанный в $\text{ro}M$ эллипсоид с центром в 0. Таким образом, если тело $M \in W_n$ задано в виде линейных неравенств вида $(a_i, y) \leq 1$, $i = 1, \dots, k$, то построение максимально по объему вписанного эллипсоида с центром в точке 0 сводится к задаче вида (39), которую мы умеем решать.

Если же центр вписанного эллипсоида не фиксирован, то получаем следующую задачу нелинейного программирования [44], [29]: найти

$$\min_{Q \succ 0} (\ln \det Q) \quad (41)$$

при ограничениях

$$\sqrt{(Qa_i, a_i)} + (a_i, b) - 1 \leq 0, \quad i = 1, \dots, k. \quad (42)$$

При подстановке $P = \sqrt{Q}$, $P \succ 0$, имеем задачу выпуклого программирования: найти

$$\min_{P \succ 0} (-2 \ln \det P), \quad (43)$$

$$|Pa_i| + (a_i, b) - 1 \leq 0, \quad i = 1, \dots, k, \quad (44)$$

которую можно свести к задаче минимизации штрафной функции:

$$\varphi(P, b) = -\ln \det P + n \max \left\{ 0, \max_i \{|Pa_i| + (a_i, b) - 1\} \right\}.$$

Получив оптимальные значения P^* и b^* , оптимальный вписанный эллипсоид находим по следующей формуле:

$$E(K^*, b^*) = E((P^*)^{-2}, b^*).$$

В работе Хачияна и Тодда [29] предложен полиномиальный алгоритм построения оптимального вписанного эллипсоида с относительной погрешностью ε , требующий

$$O \left(k^{3.5} \ln \left[\frac{kR}{\ln(1 - \varepsilon)} \right] \ln \left[\frac{n \ln R}{\ln(1 + \varepsilon)} \right] \right)$$

арифметических операций (здесь R — заранее заданное соотношение радиусов двух евклидовых шаров, один из которых описан вокруг многогранника, задаваемого системой неравенств, а другой в него вписан).

Суть метода состоит в том, что на каждом большом шаге невыпуклые неравенства относительно Q и b [40]

$$(Qa_i, a_i) \leq (1 - (b, a_i))^2 \quad (45)$$

заменяются линейными неравенствами вида

$$(Qa_i, a_i) \leq (1 - (b, a_i)) (1 - (\bar{c}, a_i)) \quad i = 1, \dots, k, \quad (46)$$

где \bar{c} — некоторое фиксированное приближение к b^* . После получения с определенной точностью оптимальных для задачи [39], [42] значений $Q^*(\bar{c})$ и $b^*(\bar{c})$, \bar{c} пересчитывается по простой формуле:

$$\bar{c}_{\text{нов}} = \frac{1}{2} (b^*(\bar{c}) + \bar{c}),$$

и цикл вычислений повторяется.

Алгоритм определения центра оптимального по объему вписанного в многогранник M эллипсоида служит основой одного из наиболее эффективных по числу итераций методов отсечения для минимизации негладких выпуклых функций $f(x)$ на выпуклом многограннике M_0 , задаваемом системой линейных неравенств. Этот метод предложен в работе [50] под названием метода вписанных эллипсоидов (МВЭ).

Суть этого метода состоит в построении последовательности вложенных многогранников $M_0 \supset M_1 \supset M_2 \dots \supset M_k \dots$, локализирующих минимум. Каждый следующий многогранник M_{k+1} получается из предыдущего M_k путем отсечения гиперплоскостью, нормалью которой служит субградиент $g_f(\tilde{x}_k^c)$, где \tilde{x}_k^c центр вписанного в M_k эллипсоида $E_{k+1} \subset M_k$, объем которого меньше объема оптимального вписанного эллипсоида не более, чем в $(1 + \gamma)$ раза (γ — достаточно малая положительная величина). Этот метод гарантирует уменьшение "рекордного" отклонения значения функции от оптимального со скоростью геометрической прогрессии со знаменателем $q \approx (q_0(1 + \gamma)/(1 - \gamma))^{1/n}$, где $q \approx 0.73 \dots$. Этот метод был опробован нами на ряде тестовых задач, при этом для приближенного нахождения центра оптимального вписанного эллипсоида использовался "простой" субградиентный алгоритм с трансформацией пространства [44], разработанный автором совместно с О.А.Березовским. Этот алгоритм в данном случае имел дополнительное преимущество, состоящее в том, что в трансформированном пространстве субоптимальный эллипсоид имел форму n -мерного шара. Поэтому после отсечения можно было использовать в качестве начального приближения то преобразование пространства, которое получилось на предыдущем цикле вычислений. Это значительно уменьшило время счета. Более подробно результаты тестирования изложены в [44].

Замечание. Сложность задачи определения оптимального по объему описанного (вписанного) эллипсоида существенно зависит от способа задания многогранника. Так, задача построения минимального по объему эллипсоида, описанного вокруг многогранника, задаваемого системой неравенств, является NP -трудной, как и задача определения максимального по объему вписанного эллипсоида в случае, когда многогранник задается как выпуклая оболочка системы точек. Даже более простая задача построения шара минимального радиуса, содержащего многогранник, заданный с помощью системы линейных неравенств, является NP -трудной. Также не найден эффективный алгоритм для определения минимального по объему эллипсоида, описанного вокруг пересечения двух или более эллипсоидов, и т.п.

4. Вычислительные эксперименты

Проведенный нами анализ показывает, что большое количество проблем матричной оптимизации приводит к необходимости решения выпуклых негладких экстремальных задач. Одним из наиболее практически эффективных методов решения такого рода задач является метод субградиентного типа с растяжением пространства в направлении разности двух последовательных субградиентов (r -алгоритм [53]). Приведем общую схему r -алгоритма применительно к решению задачи минимизации выпуклой функции $f(x)$, определенной на n -мерном пространстве E^n и обладающей свойством

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

При описании алгоритма будем использовать оператор $R_\alpha(\xi)$ растяжения (сжатия) пространства с коэффициентом $\alpha \geq 1$ ($0 < \alpha < 1$) в направлении ξ , $|\xi| = 1$, т.е.

$$R_\alpha(\xi)x = x + (\alpha - 1)(x, \xi)\xi.$$

Субградиент $f(\cdot)$ в точке \bar{x} будем обозначать $g_f(\bar{x})$.

0-й шаг. Имеем начальную точку $x_0 \in E^n$ и матрицу $B_0 = J_n$ размерности $n \times n$. Последующие шаги будут определяться способами вычисления последовательностей $\{h_k\}_{k=1}^\infty$ и $\{\alpha_k\}_{k=1}^\infty$.

1-й шаг. Вычисляем

$$x_1 = x_0 - h_1 \frac{g_f(x_0)}{|g_f(x_0)|},$$

т.е. 1-й шаг соответствует шагу субградиентного спуска. Последующие шаги будут происходить с учетом изменения метрики.

Пусть проделано k шагов, $k \geq 1$, и мы имеем x_k , $g_f(x_{k+1})$ и матрицу B_k размерности $n \times n$.

$k + 1$ -й шаг. Вычисляем

$$\begin{aligned} f(x_k); \quad g_f(x_k); \quad r_k = g_f(x_k) - g_f(x_{k-1}); \\ \xi_k = \frac{B_k r_k}{|B_k r_k|}; \quad B_{k+1} = B_k R_{1/\alpha_k}(\xi_k); \end{aligned}$$

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

(если $g_f(x_k) = 0$, то x_k — точка минимума, в противном случае итерации продолжаютя вплоть до получения признака окончания процедуры).

Каждый шаг r -алгоритма сопровождается получением новой порции информации $\{f(x_k); g_f(x_k)\}$ и простыми матрично-векторными вычислениями, требующими $O(n^2)$ арифметических операций. Выбор шаговых множителей $\{h_k\}$ и коэффициентов растяжения пространства $\{\alpha_k\}$ производится с помощью специальных процедур. В некоторых случаях принимается $\alpha_k = \bar{\alpha}$, $h_k = s_k \bar{h}$, s_k — некоторое натуральное число. В монографиях [48], [53], [54] описаны многочисленные успешные применения r -алгоритма при реализации схем декомпозиции, решения задач выпуклого программирования с использованием негладких штрафных функций, решения минимаксных задач, задач идентификации и т.п.

В работах [33], [31] описано сравнительное тестирование r -алгоритма и других процедур негладкой оптимизации, неизменно показывающее его высокую надежность и, как правило, уменьшение времени счета при достижении одних и тех же показателей точности. Практически при правильном выборе параметров r -алгоритм не хуже по числу итераций метода вписанных эллипсоидов [50]. В то же время он требует меньше дополнительных вычислений на каждой итерации.

Эти преимущества проявляются и при использовании r -алгоритма для решения выпуклых задач матричной оптимизации. Однако не нужно забывать о некоторых особенностях этих задач, основная из которых состоит в том, что минимизируемые выпуклые функции часто определены не на всем пространстве, а на некотором неявно определенном выпуклом подмножестве, связанном с условиями неотрицательной определенности (как, например, функции вида $-\ln \det A(x)$). Основные модификации r -алгоритма связаны с более осторожным выбором шагового множителя, чтобы не пересечь границу неотрицательной определенности. Проверка этого свойства осуществлялась, как правило, достаточно эффективно с помощью процедуры треугольного разложения симметричной матрицы по Холецкому [47]. Подробно модификации r -алгоритма применительно к матричным задачам описаны в монографии [54] (гл.4). В этой же монографии приведены результаты применения r -алгоритма для решения ряда задач матричной оптимизации. Так, используя материалы работ Р.Флетчера ([13]–[16], [18]), был проведен вычислительный эксперимент по решению задач образователь-

ного тестирования вида (15)–(16). Используя стандартную версию r -алгоритма, удалось решить с высокой точностью (6-7 точных значащих цифр по функционалу) две наиболее сложные тестовые задачи из работы [13]. Р. Флетчеру удалось получить приемлемые результаты для этих задач, лишь в максимальной степени учитывая их специфику, причем программное обеспечение было довольно громоздко и использовало ряд эвристических процедур.

В [54] описаны многочисленные эксперименты по получению оценок Ловаса в форме (18)–(20) для задачи о максимальном взвешенном внутренне устойчивом множестве графа с использованием модифицированного r -алгоритма. Успешно решены задачи получения оценки Ловаса с матрицами размера 70×70 (70 вершин графа) и числом неизвестных параметров 140 (число ребер графа). Другими методами задачи такой размерности решаются с большим трудом.

Ниже мы опишем еще не опубликованные результаты экспериментов по решению задач нахождения оптимальных параметров матрицы, обеспечивающих минимальное значение суммы нескольких наибольших собственных чисел. Исходные данные для тестирования взяты из работы [36]. Исследовательская программа разработана О. А. Березовским.

Рассматривается задача нахождения минимума суммы m наибольших собственных чисел матрицы $A(t)$ размерности 5×5 , вектор параметров $t = \{t_1, t_2, t_3, t_4, t_5\}$, параметры входят аффинно в $A(t)$:

$$A(t) = A_0 + \sum_{i=1}^5 t_i A_i.$$

Исходные данные, определяющие симметричные матрицы A_0, A_1, \dots, A_5 размера 5×5 , как и результаты решения с помощью специального метода для $m = 3$, содержатся в [36].

В соответствии с методом, описанным в §2, мы свели эту задачу к максиминной задаче нахождения седловой точки: найти

$$\max_{\{C : \text{tr} C = m\}} \min_t \left[\left(A_0 + \sum_{i=1}^5 t_i A_i, C \right) S(\ln \det C + \ln \det (J - C)) + \varepsilon \sum_{i=1}^5 t_i^2 \right],$$

где S – штрафной множитель, ε^* – сглаживающий множитель (по t). Внутренняя задача минимизации по t при фиксированном C решалась приравниванием частных производных по t_i к нулю. Для решения внешней задачи максимизации по C использовался r -алгоритм с модифицированным выбором шагового множителя.

Таблица 1.

m	S	ε	f	k	t^*	λ^*
1	10^{-3}	0.01	0.7067	150	-0.6126	0.7092
		0.001	0.7062	150	0.6504	0.7089
	10^{-5}	0.01	0.70912	230	-0.3242	0.0294
		0.001	0.7088	220	-0.5918	-0.6351
10^{-7}	0.01	0.709165	240	0.5891	-1.4114	
	0.001	0.7089	220	$f^* = 0.709165$		
2	10^{-3}	0.01	1.2058	130	-0.3108	-0.8370
		0.001	1.2053	150	0.1551	0.3713
	10^{-5}	0.01	1.20843	220	-0.1456	-0.1633
		0.001	1.20796	170	-0.0888	-0.3720
10^{-7}	0.01	1.208471	210	0.8990	-0.4203	
	0.001	1.20807	210	$f^* = 1.208471$		
3	10^{-3}	0.01	0.7128	180	-0.7555	0.9654
		0.001	0.7126	190	0.2960	0.4024
		0.0	0.7124	130	-0.0049	-0.6523
	10^{-5}	0.01	0.715326	180	-0.0369	-0.6526
		0.001	0.715322	200	0.0527	-0.6973
		0.0	0.715321	160	$f^* = 0.7153572$	
	10^{-7}	0.01	0.7153572	230		
		0.001	0.7153572	210		
0.0		0.7153570	170			
4	10^{-3}	0.01	0.0183	160	-0.9105	1.2083
		0.001	0.0182	180	-0.0915	0.3042
	10^{-5}	0.01	0.020297	220	0.2060	-0.7455
		0.001	0.02028	210	0.0901	-0.7458
	10^{-7}	0.01	0.020321	220	0.0776	-0.7465
		0.001	0.02031	210	$f^* = 0.020321$	

Результаты расчетов приведены в таблице 1 (здесь k – число итераций r -алгоритма, оптимизация по C проходила по $\frac{5*6}{2} - 1 = 14$ параметрам, $c_{55} = m - \sum_{i=1}^4 c_{ii}$, так как $trC = m$. При достаточно малых S и ε для всех m ($m = 1, 2, 3, 4$) получено хорошее совпадение по функционалу с оптимальным значением f^* (6-7 значащих цифр).

Таким образом, вычислительные эксперименты показывают, что современные методы негладкой оптимизации являются достаточно универсальным и гибким средством решения матричных экстремальных задач.

Отметим, что в r -алгоритме предыдущая информация запоминается лишь в форме матрицы преобразования пространства. В настоящее время нами разработаны новые алгоритмы с преобразованием пространства, учитывающие предыдущую информацию в форме кусочно-линейной аппроксимации снизу минимизируемой функции, как это делается в агрегатных ε -субградиентных методах [30]. Это позволит еще более ускорить сходимость методов негладкой оптимизации. Кроме того, оценки целевой функции снизу позволят более точно контролировать отклонение от оптимума рекордного значения целевой функции.

Заключение

Уже после сдачи статьи в печать автор побывал на XV Международном симпозиуме по математическому программированию (Энн Арбор, США). Тематика, связанная с матричной параметрической оптимизацией, там была представлена очень широко. Особенно много докладов было посвящено применению методов внутренних точек для решения этих задач. Подробный обзор новых работ потребовал бы отдельной большой статьи. Здесь автор только хотел бы выразить надежду, что здоровая конкуренция между методами внутренних точек и методами негладкой оптимизации при решении экстремальных матричных задач будет способствовать развитию обоих этих направлений. Обращаем также внимание читателей на две интересные монографии [57] и [58], вышедшие совсем недавно.

Литература

1. ALIZADEH F. Combinatorial optimization with semidefinite matrices. – In: Proceedings of the Second Annual Integer Programming and Combinatorial Optimization Conference. Carnegie-Mellon University, 1991.
2. BARNES E. R. An algorithm for partitioning the nodes of a graf. – SIAM J. Alg. and Disc. Math., 1982, v. 3.
3. BARNES E. R. An algorithm for separating patterns by ellipsoids. – IBM J. Res. Dev., 1982, v. 26, № 6.
4. BARNES A. E., HOFMAN A. J. Partitioning, spectra, and linear programming. – In: Progress in Combinatorial Optimization. New York: Academic Press, 1984.
5. BECKENBACH E. F., BELLMAN R. Inequalities. Berlin: Springer-Verlag, 1961.
6. BHATIA R. Perturbation bounds for matrix eigenvalues. – Pitman Research Notes in Mathematics series, 1987, v. 162.
7. CHU K. E. On multiple eigenvalues of matrices depending on several parameters. – SIAM J. Numerical Anal., 1990, v. 5, p. 1368–1385.
8. CULLUM J., DONATH W. E., WOLFE P. The minimization of certain nondifferentiable sums of eigenvalues of symmetric matrices. – Math. Prog. Study, 1975, v. 3, p. 35–55.
9. DONATH W. E., HOFFMAN A. J. Algorithms for partitioning graphs and computer logic based on eigenvector of connection matrices. – IBM Tech. Disclosures BuL, 1972, v. 15.
10. DONATH W. E., HOFFMAN A. J. Lower bounds for the partitioning of graphs. – IBM Res. Dev., 1973, v. 17.
11. FAN M. K. H. A quadratically convergent local algorithm on minimizing the largest eigenvalue of a symmetric matrix. – Linear Algebra and Its Applications on Numerical Linear Algebra Methods in Control, 1993, v. 188–189, p. 207–230.
12. FAN M. K. H., НЕКООИЕ B. On minimizing the largest eigenvalue of a symmetric matrix. – Linear Algebra and Its Applications on Numerical Linear Algebra Methods in Control.

13. FLETCHER R. A nonlinear programming problem in statistics (educational testing). – SIAM J. Sci. Statist. Comput., 1981, v. 2, p. 257-267.
14. FLETCHER R. Numerical experiments with an exact l_1 penalty function method. – In: Nonlinear Programming. / Ed. by O.L.Mangasarian, R.R.Meyer and S.M.Robinson. New York: Academic Press, 1981.
15. FLETCHER R. Second order corrections for nondifferentiable optimization. – Lect. Notes Math., 1982, v. 912.
16. FLETCHER R. Semi-definite matrix constraints in optimization. – Numerical Analysis Report NA/61, Univ. Dundee, Dept. Math. Sci., Scotland, 1982.
17. FILMORE P.A., WILLIAMS J.P. Some convexity theorems for matrices. – Glasgow Math. J., 1971, v. 12, p. 110–117.
18. FLETCHER R. Semidefinite matrix constraints in optimization. – SIAM J. Control Optim., 1985, v. 23, p. 493–513.
19. FRENCH R. M., ORLIN J. B. On the complexity of four polyhedral set containment problems. – Working papers, January 1984, №1527-84, MIT.
20. FRIEDLAND S., NOCEDAL J., OVERTON M.L. The formulation and analysis of numerical methods for inverse eigenvalue problems. – SIAM J. Numerical Anal., 1987, v. 24, p. 634–667.
21. GOLLAN B. Eigenvalue perturbation and nonlinear parametric optimization. – Math. Programming Study, 1987, v. 30, p. 67–81.
22. GROTSHEL M., LOVASZ L., SCHRIJVER A. Geometric Algorithm and Combinatorial Optimization. Berlin: Springer-Verlag, 1988, 362 p.
23. HAUG E.J., ROUSSELET B. Design sensitivity analysis in structural mechanics. II: Eigenvalue variation. – J. Structural Mechanics, 1980, v. 8, p. 161–186.
24. HIRIART-URRUTY J.B., SEEGER A., YE D. Sensitivity analysis for a class of convex functions defined over a space of symmetric matrices. – Lect. Notes Economics and Math. Systems, 1992, v. 382, p. 133–154.

25. HIRIART-URRUTY J.B., YE D. Sensitivity analysis of the sum of the largest eigenvalues of a symmetric matrix. – Preliminary research report, Laboratoire d'Analyse Numerique, Universite Paul Sabatier, Toulouse, June 1991.
26. HIRIART-URRUTY J.B., LEMARECHAL C. Convex Analysis and Minimization Algorithms. Vol. 1-11. Berlin: Springer-Verlag, 1994.
27. HIRIART-URRUTY J.B., YE D. Sensitivity analysis of all eigenvalues of a symmetric matrix. – In: Seminaire d'Analyse Numerique. Universite Paul Sabatier, 1992–1993, v. 11, p. 33.
28. JAYARAYAN S. A nonlinear optimization problem in educational testing. – MSc thesis, Dept. Math., Univ. Dundee, Scotland, 1979.
29. KHACHIYAN L., TODD M. On the complexity of approximating the maximal inscribed ellipsoid for a polytope. – Technical report N 893, Ithaka, New York: Cornell Univ., 39 p.
30. KIEWIEL K.C. Methods of Descent for Nondifferentiable Optimization. – Lect. Notes Math., 1985, v. 1133, p. 1–360.
31. LEMARECHAL C. Numerical experiments in nondifferentiable optimization.–CP-82- 58, IIASA, Laxenburg, 1982, p. 61–84.
32. LOVASZ L. On the Shannon capacity of a graph. – IEEE Trans. Inform. Theory, 1979, v. 25, p. 1–7.
33. MAIER R.S. Computational Results with Nonsmooth methods for Large-Scale nonlinearly-constructed optimization. – Army high performance computing research center, University of Minnesota, Preprint, 91-65, 1991, 36 p.
34. MAYNE D. Q., POLAK E. Algorithm for the design of control systems subject to singular value inequalities. – In: Algorithms and Theory in Filtering and Control. / Ed. by D.C.Sorensen and R.J-B.Wets. (Mathematical Programming Study, v. 18.) Amsterdam: North-Holland, 1982.
35. NARASIMHAN G., MANBER R. A generalization of Lovasz's sandwich theorem. In: Polyhedral Combinatorics. / Ed. by W.Cook and P. D. Seymour. 1990, p. 19–27.

36. NEKOOIE B., FAN M.K.H. A quadratically convergent local algorithm on minimizing sums of the largest eigenvalue of a symmetric matrix. – Proceedings of 31th IEEE Conference on Decision and Control. Tucson, Arizona, December 1992, p. 1915–1920.
37. NESTEROV J.E., NEMIROVSKY A.S. Self-concordant functions and polynomial-time methods in convex programming. Moscow: USSR Academy of Sciences, Central Econom. and Math. Institute, 1989.
38. OVERTON M.L. On minimizing the maximum eigenvalue of a symmetric matrix. – SIAM J. Matrix Anal. Appl., 1988, v. 9, p. 256–268.
39. OVERTON M.L., WOMERSLEY R.S. On the sum of the largest eigenvalues of a symmetric matrix. – SIAM J. Matrix Anal. Appl., 1992, v. 13, p. 41–45.
40. OVERTON M.L., WOMERSLEY R.S. Optimality conditions and duality theory for minimizing sums of the largest eigenvalues of symmetric matrices. – Math. Program., 1993, v. 62, p. 321–357.
41. POLAK E., WARDI Y. A nondifferentiable optimization algorithm for structural problems with eigenvalue inequality constraints. – J. Structural Mechanics, 1983, v. 11, p. 561–577.
42. RENEGAR J. A polynomial-time algorithm, based on Newton's method, for linear programming. – Math. Program., 1988, v. 40.
43. SHAPIRO A., FAN M.K.H. Generic analysis of optimization problems involving eigenvalues of symmetric matrices. (To appear.)
44. SHOR N.Z., BEREZOVSKI O.A. New algorithms for constructing optimal circumscribed and inscribed ellipsoids. – Optim. Methods Software, 1992, v. 1, p. 283–299.
45. TITTERINGTON D.M. Optimal design: some geometric aspects of D-optimality. – Biometrika, 1975, v. 62.
46. TSING N.K., FAN M.K.H. On analyticity of functions involving eigenvalues. – Linear Algebra Appl. (To appear.)
47. ГАНТМАХЕР Ф.Р. Теория матриц. М.: ГИТТЛ, 1953, 491 с.
48. МИХАЛЕВИЧ В.С., ТРУБИН В.А., ШОР Н.З. Оптимизационные задачи производственно-транспортного планирования. М.: Наука, 1980, 260 с.

49. НЕМИРОВСКИЙ А. С., ЮДИН Д. Б. Сложность задач и эффективность методов оптимизации. М.: Наука, 1979, 383 с.
50. ТАРАСОВ С. П., ХАЧИЯН Л. Г., ЭРЛИХ И. И. Метод вписанных эллипсоидов. – Докл. АН СССР, 1988, т. 298, №5, с. 1081–1085.
51. ХУСАИНОВ Д. Я., ИВОХИН Е. В. Об оценке решений линейных систем с использованием функций Ляпунова. – Кибернетика, 1985, № 2, с. 7–10.
52. ЧЕРНОУСЬКО Ф. Л. Оценивание фазового состояния динамических систем. Метод эллипсоидов. М.: Наука, 1988, 320 с.
53. ШОР Н. З. Методы минимизации недифференцируемых функций и их приложения. Киев: Наукова думка, 1979, 200 с.
54. ШОР Н. З., СТЕЦЕНКО С. И. Квадратичные экстремальные задачи и недифференцируемая оптимизация. Киев: Наукова думка, 1989, 204 с.
55. ХАРДИ Г. Г., ЛИТТЛВУД Д. Е., ПОЛИА Г. Неравенства. М.: ГИИЛ, 1948, 456 с.
56. БЕРЖ К. Теория графов и ее приложения. М.: ИЛ, 1962, 319 с.
57. NESTEROV YU. AND NEMIROVSKII A. Interior point polynomial methods in convex programming. Philadelphia: SIAM, 1994.
58. BOYD S., GHAOUI L., FERON E., BALAKRISHNAN V. Linear matrix inequalities in system and control theory (Final Draft). Febr. 1994, 199 p.

Использование модификации r -алгоритма для нахождения глобального минимума полиномиальных функций

Н.З. Шор, П.И. Стецюк

Кибернетика и системный анализ. – 1997. – № 4. – С. 28–49.

Определение глобального минимума полинома от нескольких переменных – весьма сложная задача уже для полиномов четвертой степени при числе переменных более 5. Так, количество локальных минимумов для полиномов четвертой степени может достигать нескольких десятков уже при $n = 6$. Один из способов решения таких задач состоит в сведении их к задачам нелинейного программирования с квадратичными целевыми функциями и квадратичными ограничениями [1, 2]. Используя функцию Лагранжа, можно получать двойственные квадратичные оценки, которые являются оценками снизу для глобального минимума полинома. Эти оценки могут быть улучшены за счет введения избыточных ограничений [3, 4, 2]. Упомянутые проблемы тесно связаны с исследованиями Гильберта о представлении неотрицательных полиномов в виде суммы квадратов полиномов меньших степеней [5].

В настоящей статье рассмотрены свойства квадратичных оценок для полиномиальных функций (разд. 2.1), алгоритмы нахождения этих оценок, основанные на методах негладкой оптимизации (разд. 2.2). Представлены результаты тестовых экспериментов для экстремальных задач, которые описываются в виде моделей нелинейного программирования квадратичного типа и для которых использование избыточных квадратичных ограничений в определенных случаях значительно улучшает двойственные оценки (разд. 2.3).

1. Двойственные квадратичные оценки для полиномиальных функций

Рассмотрим полиномиальную функцию от n вещественных переменных $P(y) = P(y_1, \dots, y_n)$, $y = (y_1, \dots, y_n) \in L^n$, L^n – n -мерное линейное пространство.

Ставится задача нахождения

$$P^* = \inf_{y \in L^n} P(y).$$

Очевидно, чтобы P^* было отлично от $-\infty$, необходимо следующее: старшие степени s_i полинома $P(y_1, \dots, y_n)$ по каждой переменной y_i , $i = 1, \dots, n$, должны быть четными и при любых значениях других переменных, коэффициент при $y_i^{s_i}$ — неотрицательным. Пусть $s_i = 2l_i$, $i = 1, \dots, n$. Обозначим через $d(P)$ максимальную тотальную степень монома, входящего в P , а через $\pi(n; s; d(P))$ (соответственно $\bar{\pi}(n; s; d(P))$) множество всех полиномов (соответственно ограниченных снизу полиномов) от n переменных степени s_i , $i = 1, \dots, n$ по каждой переменной y_i , с максимальной степенью мономов $d(P)$. Здесь $s = \{s_i\}_{i=1}^n$.

Пусть $\alpha = (\alpha_1, \dots, \alpha_n)$ — неотрицательный целочисленный вектор. Символом $R[\alpha] = R[\alpha_1, \dots, \alpha_n]$ обозначим моном вида

$$R[\alpha](y) = y_1^{\alpha_1}, \dots, y_n^{\alpha_n} = \prod_{i=1}^n y_i^{\alpha_i},$$

а символом $c_p[\alpha]$ (или просто $c[\alpha]$) — коэффициент при мономе $R[\alpha]$ в полиноме $P(y)$.

В работах [4, 6, 2] предложен способ определения глобального минимума полиномов, основанный на сведении задачи нахождения P^* к задаче квадратичного вида и нахождении Лагранжевых двойственных оценок для полученной квадратичной задачи (при этом решающую роль играют функционально избыточные ограничения).

Пусть $P(\cdot) \in \bar{\pi}(n; s; d(P))$, причем $s_i = 2l_i$, $i = 1, \dots, n$. Пусть также $\bar{P}(y) = P(y) - P^* \geq 0$ для всех $y \in L^n$. Попытаемся представить $\bar{P}(y)$ в виде суммы квадратов вещественных полиномов. Если это возможно, то старшие степени полиномов, входящих в это представление, не должны превышать l_i для каждой переменной y_i ($i = 1, \dots, n$), а их тотальная степень не должна превышать $\frac{d(P)}{2}$, т.е. если

$$\bar{P}(y) = \sum_{k=1}^N P_k^2(y), \text{ то}$$

$$P_k(y) \in \pi(n; l; \frac{d(P)}{2}), k = 1, \dots, N; \quad l = (l_1, \dots, l_n).$$

Сведение задачи нахождения P^* к квадратичной задаче основано на разложении каждого монома $R[\alpha](y) \in \pi(n; l; \frac{d(P)}{2})$ в произведение не

более двух мономов, принадлежащих $\pi(n; l; \frac{d(P)}{2})$. Это разложение может быть неоднозначным. Рассмотрим так называемое **каноническое** разложение.

Пусть $R[\alpha](y) = \prod_{i=1}^n y_i^{\alpha_i}$, $\alpha_i \leq 2l_i$; $i = 1, \dots, n$; $\sum_{k=1}^n \alpha_k \leq d(P)$.

Тогда

$R[\alpha](y) = R[\beta^{(1)}](y) \cdot R[\beta^{(2)}](y)$, причем

$$\beta_i^{(1)} = \begin{cases} \frac{\alpha_i}{2} & \text{для четных } i, \\ \frac{\alpha_i \pm 1}{2} & \text{для нечетных } i, \end{cases} \quad (1)$$

$$\beta_i^{(2)} = \alpha_i - \beta_i^{(1)}.$$

(знаки $+$, $-$ в формуле чередуются, чтобы тотальная степень мономов $R[\beta^{(1)}]$ и $R[\beta^{(2)}]$ не превышала $\frac{d(P)}{2}$). Разложение (1) назовем каноническим.

Выберем в качестве новых переменных мономы $R[\beta]$, где β принадлежит подмножеству целочисленных n -мерных неотрицательных векторов таких, что $R[\beta](\cdot) \in \pi(n; l; \frac{d(P)}{2})$.

Тогда полиному $\bar{P}(y) \in \bar{\pi}(n; 2l; d(P))$ соответствует квадратичная форма

$$\bar{P}(y) \sim \sum_{\alpha \in A} c_{\bar{P}}[\alpha] R[\beta^{(1)}] R[\beta^{(2)}],$$

где $R[\beta^{(1)}], R[\beta^{(2)}]$ получаются из канонического разложения $R[\alpha](\cdot)$, множество A состоит из векторов α таких, что мономы $R[\alpha]$ входят с ненулевыми коэффициентами в $\bar{P}(y)$.

Переменные $R(\beta)$ связаны между собой системой (Q) следующих квадратичных соотношений:

$$K_{\alpha}(\rho) = K_{\beta^{(3)}, \beta^{(4)}}(\rho) = R[\beta^{(1)}] \cdot R[\beta^{(2)}] - R[\beta^{(3)}] \cdot R[\beta^{(4)}] = 0. \quad (2)$$

Здесь $\beta^{(1)}, \beta^{(2)}$ соответствуют каноническому разложению $R[\alpha] = R[\beta^{(1)} + \beta^{(2)}]$; $\beta^{(3)}, \beta^{(4)}$ удовлетворяют следующим условиям: $\beta^{(3)} + \beta^{(4)} = \alpha$; $(\beta^{(3)} \geq \beta^{(4)})$ лексикографически; $R[\beta^{(3)}], R[\beta^{(4)}] \in \pi(n; l; \frac{d(P)}{2})$; $\alpha \in M(n; 2l; d(P))$, где $M(n; 2l; d(P))$ множество неотрицательных n -мерных векторов таких, что $R[\alpha](y) \in \pi(n; 2l; d(P))$. Совокупность всевозможных допустимых соотношений вида (2) обозначим $Q[n; 2l; d(P)]$, множество допустимых индексов β переменных $\{R(\beta)\}$ обозначим $\bar{M}[n; l; \frac{d(P)}{2}]$, квадратичный полином от переменных $R[\beta]$, стоящий в левой части тождества (2), обозначим $K_{\beta^{(3)}, \beta^{(4)}}(\rho)$ (здесь ρ – вектор переменных $\{R(\beta)\}_{\beta \in \bar{M}[n, l, \frac{d(P)}{2}]}$).

Пусть $P(y) = \sum_{\alpha \in A} c_p[\alpha] \cdot R[\beta^{(1)}](y) \cdot R[\beta^{(2)}](y)$, где $R[\beta^{(1)}](y) \cdot R[\beta^{(2)}](y)$ - каноническое разложение $R[\alpha](y)$. Так как соотношения (2) являются тождественными для $y \in L^n$, то

$$P(y) = \sum_{\alpha \in A} c_p[\alpha] \cdot R[\beta^{(1)}](y) \cdot R[\beta^{(2)}](y) + \sum_{(\beta^{(3)}, \beta^{(4)}) \in Q[n; 2l; d(P)]} \lambda[\beta^{(3)}, \beta^{(4)}] K_{\beta^{(3)}, \beta^{(4)}}[\rho](y) \quad (3)$$

при произвольных множителях $\lambda[\beta^{(3)}, \beta^{(4)}]$ для всех $y \in L^n$.

Задаче нахождения $P^* = \inf_{y \in L^n} P(y)$ можно сопоставить эквивалентную в определенном смысле задачу квадратичного типа: найти

$$\inf K_0(\rho) = \inf \sum_{\alpha \in A} c_p[\alpha] R[\beta^{(1)}](y) \cdot R[\beta^{(2)}] \quad (4)$$

при ограничениях

$$K_{\beta^{(3)}, \beta^{(4)}}(\rho) = 0 \text{ для всех } (\beta^{(3)}, \beta^{(4)}) \in Q(n; 2l; d(P)). \quad (5)$$

Рассмотрим характерные примеры.

1. $n = 1$; $2l = 6$ ($l = 3$); $d(P) = 6$. Речь идет о нахождении глобального минимума полинома 6-й степени от одной переменной:

$$P(y_1) = c[6]y_1^6 + c[5]y_1^5 + c[4]y_1^4 + \dots = \sum_{i=1}^6 c[i]y_1^i.$$

Мономы y_1^i будем обозначать $R[i](y_1)$ ($i = 0, 1, \dots, 6$). В соответствии с общей конструкцией введем новые переменные: $R[0], R[1], R[2], R[3]$. Получим следующие канонические разложения:

$$R[6] = (R[3])^2; \quad R[5] = R[3] \cdot R[2]; \quad R[4] = (R[2])^2;$$

$$R[3] = R[2] \cdot R[1]; \quad R[2] = (R[1])^2; \quad R[1] = R[1] \cdot R[0] = R[1].$$

Введем обозначения: $x_i = R[i]$, $i = 1, 2, 3$. Тогда задача нахождения $\inf P(y_1)$ сводится к задаче квадратичного типа от трех переменных: найти

$$\min K_0(x) = \min(c[6]x_3^2 + c[5]x_3x_2 + c[4]x_2^2 + c[3]x_2x_1 + c[2]x_1^2 + c[1]x_1 + c[0]) \quad (6)$$

при ограничениях

$$K_1(x) = x_1^2 - x_2 = 0, \quad (7)$$

$$K_2(x) = x_2x_1 - x_3 = 0, \quad (8)$$

$$K_3(x) = x_2^2 - x_3x_1 = 0. \quad (9)$$

Очевидно, что задача нахождения $\min P(y_1)$ эквивалентна в обычном смысле задаче минимизации $K_0(x)$ при ограничениях (7), (8). Ограничение (9) является избыточным, но оно оказывается целесообразным для улучшения двойственных оценок квадратичной задачи.

Функции $K_i(x)$, $i = 1, 2, 3$ при подстановках $x_2 = y_1^2$, $x_3 = y_1^3$, $x_1 = y_1$ тождественно равны нулю для произвольного y_1 . Поэтому $L(x, \lambda) = K_0(x) + \sum_{i=1}^3 \lambda_i K_i(x)$ после указанных подстановок преобразуется в $P_1(y)$ при произвольных λ_i , $i = 1, 2, 3$. С другой стороны, $L(x, \lambda)$ можем рассматривать как функцию Лагранжа задачи нелинейного программирования квадратичного типа (6)–(9). Этой задаче соответствуют маргинальная функция

$$\psi(\lambda) = \inf_x L(x, \lambda)$$

и двойственная оценка

$$\psi^* = \sup_{\lambda \in \text{dom } \psi} \psi(\lambda)$$

(если $\text{dom } \psi = \emptyset$, то будем считать, что $\psi^* = -\infty$). Если $\psi^* > -\infty$, то мы получаем нетривиальную оценку снизу для P^* :

$$\psi^* \leq P^*.$$

Оказывается, для нашего примера при $c[6] > 0$, $\psi^* = P^*$ для любого полинома шестой степени [3].

Рассмотрим общий случай (4), (5). Выражение

$$L(\rho, \lambda) = K_0(\rho) + \sum_{(\beta^{(3)}, \beta^{(4)}) \in Q[n; 2l; d(P)]} \lambda[\beta^{(3)}, \beta^{(4)}] K_{\beta^{(3)}, \beta^{(4)}}(\rho)$$

можно рассматривать как функцию Лагранжа нелинейной задачи квадратичного программирования (4), (5).

Маргинальная функция $\psi(\lambda) = \inf_{\rho} L(\rho, \lambda)$ либо имеет пустую область определения, либо является вогнутой, при этом $\text{dom } \psi$ – выпуклое множество.

Пусть $\psi^* = \sup_{\lambda \in \text{dom } \psi} \psi(\lambda)$ (если $\text{dom } \psi = \emptyset$, считаем, что $\psi^* = -\infty$). Очевидно, что $\psi(\lambda) \leq P^*$ для любого $\lambda \in \text{dom } \psi$, поэтому $\psi^* \leq P^*$.

Большой интерес представляют условия, при которых $\psi^* = P^*$. Как оказалось, этот вопрос тесно связан с ранними работами знаменитого Давида Гильберта по представлению неотрицательных полиномов в виде суммы квадратов вещественных полиномов [7] и последующими исследованиями по 17-ой проблеме Гильберта о представлении неотрицательных рациональных функций в виде суммы квадратов рациональных вещественных функций [8, 9, 5].

В [4] доказана следующая теорема (см. также [2, 10]).

Теорема 1. Пусть $P(y) \in \overline{\pi}(n, s, d(P))$ и $\inf_{y \in L^n} P(y) = P^* > -\infty$. Тогда двойственная Лагранжева оценка ψ^* , полученная для соответствующей задачи квадратичного типа (4), (5) равна P^* тогда и только тогда, когда неотрицательный полином $\overline{P}(y) = P(y) - P^*$ представим в виде суммы квадратов вещественных полиномов, принадлежащих классу $\pi(n, l, \frac{d(P)}{2})$.

Д. Гильберт рассматривал однородные полиномиальные формы степени $2l$ от n переменных и поставил вопрос, для каких комбинаций l и n любой однородный неотрицательный полином степени $2l$ от n переменных разложим в сумму квадратов вещественных полиномиальных форм степени l . Он выделил 3 случая, два из которых почти тривиальны.

1. $n = 2$; l - произвольное натуральное число (класс $H(2; 2l)$, т.е. однородные формы от двух переменных четной степени $2l$). Этот класс полиномов по своим свойствам подобен неоднородным полиномам от одной переменной степени $2l$. Легко показать, что любой неотрицательный полином из класса $H(2, 2l)$ может быть представлен в виде суммы квадратов не более двух вещественных полиномов.

2. n - произвольно, $l = 1$ (класс $H(n, 2)$, т.е. класс квадратичных форм от n переменных). Известно, что любую неотрицательно определенную форму от n переменных можно представить в виде суммы не более n квадратов линейных форм.

3. Однородные полиномы от трех переменных четвертой степени (класс $H(3, 4)$). Гильберт показал, что неотрицательные полиномы из класса $H(3, 4)$ могут быть представлены в виде суммы не более трех квадратов квадратичных форм. Доказательство этого факта, данное Гильбертом, довольно сложно. Приведем предложенный нами вариант доказательства ослабленной теоремы Гильберта.

Теорема 2. Любой неотрицательный полином из класса $H(3, 4)$ представим в виде суммы квадратов квадратичных форм.

Доказательство. Совокупность полиномов $H(3, 4)$ можно рассматривать как 15-мерное линейное пространство, поскольку полином такого вида содержит 15 параметров-коэффициентов. Любой элемент $H(3, 4)$ может быть записан в следующем виде:

$$\begin{aligned}
 P(x) = P(x_1, x_2, x_3) = & \\
 & a_{400}x_1^4 + a_{040}x_2^4 + a_{004}x_3^4 + a_{310}x_1^3x_2 + a_{301}x_1^3x_3 + \\
 & a_{130}x_1x_2^3 + a_{103}x_1x_3^3 + a_{031}x_2^3x_3 + a_{013}x_2x_3^3 + a_{220}x_1^2x_2^2 + \\
 & a_{202}x_1^2x_3^2 + a_{022}x_2^2x_3^2 + a_{211}x_1^2x_2x_3 + \\
 & a_{121}x_1x_2^2x_3 + a_{112}x_1x_2x_3^2. \tag{10}
 \end{aligned}$$

Назовем неотрицательный полином из $H(3, 4)$ 3-нулевым, если существуют три линейно независимых вектора η_1, η_2, η_3 такие, что $P(\eta) = 0, i = 1, 2, 3$. Перейдем к новой системе координат, в которой η_1, η_2, η_3 являются осями координат. Обозначим $y = (y_1, y_2, y_3)$ координаты точек в новой системе координат, причем x связан с y линейными соотношениями:

$$x_1 = (l_1, y); \quad x_2 = (l_2, y); \quad x_3 = (l_3, y),$$

т.е. $x = Ay$, где A - невырожденный линейный оператор. После подстановки $x = Ay$ в $P(x)$ мы получим полином $\bar{P}(y) = P(Ay)$.

Ясно, что полином $\bar{P}(y)$ также принадлежит классу $H(3, 4)$ и неотрицателен. Обозначим его коэффициенты при мономах $x_1^i x_2^j x_3^k$ в виде \bar{a}_{ijk} ($i, j, k \in \{0, 1, 2, 3, 4\}, i + j + k = 4$).

Поскольку $\bar{P}(1, 0, 0) = \bar{P}(0, 1, 0) = \bar{P}(0, 0, 1) = 0$, то $\bar{a}_{400} = \bar{a}_{040} = \bar{a}_{004} = 0$. Далее, с учетом условия неотрицательности \bar{P} , получаем, что \bar{a}_{310} , а также все коэффициенты вида \bar{a}_{ijk} , где i, j, k - всевозможные перестановки индексов 3, 1, 0, равны нулю. Таким образом, $\bar{P}(y)$ имеет вид:

$$\bar{P}(y) = \bar{a}_{220}y_1^2y_2^2 + \bar{a}_{202}y_1^2y_3^2 + \bar{a}_{022}y_2^2y_3^2 + \bar{a}_{211}y_1^2y_2y_3 + \bar{a}_{121}y_1y_2^2y_3 + \bar{a}_{112}y_1y_2y_3^2.$$

Выполним следующие подстановки:

$$z_1 = y_1y_2; \quad z_2 = y_1y_3; \quad z_3 = y_2y_3.$$

Тогда $\bar{P}(y) = \bar{a}_{220}z_3^2 + \bar{a}_{202}z_2^2 + \bar{a}_{022}z_3^2 + \bar{a}_{211}z_1z_2 + \bar{a}_{121}z_1z_3 + \bar{a}_{112}z_2z_3 = K(z)$, где $K(z)$ - квадратичная форма от переменных $z = (z_1, z_2, z_3)$.

Легко доказывается следующая лемма.

Лемма 1. *Полином $\overline{P}(y)$ принимает неотрицательные значения для любого $y \in L^3$ только в том случае, когда соответствующая квадратичная форма $K(z)$ является неотрицательно определенной.*

Теорема 3. *Любой неотрицательный 3-нулевой полином $P(y)$, принадлежащий $H(3, 4)$, представим в виде суммы не более трех полиномов.*

Доказательство. Из леммы 1 следует, что если $P(y) - 3$ -нулевой, и $P(y) \geq 0$ для всех y , то соответствующая квадратичная форма $K(z)$ является неотрицательно определенной, а значит, она представима в виде суммы квадратов не более трех линейных форм. При подстановке в $K(z)$ полиномов от y вместо переменных z_i полином $P(y)$ выражается в виде суммы не более трех квадратов квадратичных форм вида $a_1 x_1 x_2 + a_2 x_1 x_3 + a_3 x_2 x_3$.

Теорема доказана.

Лемма 2. *Пусть $P(y) \in H(3, 4)$, неотрицателен, но не является 3-нулевым и размерность линейной оболочки векторов $\{y : P(y) = 0\}$ равна k , где $k < 3$. Тогда его можно представить в виде $P(y) = \overline{P}(y) + (\eta, y)^4$, $\eta \in L^3$, причем линейная оболочка множества $\{y : \overline{P}(y) = 0\}$ имеет размерность больше k .*

Доказательство. Пусть $Y_P[0] = \{y : \overline{P}(y) = 0\}$ и размерность $\text{Lin}Y_P(0) = k < 3$ (Lin - символ линейной оболочки). Тогда существует вектор $\eta \neq 0$ такой, что $(\eta, y) = 0$ для всех $y \in \text{Lin}Y_P(0)$.

Рассмотрим параметрическое множество полиномов вида:

$$P_t(y) = P(y) - t(\eta, y)^4.$$

Ясно, что $P_t(y) = P(y)$ для $y \in \text{Lin}Y_P(0)$. Очевидно, что существуют $\bar{t} > 0$, и $\bar{y} \in \text{Lin}Y_P[0]$ такие, что $P_{\bar{t}}(y) \geq 0$ для всех y и $P(\bar{y})$. Обозначим $P_{\bar{t}}(y)$ как $\overline{P}(y)$. Так как $Y_{\overline{P}}(0) \supseteq Y_P(0) \cup \bar{y}$, $\bar{y} \notin \text{Lin}Y_P(0)$, то размерность $\text{Lin}Y_{\overline{P}}(0)$ больше или равна $(k + 1)$.

Лемма доказана.

Таким образом, любой неотрицательный полином из $H(3, 4)$ может быть представлен в виде суммы 3-нулевого полинома и нескольких полиномов, представляющих четвертые степени линейных функций.

Из теоремы 3 и леммы 2 непосредственно следует, что любой неотрицательный однородный полином $P(x)$ от трёх переменных степени четыре представим в виде суммы квадратов квадратичных форм.

Теорема 2 доказана.

2. Алгоритм нахождения двойственной оценки задачи квадратичного типа

Рассмотрим задачу квадратичного типа: найти

$$\inf K_0(x), \quad x \in L^n \quad (11)$$

при ограничениях

$$K_i(x) = 0, \quad i = 1, \dots, m, \quad (12)$$

где $K_\nu(x)$ – квадратичные или линейные функции, $\nu = 0, 1, \dots, m$.

Пусть $L(x, u) = K_0(x) + \sum_{i=1}^m u_i K_i(x)$ – функция Лагранжа задачи (11), (12), которую можно записать в следующем виде:

$$L(x, u) = (Q(u)x, x) + (l(u), x) + c(u),$$

где $(Q(u)x, x)$ – квадратичная по x часть функции Лагранжа, $(l(u), x)$ – линейная по x функция, $c(u)$ от x не зависит. Пусть $\psi(u) = \inf_x L(x, u)$. $\psi(u)$ – вогнутая функция, не обязательно собственная. Если $\text{dom } \psi \neq \emptyset$, $\bar{u} \in \text{dom } \psi$ только в том случае, если $Q(\bar{u})$ неотрицательно определена по x и система линейных уравнений

$$2Q(\bar{u})x + l(\bar{u}) = 0 \quad (13)$$

разрешима по x .

Пусть $\Omega = \{u : Q(u) \succ 0\}$, $\bar{\Omega} = \{u : Q(u) \succeq 0\}$. (Здесь символ $\succ 0$ ($\succeq 0$) означает положительную (неотрицательную) определенность).

Если $\Omega \neq \emptyset$, то $\text{dom } \psi \supseteq \Omega$. Функция $\psi(u)$ непрерывно дифференцируема на Ω и ее градиент выражается формулой:

$$g_\psi(\bar{u}) = \{K_i(x(\bar{u}))\}_{i=1}^m,$$

где $x(\bar{u})$ – решение системы (13). Двойственной оценкой задачи (11), (12) называется

$$\psi^* = \sup_{u \in \text{dom } \psi} \psi(u).$$

Если $\max_{u \in \text{dom } \psi} \psi(u)$ достигается на $u^* \in \Omega$, то $g_\psi(u^*) = 0$, т.е. $K_i(x(u^*)) = 0$, $i = 1, \dots, m$. В этом случае $x(u^*)$ является допустимой и оптимальной точкой и $K_0(x(u^*)) = \psi^*$.

Для нахождения ψ^* нужно сконструировать алгоритмы, дающие возможность решать следующие подзадачи:

а) проверять, является ли область Ω не пустой, и если Ω не пуста, находить точку $\bar{u} \in \Omega$;

б) достаточно эффективно максимизировать $\psi(u)$ на Ω ;

с) после каждой итерации алгоритма отслеживать выход текущей точки за пределы Ω , и если это произошло, то алгоритм должен снова найти допустимую точку в Ω .

Первая подзадача сводится к нахождению $\lambda^* = \max_u \lambda_n[Q(u)]$, где $\lambda_n(A)$ означает минимальное собственное число симметричной матрицы A размерности $n \times n$. Если $\lambda^* \leq 0$, то Ω пусто. Если в процессе максимизации $\lambda_n[Q(u)]$ получено \bar{u} такое, что $\lambda_n[Q(\bar{u})] > 0$, то точка $\bar{u} \in \Omega$. Поскольку $\lambda_n[Q(u)]$ является вогнутой негладкой функцией от u , то для ее максимизации можно применять метод субградиентного типа с растяжением пространства в направлении разности двух последовательных градиентов (так называемый r -алгоритм [11, 12]), который практически является одним из наиболее эффективных методов негладкой оптимизации.

Этот метод можно применять для решения подзадач б и с. При их решении проверка выхода текущей точки \bar{u} за пределы Ω осуществляется путем разложения матрицы $Q(\bar{u})$ по Холецкому. Если $\bar{u} \notin \Omega$, то включается ветвь алгоритма, решающая первую подзадачу. В случае, если $\psi(u)$ имеет единственную точку максимума $u^* \in \Omega$, то после определенного числа итераций r -алгоритма, когда длины шагов становятся достаточно малыми, точки максимизирующей последовательности не выходят области Ω .

Осложнения возникают тогда, когда точка максимума функции $\psi(u)$ либо определяется неоднозначно, либо находится на границе области Ω . Первый случай типичен для квадратичных задач, когда вводятся избыточные ограничения, с целью улучшения двойственных оценок. Второй случай встречается довольно часто, например, в задачах, где двойственная оценка не совпадает с оптимальным значением задачи (11), (12), а также, когда оптимальная точка задачи (11), (12) не единственная.

Пусть в квадратичной задаче необходимо найти

$$\inf K_0(x), \quad x \in L^n \quad (14)$$

при ограничениях

$$K_i(x) = 0, \quad i = 1, \dots, m, \quad (15)$$

где часть ограничений избыточна. Без ограничения общности будем считать, что первые k ограничений, $k < m$, являются "основными", а ограничения с индексами $i = k+1, \dots, m$ — избыточными, т.е. $K_i(x) = 0$,

$i \in \{k+1, \dots, m\}$ во всех точках x , удовлетворяющих условию

$$K_s(x) = 0, \quad s = 1, \dots, k.$$

Рассмотрим "укороченную" функцию Лагранжа $L^{(k)}(x, u^{(k)}) = K_0(x) + \sum_{i=1}^k u_i^{(k)} K_i(x)$. Предположим, что задача (11), (12) имеет оптимальную точку x^* , которая является стационарной по x точкой функции $L^{(k)}$ при $u^{(k)} = (u^{(k)})^*$. Тогда при произвольных множителях $\bar{u}_{k+1}, \dots, \bar{u}_m$ существуют такие $\bar{u}_1, \bar{u}_2, \dots, \bar{u}_k$, что точка x^* будет стационарной точкой функции Лагранжа "полной" задачи (14), (15) при значениях множителей $\bar{u} = \{\bar{u}_1, \bar{u}_2, \dots, \bar{u}_k, \bar{u}_{k+1}, \dots, \bar{u}_m\}$, т.е. совокупность оптимальных множителей Лагранжа задачи (14), (15) образует $(m-k)$ - мерное линейное многообразие, проходящее через точку

$$\bar{u} = \{(u_i^{(k)})^*_{i=1, \dots, k}; \overbrace{0, \dots, 0}^{(m-k) \text{ раз}}\}. \text{ Обозначим его символом } M^*[m-k](x^*).$$

Рассмотрим $M_\Omega^*[x^*] = M^*[m-k](x^*) \cap \Omega$. Если $M_\Omega^*[x^*] \neq \emptyset$, то при любом $\bar{u} \in M_\Omega^*[x^*]$ решение $x(\bar{u})$ задачи $\min_x L(x, \bar{u})$ даст единственную оптимальную точку задачи (11), (12). Если решение задачи (14), (15) не единственно, то $M_\Omega^*[x^*] = \emptyset$, и $\psi^* = \sup_u \psi(u)$ может достигаться лишь на границе множества Ω , при этом не обязательно, чтобы оценка ψ^* совпадала с оптимальным значением квадратичной задачи (14), (15).

Пусть оптимальное значение вектора двойственных переменных u^* находится на границе области Ω , т.е. $u^* \in \text{dom } \psi \setminus \Omega$. Тогда при приближении \bar{u} к u^* , $\bar{u} \in \Omega$, решение системы (13) становится все менее устойчивым. Для регуляризации решения можно использовать ϵ -возмущения диагональных элементов матриц $Q(u)$, рассматривая $\tilde{Q}_\epsilon(u) = Q(u) + \epsilon I$, где $\epsilon > 0$, I - единичная матрица.

Другой способ преодоления этой неустойчивости состоит в следующем. Рассмотрим функцию:

$$\tilde{\psi}(u) = \begin{cases} \psi(u), & u \in \Omega, \\ \lambda_n[Q(u)], & \text{при } \lambda_n[Q(u)] \leq 0. \end{cases} \quad (16)$$

и будем применять вариант r -алгоритма с адаптивной регулировкой шагового множителя, используя в качестве субградиента на каждом шаге субградиент от функции $\tilde{\psi}(u)$ в соответствующей точке. Такой способ можно рассматривать как модификацию $r(\alpha)$ -алгоритма. Он дает возможность параллельного решения рассмотренных выше подзадач а, б и с и, кроме того, позволяет избежать осложнений, связанных как с неоднозначностью u^* , так и с принадлежностью u^* границе Ω . Подробное описание исследовательской программы и результатов вычис-

лительных экспериментов при нахождении (с помощью модификации $r(\alpha)$ -алгоритма) u^* для квадратичных задач дано ниже.

3. Описание вычислительных экспериментов нахождения двойственных оценок для задач квадратичного типа

Описание $r(\alpha)$ -алгоритма Поскольку на основе метода субградиентного типа с растяжением пространства в направлении разности двух последовательных субградиентов можно построить семейство алгоритмов, используя различные способы регулировки шагового множителя и коэффициента растяжения пространства, нужно уточнить, какой именно модификацией алгоритма мы пользовались при выполнении численных экспериментов. Это так называемый $r(\alpha)$ -алгоритм с постоянным коэффициентом растяжения пространства α и адаптивным способом регулировки шагового множителя [2, 12], разработанный в Институте кибернетики НАН Украины и хорошо зарекомендовавший себя в практических расчетах.

При описании вычислительной схемы $r(\alpha)$ -алгоритма применительно к решению задачи минимизации выпуклой функции $f(x)$, определенной на n -мерном пространстве E^n и обладающей свойством

$$\lim_{|x| \rightarrow +\infty} f(x) \rightarrow +\infty,$$

будем использовать оператор $R_\alpha(\xi)$ растяжения (сжатия) пространства с коэффициентом $\alpha \geq 1$ ($0 < \alpha < 1$) в направлении $\xi \in E^n$, $\|\xi\| = 1$. В матричной форме он представим формулой:

$$R_\alpha(\xi) = I_n + (\alpha - 1)\xi\xi^T,$$

где I_n – единичная матрица размерности $n \times n$. Субградиент функции $f(x)$ в точке x_k будем обозначать $g_f(x_k)$, а полученную с помощью $r(\alpha)$ -алгоритма приближенную точку минимума $f(x)$ при выполнении условия останова обозначим x_r^* .

При комментировании основных операций в вычислительной схеме $r(\alpha)$ -алгоритма (комментарии выделены скобками) будем использовать следующие обозначения: X – исходное пространство переменных (аргументов) функции $f(x)$; A – неособенная матрица размерности

$n \times n$, задающая невырожденный линейный оператор A , который действует из E^n в E^n ; B – матрица обратного преобразования пространства ($B=A^{-1}$); Y – преобразованное линейным оператором A пространство переменных (аргументов) функции $\varphi(y) = f(By)$; $g_\varphi(y_k)$ – субградиент функции $\varphi(y)$ в точке y_k , который вычисляется по формуле $g_\varphi(y) = B^T g_f(x_k)$.

$r(\alpha)$ -алгоритм использует: коэффициент растяжения пространства α ($\alpha > 1$); параметры адаптивной регулировки шагового множителя (в данном случае в направлении нормированного субградиента в преобразованном пространстве переменных) h_0, q_1, n_h, q_2 (h_0 – величина начального шага, q_1 – коэффициент уменьшения шага ($q_1 < 1$), если условие спуска по направлению выполняется за один шаг, q_2 – коэффициент увеличения шага ($q_2 > 1$), число n_h задает число шагов одномерного спуска, после которых шаговый множитель будет увеличиваться в q_2 раз ($n_h > 1$)); параметры ε_x и ε_g , задающие условия останова алгоритма (ε_x задает останов по отклонению аргумента; останов алгоритма в точке x_{k+1} происходит, если выполнено условие $\|x_{k+1} - x_k\| \leq \varepsilon_x$; ε_g задает останов по норме субградиента; останов алгоритма в точке x_{k+1} происходит, если выполнено условие $\|g_f(x_{k+1})\| \leq \varepsilon_g$).

Вычислительная схема $r(\alpha)$ -алгоритма принимает следующий вид.

В начале процесса имеем начальную точку $x_0 \in E^n$, параметры $\alpha, h_0, q_1, n_h, q_2, \varepsilon_x, \varepsilon_g$. Вычислим $f(x_0)$ и $g_f(x_0)$. Если $\|g_f(x_0)\| \leq \varepsilon_g$, то $x_r^* = x_0$ и останов. В противном случае полагаем $B_0 = I_n$ (матрица размерности $n \times n$), $\tilde{g}_0 = g_f(x_0)$, $g_0 = g_f(x_0)$.

Пусть в результате вычислений на k -й итерации процесса получены величины $x_k, h_k, \tilde{g}_k, g_k = g_f(x_k)$ и матрица B_k размерности $n \times n$. (Здесь h_k – текущее значение шагового множителя в направлении нормированного субградиента в преобразованном пространстве переменных $Y_k = A_k X = B_k^{-1} X$. Вектор \tilde{g}_k используется для обозначения субградиента на предыдущей итерации метода, т.е. субградиента функции $\varphi_k(y) = f(B_k y)$ в точке $y_{k-1} = B_k^{-1} x_{k-1}$. Исключением является итерация при $k = 0$, где \tilde{g}_0 совпадает с субградиентом $f(x)$ в точке x_0 . Это сделано для удобства при программировании метода (см. п. 1.3.)). Тогда переход к $(k+1)$ -й итерации процесса состоит из следующих операций.

1. Подготовка к операции растяжения пространства

Вычисляем следующие величины.

1.1. $g_k^* = B_k^T g_k$ (соответствует $g_{\varphi_k}(y_k)$, где $y_k = A_k x_k$, т.е. субградиенту функции $\varphi_k(y) = f(B_k y)$ в точке $y_k = A_k x_k = B_k^{-1} x_k$ преобразованного пространства аргументов $Y_k = A_k X = B_k^{-1} X$).

1.2. $r_k = g_k^* - \tilde{g}_k$ (соответствует вектору разности двух последо-

вательных субградиентов в преобразованном пространстве аргументов $Y_k = A_k X$, т.е. $r_k = g_{\varphi_k}(y_k) - g_{\varphi_k}(y_{k-1}) = B_k^T(g_f(x_k) - g_f(x_{k-1}))$.

Шаг 1.3. $\xi_k = \begin{cases} 0, & \text{если } \|r_k\| \leq \varepsilon_0, \\ \frac{r_k}{\|r_k\|}, & \text{если } \|r_k\| > \varepsilon_0. \end{cases}$ (нормировка вектора r_k для

выполнения очередной операции растяжения пространства в направлении разности двух последовательных субградиентов).

В п. 1.3 ε_0 – точность представления нуля для ЭВМ. При нахождении x_r^* с точностью $\varepsilon_x = 10^{-6}$ либо $\varepsilon_g = 10^{-10}$ достаточно положить $\varepsilon_0 = 10^{-20}$. Отметим, что $R_\alpha(\xi_k) = I_n$ при $\xi_k = 0$, т.е. растяжение пространства не будет выполняться. Это всегда имеет место на итерации метода при $k = 0$, поскольку $\tilde{g}_0 = g_f(x_0)$. Кроме того, $\xi_k = 0$ может иметь место на "предельных" шагах метода при выборе либо большого значения α , либо слишком малых значений ε_x ($\varepsilon_x \sim 10^{-10}$) и ε_g ($\varepsilon_g \sim 10^{-10}$). Однако такая ситуация имеет "вырожденный" характер, так как получение x_r^* с достаточно высокой точностью ($\sim 10^{-6}$) обеспечивает выбор параметров $\varepsilon_x = 10^{-6}$, $\varepsilon_g = 10^{-6}$ и $\alpha \in [2, 3]$. Тем не менее, если значение x_r^* должно быть более точным, то в этом случае можно проводить процедуру "восстановления" процесса, положив $B_{k+1} = I_n$ и установив величину шага h_{k+1} порядка $\|x_k - x_{k-1}\|$.

2. Операция растяжения пространства

Вычисляем следующие величины

2.1. $B_{k+1} = B_k + (\beta - 1)(B_k \xi_k) \xi_k^T$, где $\beta = \frac{1}{\alpha}$ (пересчет матрицы обратного преобразования пространства при переходе в очередное преобразованное пространство аргументов, который следует из соотношения $B_{k+1} = (R_\alpha(\xi_k) A_k)^{-1} = B_k R_\beta(\xi_k) = B_k (I_n + (\beta - 1) \xi_k \xi_k^T)$.

2.2. $\tilde{g}_{k+1} = g_k^* + (\beta - 1)(g_k^*, \xi_k) \xi_k^T$ (пересчет текущего субградиента при переходе в очередное преобразованное пространство аргументов, который следует из соотношения $\tilde{g}_{k+1} = g_{\varphi_{k+1}}(y_k) = B_{k+1}^T g_f(x_k) = R_\beta(\xi_k) B_k^T g_f(x_k) = R_\beta(\xi_k) g_k^* = (I_n + (\beta - 1) \xi_k \xi_k^T) g_k^*$.

3. Процедура одномерного спуска по направлению

3.1. Вычисляем $p_{k+1} = B_{k+1} \frac{\tilde{g}_{k+1}}{\|\tilde{g}_{k+1}\|}$ (направление спуска, соответствующее направлению нормированного субградиента в преобразованном пространстве аргументов $Y_{k+1} = A_{k+1} X = B_{k+1}^{-1} X$).

3.2. Осуществляем спуск по направлению $-p_{k+1}$ с адаптивной регуляровкой шага.

Положим $z_0 = x_k$, $l = 0$, $\tilde{h}_0 = h_k$. Вычисляем

$$z_{l+1} = z_l - \tilde{h}_l p_{k+1},$$

(здесь $\tilde{h}_l = h_k q_2^{\lfloor \frac{l}{n_h} \rfloor}$ и $\lceil \cdot \rceil$ — целая часть рационального числа) до тех пор, пока не выполнится одно из следующих условий:

3.2.1. $(p_{k+1}, g_f(z_{l+1})) \leq 0$ (задает условие спуска по направлению, его выполнение обеспечивается в силу предположения $\lim_{|x| \rightarrow +\infty} f(x) \rightarrow +\infty$).

3.2.2. $\|g_f(z_{l+1})\| \leq \varepsilon_g$, тогда $x_r^* = z_{l+1}$ и останов (задает останов по норме субградиента и в основном имеет место для гладких функций).

Если условие спуска (п.3.2.1) выполнилось на первом шаге, т.е. при $l = 0$, то полагаем $\tilde{h}_0 = q_1 h_k$ (уменьшение шагового множителя, которое используется, как правило, при минимизации гладких функций).

4. Переход к очередному шагу

4.1. Полагаем $x_{k+1} = z_{l+1}$, $g_{k+1} = g_f(z_{l+1})$ и $h_{k+1} = \tilde{h}_l$.

4.2. Если $\|x_{k+1} - x_k\| \leq \varepsilon_x$, то $x_r^* = x_{k+1}$ и останов (по отклонению аргумента). В противном случае переходим $(k + 1)$ -й итерации процесса с новыми значениями величин x_{k+1} , h_{k+1} , \tilde{g}_{k+1} , $g_{k+1} = g_f(x_{k+1})$ и матрицы B_{k+1} .

Для B -формы $r(\alpha)$ -алгоритма (т.е. использующего коррекцию матрицы обратного преобразования B) приведенная вычислительная схема является наиболее экономной по числу арифметических операций. Именно она лежит в основе практически используемой программной реализации $r(\alpha)$ -алгоритма [13] и требует на одной итерации всего три операции умножения матрицы на вектор ($3n^2$ арифметических операций умножения и столько же операций сложения) и одну операцию одноранговой коррекции матрицы B_{k+1} (n^2 арифметических операций умножения и столько же операций сложения). Реализация B -формы $r(\alpha)$ -алгоритма в рамках общей схемы r -алгоритмов, описанной в работе [14], потребовала бы одну дополнительную операцию умножения матрицы на вектор, а трудоемкость одной итерации составила бы $5n^2$ арифметических операций умножения вместо $4n^2$ для приведенной вычислительной схемы $r(\alpha)$ -алгоритма.

Вычислительная эффективность $r(\alpha)$ -алгоритма зависит от коэффициента растяжения пространства и параметров адаптивной регуляции шагового множителя (подробные рекомендации даны в [2], с. 45–47). Суть выбора параметров состоит в том, чтобы адаптивный способ регуляции шагового множителя позволял увеличивать точность поиска минимума функции по направлению в процессе счета и при этом число шагов по направлению не должно быть большим. При минимизации негладких функций это обеспечивает следующий выбор параметров: $\alpha = 2 \div 3$, $h_0 = 1.0$, $q_1 = 1.0$, $q_2 = 1.1 \div 1.2$, $n_h = 2 \div 3$. Если известна априорная оценка расстояния от начальной точки x_0 до точ-

ки минимума x^* , то начальный шаг h_0 целесообразно выбирать порядка $\|x_0 - x^*\|$. При минимизации гладких функций рекомендуемые параметры такие же, за исключением q_1 ($q_1 = 0.8 \div 0.95$). Это обусловлено тем, что дополнительное измельчение шага способствует увеличению точности поиска минимума функции по направлению, что при минимизации гладких функций обеспечивает более быструю скорость сходимости.

При таком выборе параметров $r(\alpha)$ -алгоритм, как правило, дает следующие результаты: число спусков по направлению редко превосходит два, за n итераций точность по функционалу улучшается в три-пять раз. Выбирая в качестве критериев останова $\varepsilon_x, \varepsilon_g \sim 10^{-6} \div 10^{-5}$ при минимизации выпуклой функции, даже существенно овражной структуры, можно обеспечить нахождение x_r^* со значением целевой функции, достаточно близким к оптимальному ($[f(x_r^*) - f(x^*)]/|f(x^*)| + 1] \sim 10^{-6} \div 10^{-5}$ – для негладких и $[f(x_r^*) - f(x^*)]/|f(x^*)| + 1] \sim 10^{-12} \div 10^{-10}$ – для гладких функций). Это подтверждается результатами многочисленных тестовых и реальных расчетов.

Замечание. Следует отметить, что в ряде работ использовались другие программные реализации r -алгоритмов, существенно отличные от описанного выше $r(\alpha)$ -алгоритма. Например, в работе [15] применялась регулировка шагового множителя по формуле геометрической прогрессии ($t_k = t_0 q^k, q < 1$), что в общем случае может привести к предельной точке, которая не обязательно окажется точкой минимума при выбранных значениях t_0 и q . Тестируемый в [16] "вариант r -алгоритма" отличается от $r(\alpha)$ -алгоритма как регулировкой шагового множителя, так и выбором направления растяжения пространства. Используемый в [16] способ регулировки сильно измельчает шаговый множитель, что способствует увеличению числа шагов одномерного спуска в направлении убывания функции и отрицательно сказывается на общем количестве вычислений $f(x)$ и $g_f(x)$. Как правило, для $r(\alpha)$ -алгоритма характерна более устойчивая и быстрая сходимость, чем для вариантов r -алгоритмов из [15, 16].

Тестовые эксперименты. Использование модификации $r(\alpha)$ -алгоритма применительно к нахождению двойственных оценок в задачах квадратичного типа (14), (15) будет подразумевать подстановку вместо выпуклой функции $f(x)$ функции $-\tilde{\psi}(u)$, определенной согласно (16). Суперградиент вогнутой функции $\tilde{\psi}(u)$ вычисляется

$$\text{supergrad}_u \tilde{\psi}(u) = \begin{cases} \text{grad}_u \psi(u), & u \in \Omega, \\ \text{supergrad}_u \lambda_n [Q(u)], & \lambda_n [Q(u)] \leq 0. \end{cases} \quad (17)$$

Здесь

$$\text{grad}_u \psi(u) = K_i(x(\bar{u})), \quad (18)$$

где $x(\bar{u})$ – решение системы уравнений (13), и

$$\text{grad}_u \lambda_n[Q(u)] = \sum_{i,j} \frac{\partial^2 K_l(x)}{\partial x_i \partial x_j} \xi_i \xi_j, \quad l = 1, \dots, m, \quad (19)$$

где ξ – произвольный нормированный собственный вектор, относящийся к минимальному собственному числу матрицы $Q(u)$ ($\lambda_n[Q(u)]$).

Цель тестовых экспериментов – изучение эффективности этой модификации $r(\alpha)$ -алгоритма при нахождении двойственных оценок в задачах квадратичного типа (14), (15). Основные исследуемые характеристики: iter – число итераций модификации $r(\alpha)$ -алгоритма; $n_{\tilde{\psi}}$ – общее число вычислений суперградиента $\tilde{\psi}(u)$ по формуле (17); n_λ – число вычислений суперградиента $\tilde{\psi}(u)$, связанных с его вычислением по формуле (19), характеризующие затраты модификации $r(\alpha)$ -алгоритма на возврат в область положительной определенности матрицы $Q(u)$.

При вычислении суперградиента $\tilde{\psi}(u)$ использованы следующие фортрановские подпрограммы:

для решения системы уравнений (13) с симметричной положительно определенной матрицей $Q(u)$ – подпрограммы *DPPFA* (реализует факторизацию матрицы $Q(u)$, если она положительно определена, и в случае, если $Q(u)$ не является положительно определенной, сигнализирует об этом) и *DPPSL* (определяет решение системы, используя факторизованную программой *DPPFA* матрицу $Q(u)$) из [17];

для нахождения минимального собственного числа $\lambda_n[Q(u)]$ матрицы $Q(u)$ и соответствующего этому числу собственного вектора ξ применена подпрограмма *EIGEN* из [18].

Первая серия экспериментов связана с тестовыми задачами квадратичного типа, которые возникают при попытке представления неотрицательной полиномиальной функции $P(x) - P^*$ в виде суммы квадратов, а также при нахождении ее глобального минимума. Тестовые задачи формировались следующим образом.

Как отмечалось в разд. 1, любой неотрицательный однородный полином $\bar{P}(x_1, x_2, x_3)$ от трех переменных четвертой степени представим в виде суммы квадратов. Пусть $\bar{P}(x_1, x_2, x_3) \in H(3, 4)$. Рассмотрим функцию $f\left(\frac{x_1}{x_3}, \frac{x_2}{x_3}\right) = \frac{\bar{P}(x_1, x_2, x_3)}{x_3^4}$. Сделав подстановку: $\frac{x_1}{x_3} = y_1$; $\frac{x_2}{x_3} = y_2$, получим неоднородный полином четвертой степени от двух переменных $P(y_1, y_2)$. Если полином $\bar{P}(x_1, x_2, x_3)$ неотрицателен, то в силу теоремы

Гильберта он представим в виде суммы квадратов, а значит, и производный от него полином $P(y_1, y_2)$ также представим в виде суммы квадратов полиномов. Отсюда следует, что любой ограниченный снизу полином четвертой степени от двух переменных $P(y_1, y_2)$ может быть представлен в виде:

$$P(y_1, y_2) = \sum_{i \in I} Q_i^2(y_1, y_2) + P^*,$$

где $Q_i(y_1, y_2)$ – полиномы, степень которых не превышает двух, $P^* = \inf_{y \in L^2} P(y_1, y_2)$.

Для тестовых экспериментов рассматривались два типа полиномов:

$$P_\gamma(y_1, y_2) = (y_1 + 1)^4 - \gamma y_1^2 y_2^2 + (y_1 + 2)^4,$$

$$P_{a_1, a_2}(y_1, y_2) = a_1 y_1^4 - 2y_1^3 + y_1^2 + a_2 y_2^4 - 2y_2^3 + y_2^2$$

с параметрами γ ($\gamma \in [0, 2]$) и a_1, a_2 ($a_1, a_2 \in [1 - \varepsilon, 1 + \varepsilon]$), где $\varepsilon > 0$ – возмущающий параметр.

Для указанных полиномов задача нахождения глобального минимума сводилась к задаче минимизации квадратичной функции от пяти переменных: $x_1 = y_1$, $x_2 = y_2$, $x_3 = y_1^2$, $x_4 = y_2^2$, $x_5 = y_1 y_2$ при следующих ограничениях:

$$x_1^2 - x_3 = 0; \quad x_2^2 - x_4 = 0; \quad x_1 x_2 - x_5 = 0;$$

$$x_5^2 - x_3 x_4 = 0; \quad x_3 x_2 - x_1 x_5 = 0; \quad x_4 x_1 - x_2 x_5 = 0.$$

Функция Лагранжа этих квадратичных задач имеет вид

$$L(x, u) = K_0(x) + u_1(x_1^2 - x_3) + u_2(x_2^2 - x_4) + u_3(x_1 x_2 - x_5) + \\ + u_4(x_5^2 - x_3 x_4) + u_5(x_3 x_2 - x_1 x_5) + u_6(x_4 x_1 - x_2 x_5),$$

где $K_0(x) = 6x_1^2 + 4x_1 x_3 + 24x_2^2 + 8x_2 x_4 + x_3^2 - \gamma x_3 x_4 + x_4^2 + 4x_1 + 32x_2 + 17$ для полинома $P_\gamma(y_1, y_2)$ и $K_0(x) = x_1^2 - 2x_1 x_3 + a_1 x_3^2 + x_2^2 - 2x_2 x_4 + a_2 x_4^2$ для полинома $P_{a_1, a_2}(y_1, y_2)$.

При нахождении двойственных оценок квадратичных задач, соответствующих полиномам $P_\gamma(y_1, y_2)$ и $P_{a_1, a_2}(y_1, y_2)$, применялась модификация $r(\alpha)$ -алгоритма со следующими параметрами: $\alpha = 2$, $h_0 = 1.0$, $q_1 = 1.0$, $n_h = 3$, $q_2 = 1.1$. В качестве начальной стартовой точки выбиралась $u_0 = (1, 1, 1, 1, 1, 1)$. Работа алгоритма прекращалась при выполнении одного из условий останова: $\varepsilon_x = 10^{-5}$ либо $\varepsilon_g = 10^{-5}$.

Результаты вычислений для $P_\gamma(y_1, y_2)$ при значениях параметра $\gamma \in [0, 1.9]$ приведены в табл. 1. Здесь в дополнение к вышепринятым обозначениям iter , $n_{\tilde{\psi}}$ и n_λ использованы следующие: ψ_r^* – экстремальное значение $\psi(u)$ в точке u_r^* , которая получена модификацией $r(\alpha)$ -алгоритма (с точностью до критерия останова); y_1^* , y_2^* – соответствующие u_r^* оптимальные значения компонент $x_1(u_r^*)$ и $x_2(u_r^*)$ вектора решений системы уравнений (13) при $\bar{u} = u_r^*$; $\lambda_n[Q(u_r^*)]$ – значение минимального собственного числа матрицы $Q(u)$ в точке u_r^* ; ε_1^* – показатель точности, с которой выполнены необходимые условия экстремума для полинома $P(y_1, y_2)$ (в данном случае для $P_\gamma(y_1, y_2)$) в точке $y^* = (y_1^*, y_2^*)$. Точность ε_1^* вычислялась по формуле:

$$\varepsilon_1^* = \max \left\{ \left| \frac{\partial P(y_1, y_2)}{\partial y_1} (y_1^*, y_2^*) \right|, \left| \frac{\partial P(y_1, y_2)}{\partial y_2} (y_1^*, y_2^*) \right| \right\}. \quad (20)$$

Как видно из табл. 1., модификация $r(\alpha)$ -алгоритма работает устойчиво и находит глобальный минимум с высокой точностью. При этом для всех тестовых задач останов алгоритма происходил по критерию нормы субградиента (суперградиента), поскольку u_γ^* при $\gamma \in [0, 1.9]$ находится внутри области Ω , о чем свидетельствует столбец $\lambda_n[Q(u_r^*)]$. В силу этого обеспечивается достаточно равномерная точность ($\sim 10^{-6}$) выполнения необходимых условий экстремума для полинома $P_\gamma(y_1, y_2)$ в точке $y^* = (y_1^*, y_2^*)$ (см. столбец ε_1^*).

Следует отметить поведение алгоритма вблизи критического значения параметра $\gamma = 2$. При $\gamma \rightarrow 2$ ($\gamma < 2$) u_γ^* приближается к границе области Ω . Результаты вычислений показаны в табл. 2. Здесь по отношению к табл. 1. сделаны изменения, которые обусловлены большими значениями модулей y_1^* и y_2^* . Графа ε_1^* заменена на $\varepsilon_1^* = \max \left\{ \frac{\varepsilon_1^*}{|y_1^*|^3}, \frac{\varepsilon_1^*}{|y_2^*|^3} \right\}$ – относительную точность до $\max\{|y_1^*|^3, |y_2^*|^3\}$ выполнения необходимых условий экстремума полинома $P_\gamma(y_1, y_2)$ в точке $y^* = (y_1^*, y_2^*)$, где ε_1^* вычисляется по формуле (20). Аналогично колонка y_2^* заменена на колонку разности y_2^* и y_1^* . Отметим, что при приближении γ к двум количество итераций, в которых суперградиент $\tilde{\psi}(u)$ вычисляется по формуле (19), увеличивается. Это означает (см. столбец $\lambda_n[Q(u_r^*)]$), что u_γ^* при приближении к границе области Ω . При $\gamma = 2$ все итерации связаны с вычислением суперградиента $\tilde{\psi}(u)$ по формуле (19). Несмотря на "вырожденный" в вычислительном отношении характер тестовых задач ($\psi^* \rightarrow -\infty$ при $\gamma \rightarrow 2$) модификация $r(\alpha)$ -алгоритма является довольно устойчивой (для всех тестовых задач из табл. 2. срабатывал критерий останова по отклонению аргумента).

Табл. 1. Тесты для полинома $P_\gamma(y_1, y_2)$ при $\gamma \in [0, 1, 9]$.

γ	$iter(n_{\tilde{\psi}}, n_\lambda)$	ψ_r^*	ε_1^*	y_1^*	y_2^*	$\lambda_n Q(u_r^*) $
.0	82(119,34)	-39888 * 10 ⁻¹¹	9.41 * 10 ⁻⁶	-1.0001	-1.9995	6.31 * 10 ⁻⁸
.2	69(93,5)	-5.9166	4.56 * 10 ⁻⁶	-2.3399	-3.2063	3.47 * 10 ⁻²
.3	67(96,4)	-13.471	4.32 * 10 ⁻⁶	-2.7468	-3.5967	4.13 * 10 ⁻²
.4	74(100,4)	-26.140	7.81 * 10 ⁻⁶	-3.1634	-4.0006	4.98 * 10 ⁻²
.5	68(94,5)	-46.605	9.12 * 10 ⁻⁶	-3.6080	-4.4347	5.44 * 10 ⁻²
.6	71(91,5)	-79.096	7.16 * 10 ⁻⁶	-4.0953	-4.9131	6.16 * 10 ⁻²
.7	70(94,5)	-130.42	3.93 * 10 ⁻⁶	-4.6406	-5.4505	6.43 * 10 ⁻²
.8	67(97,6)	-211.80	5.80 * 10 ⁻⁶	-5.2623	-6.0652	6.59 * 10 ⁻²
.9	80(120,6)	-342.36	5.52 * 10 ⁻⁶	-5.9842	-6.7808	6.65 * 10 ⁻²
1.0	82(118,12)	-555.83	3.70 * 10 ⁻⁶	-6.8386	-7.6294	6.46 * 10 ⁻²
1.1	82(123,14)	-914.53	8.93 * 10 ⁻⁶	-7.8716	-8.6572	6.34 * 10 ⁻²
1.2	90(130,16)	-1540.1	5.24 * 10 ⁻⁶	-9.1520	-9.9326	6.06 * 10 ⁻²
1.3	92(137,23)	-2686.8	4.54 * 10 ⁻⁶	-10.787	-11.563	5.51 * 10 ⁻²
1.4	91(148,26)	-4935.4	8.13 * 10 ⁻⁶	-12.956	-13.728	5.00 * 10 ⁻²
1.5	103(156,30)	-9777.9	3.59 * 10 ⁻⁶	-15.981	-16.748	4.43 * 10 ⁻²
1.6	103(166,38)	-21724.	1.71 * 10 ⁻⁶	-20.504	-21.267	3.66 * 10 ⁻²
1.7	108(176,44)	-58173.	1.08 * 10 ⁻⁶	-28.025	-28.785	2.93 * 10 ⁻²
1.8	118(198,57)	-22042 * 10 ⁶	4.28 * 10 ⁻⁶	-43.046	-43.802	1.96 * 10 ⁻²
1.9	176(308,74)	-19687 * 10 ⁷	8.23 * 10 ⁻⁶	-88.065	-88.818	9.56 * 10 ⁻³

Табл. 2. Тесты для полинома $P_\gamma(y_1, y_2)$ при $\gamma \rightarrow 2$.

γ	$iIter(n_\psi, n_\lambda)$	ψ_r^*	ε_2^*	y_1^*	$y_2^* - y_1^*$	$\lambda_n[Q(u_r^*)]$
1.95000000	189(345,89)	$-1.6610 * 10^7$	$4.2730 * 10^{-12}$	-178.07	-75158	$5.45 * 10^{-3}$
1.97500000	204(378,107)	$-1.3640 * 10^8$	$2.2521 * 10^{-12}$	-358.08	-75079	$2.65 * 10^{-3}$
1.98750000	205(384,132)	$-1.1054 * 10^9$	$3.9227 * 10^{-8}$	-718.08	-75040	$9.94 * 10^{-4}$
1.99375000	196(364,147)	$-8.9005 * 10^9$	$2.6580 * 10^{-9}$	-1438.1	-75019	$6.06 * 10^{-4}$
1.99687500	218(416,177)	$-7.1434 * 10^{10}$	$1.5834 * 10^{-9}$	-2878.1	-75010	$2.47 * 10^{-4}$
1.99843750	212(421,177)	$-5.7239 * 10^{11}$	$1.5242 * 10^{-11}$	-5758.1	-75005	$1.29 * 10^{-4}$
1.99921875	228(484,207)	$-4.5828 * 10^{12}$	$1.9915 * 10^{-8}$	-11518.	-74997	$6.07 * 10^{-5}$
2.00000000	153(523,523)	--	--	--	--	$-1.08 * 10^{-2}$

Эксперименты для первого семейства $P_\gamma(y_1, y_2)$ отражают поведение модификации $r(\alpha)$ -алгоритма в случае единственного глобального минимума. Второй же полином $P_{a_1, a_2}(y_1, y_2)$ при $a_1 = a_2 = 1$ имеет четыре глобальных минимума: $(0, 0)$, $(0, 1)$, $(1, 0)$, $(1, 1)$. Поэтому оптимальное значение множителей Лагранжа u^* квадратичной задачи находится на границе области Ω . Это вызывает неустойчивость алгоритма при подходе к u^* . Для выделения различных глобальных минимумов будем вводить малые ε -возмущения в параметры задачи a_1 и a_2 .

Результаты работы модификации $r(\alpha)$ -алгоритма по "выделению" различных глобальных минимумов с возмущенными параметрами a_1 и a_2 при разных $\varepsilon = \{0.1, 0.01, 0.001\}$ представлены в табл. 3. Здесь все обозначения такие же, как и в табл. 1. Как видим, "выделение" глобальных минимумов происходит устойчиво, необходимые условия экстремума для полинома $P_{a_1, a_2}(y_1, y_2)$ при $a_1, a_2 = 1 \pm \varepsilon$ в в точке $y^* = (y_1^*, y_2^*)$ выполняются с достаточной точностью (см. значения ε_1^*). В то же время имеется некоторый запас по "расстоянию" от u^* до границы области Ω (см. значения $\lambda_n[Q(u_r^*)]$), в результате количество вычислений суперградиента $\tilde{\psi}(u)$ по формуле (19) оказалось сравнительно небольшим по отношению к общему числу вычислений суперградиента $\tilde{\psi}(u)$.

Вторая серия экспериментов связана с задачей нахождения максимального независимого множества вершин графа $G = (V, E)$. Эта задача сводится к задаче квадратичного типа:

$$\nu(G) = \max \sum_{j=1}^n x_j \quad (21)$$

при ограничениях:

$$x_i x_j = 0; \quad (i, j) \in E, \quad (22)$$

$$x_k^2 - x_k = 0; \quad k = 1, \dots, n \quad (23)$$

(здесь $V = \{1, \dots, n\}$ – множество вершин графа, E – множество его ребер).

Рассмотрим два способа нахождения оптимальных оценок $\nu(G)$ в задаче (21)–(23). Первый способ изложен в разд. 2. Для него мы проверим работу модификации $r(\alpha)$ -алгоритма при нахождении двойственных оценок ψ^* в задачах (21)–(23) для двух типов графов:

– G_1^n – циклический граф с числом вершин n и весовыми коэффициентами ребер, равными единице. Оптимальное значение ψ^* для задачи (21)–(23), соответствующей графу G_1^n , равно $n/2$, если n – четное, и $\frac{ncos(\pi/n)}{1 + cos(\pi/n)}$, если n – нечетное. Учитывая, что для графа G_1^n

Табл. 3. Тесты для полинома $P_{a_1, a_2}(y_1, y_2)$.

a_1	a_2	$iter(n_{\bar{y}}, n_A)$	ψ_r^*	ε_1^*	y_1^*	y_2^*	$\lambda_n[Q(u_r^*)]$
1.000	1.000	152(244,58)	-4.11 * 10 ⁻⁸	9.20 * 10 ⁻²	.59548	.44913	1.12 * 10 ⁻⁸
1.100	1.100	63(84, 8)	-3.62 * 10 ⁻¹²	1.11 * 10 ⁻⁵	5.55 * 10 ⁻⁶	5.39 * 10 ⁻⁶	2.49 * 10 ⁻²
.9000	.9000	73(100,11)	-.29964	1.02 * 10 ⁻⁵	1.2060	1.2060	1.49 * 10 ⁻²
1.100	.9000	67(93, 7)	-.14982	6.85 * 10 ⁻⁶	1.54 * 10 ⁻⁶	1.2060	2.51 * 10 ⁻²
.9000	1.100	71(104,12)	-.14982	2.08 * 10 ⁻⁵	1.2060	-6.10 * 10 ⁻⁷	2.02 * 10 ⁻²
1.010	1.010	88(120,13)	-5.36 * 10 ⁻¹³	1.01 * 10 ⁻⁵	-5.04 * 10 ⁻⁶	5.06 * 10 ⁻⁶	2.69 * 10 ⁻³
.9900	.9900	89(127,18)	-.020816	4.95 * 10 ⁻⁶	1.0200	1.0200	1.36 * 10 ⁻³
1.010	.9900	89(139,13)	-.010408	2.00 * 10 ⁻⁵	4.95 * 10 ⁻⁶	1.0200	2.03 * 10 ⁻³
.9900	1.010	83(110,15)	-.010408	1.18 * 10 ⁻⁵	1.0200	1.85 * 10 ⁻⁶	1.83 * 10 ⁻³
1.001	1.001	100(141,26)	-2.70 * 10 ⁻¹⁴	1.20 * 10 ⁻⁵	-6.01 * 10 ⁻⁶	-1.48 * 10 ⁻⁶	2.33 * 10 ⁻⁴
.9990	.9990	110(162,21)	-.00201	4.06 * 10 ⁻⁶	1.0020	1.0020	1.93 * 10 ⁻⁴
1.001	.9990	109(159,22)	-.00100	1.02 * 10 ⁻⁵	-2.72 * 10 ⁻⁶	1.0020	2.62 * 10 ⁻⁴
.9990	1.001	107(160,19)	-.00100	6.71 * 10 ⁻⁶	1.0020	3.21 * 10 ⁻⁶	1.70 * 10 ⁻⁴

существует группа автоморфизмов, обеспечивающая отображение множества вершин на себя, при котором сохраняются ребра графа, нахождение оценки ψ^* для задачи (21)–(23), соответствующей графу G_1^n , сводится к максимизации $\psi(u)$ от двух переменных, соответствующих группе ограничений (23) и группе ограничений (22).

– G_{1+}^n – дополнительный к G_1^n граф с единичными весовыми коэффициентами ребер. Для задачи (21)–(23), соответствующей графу G_{1+}^n , $\psi^* = 2$, если n – четное, и $\psi^* = \frac{1 + \cos(\pi/n)}{\cos(\pi/n)}$, если n – нечетное. Аналогично, как и для G_1^n , нахождение оценки ψ^* для задачи (21)–(23), соответствующей графу G_{1+}^n , сводится к максимизации $\psi(u)$ от $\frac{n}{2}$ [переменных (здесь)] [– целая часть рационального числа]. Первая переменная соответствует группе ограничений (23), вторая – группе ограничений (22), у которых расстояние между вершинами вдоль цикла равно двум, для третьей расстояние равно трем и так далее, для $\frac{n}{2}$ [–ой переменной] расстояние равно $\frac{n}{2}$ [

Вычислительные затраты модификации $r(\alpha)$ -алгоритма (в терминах числа итераций, общего количества вычислений суперградиента $\check{\psi}(u)$ и количества вычислений суперградиента $\dot{\psi}(u)$ по формуле (19)) при решении задач (21)–(23), соответствующих графам G_1^n и G_{1+}^n , при различных значениях n приведены в табл. 4. Здесь $\psi^* - \psi_r^*$ – разность между известным оптимальным значением $\psi(u)$ и значением $\psi(u)$, полученным модификацией $r(\alpha)$ -алгоритма с точностью до условия останова. Параметры для модификации $r(\alpha)$ -алгоритма такие же, как и в первой серии экспериментов, т.е. $\alpha = 2$, $h_0 = 1.0$, $q_1 = 1.0$, $n_h = 3$, $q_2 = 1.1$, $\varepsilon_x = 10^{-5}$, $\varepsilon_g = 10^{-5}$. В качестве начальной стартовой точки использовались: $u_0 = (1, 1)$ – для задач, соответствующих G_1^n , и $\frac{n-2}{2}$ [раз

$u_0 = (1, \overbrace{1, \dots, 1}^{\frac{n-2}{2}})$ – для задач, соответствующих G_{1+}^n . Из табл. 4. видно, что для всех тестовых задач модификация $r(\alpha)$ -алгоритма гарантирует нахождение ψ_r^* с достаточно хорошей точностью.

Второй способ нахождения оптимальных оценок $\nu(G)$ в задачах (21)–(23) состоит в сведении их к задачам минимизации выпуклой функции на классе симметричных матриц. В качестве верхней оценки для $\nu(G)$ можно взять [19]

$$\nu_1^*(G) = \min_y \lambda_1(A(y)), \quad (24)$$

где $\lambda_1(A)$ – максимальное собственное число матрицы A , $A(y)$ – класс симметричных матриц, у которых (i, j) -й элемент принимает произволь-

Табл. 4. Тесты для графов G_1^n и G_{1+}^n .

n	циклический граф G_1^n		дополнительный граф G_{1+}^n	
	$\psi^* - \psi_r^*$	$iter(n_{\bar{\psi}}, n_{\lambda})$	$\psi^* - \psi_r^*$	$iter(n_{\bar{\psi}}, n_{\lambda})$
5	$1.51 * 10^{-10}$	51(83,31)	$1.51 * 10^{-10}$	51(83,31)
6	$2.13 * 10^{-10}$	52(88,33)	$8.99 * 10^{-8}$	64(93,40)
7	$3.37 * 10^{-10}$	50(75,30)	$3.87 * 10^{-9}$	79(119,53)
8	$3.74 * 10^{-10}$	50(89,33)	$9.49 * 10^{-8}$	90(122,56)
9	$9.75 * 10^{-11}$	51(79,30)	$2.46 * 10^{-9}$	116(177,77)
10	$3.41 * 10^{-9}$	46(76,29)	$2.73 * 10^{-7}$	111(149,71)
15	$6.42 * 10^{-6}$	30(54,20)	$4.03 * 10^{-7}$	177(255,116)
16	$1.88 * 10^{-9}$	50(86,31)	$1.95 * 10^{-7}$	202(261,133)
19	$8.04 * 10^{-7}$	38(63,24)	$3.42 * 10^{-6}$	216(311,151)
20	$1.04 * 10^{-9}$	51(82,30)	$1.31 * 10^{-6}$	257(347,173)

ное значение y_{ij} при $(i, j) \in E$ и равен единице при $(i, j) \notin E$.

Оценку (24) можно улучшить, если использовать избыточные ограничения. Пусть к ограничениям (22), (23) добавлены избыточные ограничения

$$x_i x_j \geq 0, \quad (i, j) \notin E, \quad i \neq j.$$

Новой постановке задачи соответствует более точная, чем (24), оценка:

$$\nu_2^*(G) = \min_z \lambda_1(A(z)), \quad (25)$$

где $A(z)$ - класс симметричных матриц, у которых (i, j) -й элемент принимает произвольное значение z_{ij} при $(i, j) \in E$, $z_{ij} \geq 1$ при $(i, j) \notin E$ и равен единице при $i = j$.

При решении задач (24) и (25) можно применять $r(\alpha)$ -алгоритм. Для тестов рассмотрим задачу, связанную с графом G_2^{64} , в котором вершины соответствуют бинарным кодовым словам длины шесть и ребрами соединены кодовые слова, между которыми расстояние по Хеммингу не превышает трех. Учитывая, что для графа G_2^{64} существует группа автоморфизмов, обеспечивающая отображение множества вершин на себя, при котором сохраняются ребра графа, нахождение первой оценки сводится к минимизации выпуклой функции (24) от трех переменных, на классе симметричных матриц размера 64×64 . Первая переменная соответствует ребрам, связывающим кодовые слова, расстояние по Хемингу между которыми равно единице, вторая - с расстоянием два, третья - с расстоянием три. Оптимальное значение оценки $\nu_1^*(G_2^{64}) = 16/3$. Нахо-

Табл. 5. Нахождение оценок для графа G_2^{64} .

Оценка $\nu_2(G_2^{64})$ ($\nu_2^*(G_2^{64})=4$)			Оценка $\nu_1(G_2^{64})$ ($\nu_1^*(G_2^{64}) = 16/3$)		
$iter$	$f(x_{iter})$	N_λ	$iter$	$f(x_{iter})$	N_λ
0	64.00000	0(0)	0	64.00000	0(0)
5	7.609600	14(14)	5	6.914880	14(14)
10	8.147414	0(0)	10	5.937764	0(0)
15	4.777998	0(0)	15	5.403329	0(0)
20	4.197216	19(19)	20	5.336638	20(19)
25	4.085952	19(19)	25	5.333766	21(19)
30	4.038819	19(19)	29	5.333389	21(19)
35	4.014916	19(19)	–	–	–
40	4.011060	34(14)	–	–	–
45	4.003007	50(14)	–	–	–
50	4.001437	51(19)	–	–	–
55	4.000639	51(34)	–	–	–
60	4.000307	51(15)	–	–	–
61	4.000280	51(29)	–	–	–

ждение более точной оценки, согласно (25), (здесь $\nu_2^*(G_2^{64}) = 4$) сводится к минимизации выпуклой функции от шести переменных.

Изменение значений функционала по итерациям $r(\alpha)$ -алгоритма при нахождении оценок (24) и (25) для графа G_2^{64} приведено в табл. 5. Параметры и критерии останова $r(\alpha)$ -алгоритма выбирались следующими: $\alpha = 2$, $h_0 = 1.0$, $q_1 = 0.95$, $nh = 3$, $q_2 = 1.1$, $\varepsilon_x = 10^{-5}$, $\varepsilon_g = 10^{-5}$. В качестве начальных стартовых точек использовались $y_0 = (0, 0, 0)$ и $z_0 = (0, 0, 0, 0, 0, 0)$. Как видно из таблицы, $r(\alpha)$ -алгоритм гарантирует довольно хорошую точность решения задач по функционалу за сравнительно небольшое число вычислений функции и субградиента. По итерациям $r(\alpha)$ -алгоритма в табл. 5. приведено (в столбце N_λ) число собственных значений матрицы размером 64×64 , которые отличаются от максимального не более, чем на 0.01 и в скобках не более чем на 0.0001. Этот показатель характеризует степень овражности минимизируемой функции и, как видно из таблицы, он довольно высок как для первой, так и для второй из задач. Поэтому работу $r(\alpha)$ -алгоритма можно считать достаточно эффективной.

Общие замечания. Приведенные выше результаты тестовых экспериментов представляют лишь незначительную часть из тех, которые были проделаны для $r(\alpha)$ -алгоритма и его модификации применительно к задачам матричной оптимизации. Ряд экспериментов проделан и

для матричных задач, к которым сводятся такие задачи на графах, как нахождение максимального взвешенного независимого множества вершин графа, нахождение максимального разреза графа (в частности, для тестовых примеров из [20]). Учитывались и такие вопросы, как устойчивость методов в зависимости от выбора параметров и от выбора начальной стартовой точки. В целом проведенные тестовые эксперименты свидетельствуют о довольно устойчивой работе методов негладкой оптимизации в матричных негладких задачах.

Отметим, что вычислительный опыт решения задач квадратичного типа (14), (15) по методике, изложенной в разд. 2, позволил обнаружить и некоторые проблемы, связанные с модификацией $r(\alpha)$ -алгоритма. Так, если начальная точка выбрана далеко от границы области Ω , на первых итерациях метода имеют место большие затраты по количеству вычислений суперградиента $\tilde{\psi}(u)$ по формуле (19). Это связано с некоторым "рассогласованием" адаптивного способа регулировки шагового множителя, т.е. в начале процесса, когда длина шага велика, спуск по направлению может выводить далеко за пределы области Ω . Для того, чтобы шаг уменьшился за счет операции растяжения пространства, требуется выполнить достаточное число итераций. Подобный эффект наблюдается и в случае, когда u^* находится на границе множества Ω . Здесь "рассогласование" шага связано с тем, что норма суперградиента $\tilde{\psi}(u)$ при вычислении по формуле (18) может стремиться к нулю, а при вычислении по формуле (19) не стремится к нулю. В то же время величина шага h_k одинакова для обеих направлений, что приводит к неравномерным сдвигам по этим направлениям и препятствует останову по норме отклонения ε_x .

Один из возможных вариантов исключения такого "рассогласования" шага состоит в создании "комбинированных" алгоритмов с растяжением пространства, где спуск по направлению осуществляется с помощью адаптивного способа регулировки шага, если суперградиент $\tilde{\psi}(u)$ вычисляется по формуле (18), и с помощью классической фейеровской регулировки шагового множителя в преобразованном пространстве аргументов, если суперградиент $\tilde{\psi}(u)$ вычисляется по формуле (19). Для второго случая можно применять операцию растяжения пространства в направлении разности двух последовательных суперградиентов, вычисленных согласно классическому фейеровскому шагу. При построении "комбинированных" алгоритмов наряду с оператором растяжения пространства можно использовать и одноранговые операторы из [21], которые обеспечивают преобразование по типу r -алгоритмов и позволяют проводить внешнюю аппроксимацию множества экстремумов мо-

нотонно уменьшающимися по объему эллипсоидами. Разработка "комбинированных" алгоритмов целесообразна по той причине, что как и модификация $r(\alpha)$ -алгоритма, такие методы позволяют при решении общей задачи выпуклого программирования избежать введения негладких штрафов.

Литература

1. PARDALOS, P.M. AND ROSEN J.B. Constrained Global Optimization: Algorithms and Applications // Lecture Notes in Computer Science 268. – New York: Springer-Verlag, 1987. – 268. – 143 p.
2. ШОР Н.З., СТЕЦЕНКО С.И. Квадратичные экстремальные задачи и недифференцируемая оптимизация. – Киев: Наук. думка, 1989. – 208 с.
3. ШОР Н.З. Об одном подходе к получению глобальных экстремумов в полиномиальных задачах математического программирования // Кибернетика. – 1987. – № 5. – С. 102–106.
4. ШОР Н.З. Об одном классе оценок глобального минимума полиномиальных функций // Кибернетика. – 1987. – № 6. – С. 9–11.
5. Проблемы Гильберта / Под. ред. П.С.Александрова. – М.: Наука, 1969. – 238 с.
6. ШОР Н.З. Задачи квадратичного типа // Изв. АН СССР. Техн. кибернетика. – 1987. – № 1. – С. 128–139.
7. HILBERT D. Über die Darstellung definiter Formen als Summen von Formen quadraten // Math. Ann. – Leipzig, 1888. – Bd. 22. – S. 342–350.
8. ARTIN E. Über die Zerlegung definiter Functionen in Quadrate // Hamburg Abh. – 1927. – № 5. – S. 100–115.
9. POLYA G. Über positive Darstellung von Polynomen // Vierteljahrsschrift d. naturforschenden Gesellsch., Zurich 73. – 1928. – S. 141–145.
10. SHOR N.Z. Dual Estimates in Multiextremal Problems // Journal of Global Optimization. – 1992. – № 2. – P. 411–418

11. ШОР Н.З., ЖУРБЕНКО Н.Г. Метод минимизации, использующий операцию растяжения пространства в направлении разности двух последовательных градиентов // Кибернетика, 1971. – № 3. – С. 51–59.
12. ШОР Н.З. Методы минимизации недифференцируемых функций и их приложения. - Киев: Наук. думка, 1979. – 199 с.
13. ЖУРБЕНКО Н.Г., МАРЧУК Т.В. Алгоритм минимизации негладких функционалов ($r(\alpha)$ -алгоритм) // АН УССР. РФАП. – 1976. – № 22. – 326 с.
14. ШОР Н.З. Минимизация матричных функций и недифференцируемая оптимизация // Обзорение прикладной и промышленной математики. – 1995. – Т. 2. – С. 113–138.
15. LEMARECHAL C. Numerical experiments in nonsmooth optimization. In: Progress in nondifferentiable optimization / Ed. E.A.Nurminski. **CP-82-S2**, Intern. Inst. for Applied System Analysis – Laxenburg, Austria, 1982. – P. 61–84.
16. НЕСТЕРОВА С.И., СКОКОВ В.А. Численный анализ программ негладкой безусловной оптимизации // Экономика и мат. методы / – 1994. – **30**. – № 2. – С. 136–145.
17. Программный комплекс решения систем линейных алгебраических уравнений с автоматическим выбором метода решения (ПК АРАС) / И.Н. Молчанов, Л.Д. Николенко, М.Ф. Яковлев и др. – ГРФАП, 1984 г. – № 007135.
18. Сборник научных программ на ФОРТРАНе. – Руководство для программиста. – Вып. 2. Матричная алгебра и линейная алгебра. – М.: Статистика, 1974. – 223 с.
19. LOVASZ L. On the Shannon capacity of a graph // IEEE Trans. Inform. Theory. – 1979. – **25**. – № 1. – P. 1–7.
20. ШОР Н.З., БЕРЕЗОВСКИЙ О.А. Новые алгоритмы решения задачи о максимальном разрезе графа // Кибернетика и систем. анализ. – 1995. – № 2. – С. 100–106.
21. СТЕЦЮК П.И. Ортогонализирующие линейные операторы в выпуклом программировании. I // Кибернетика и систем. анализ. – 1997. – № 3. – С. 97–119.

Роль избыточных ограничений в улучшении двойственных оценок для полиномиальных оптимизационных задач

Н. З. Шор

Кибернетика и системный анализ. – 1998. – № 4. – С. 106–121.

При решении задач нелинейного программирования с полиномиальными целевыми функциями и ограничениями (к этому классу относятся также полиномиальные булевы задачи) возникает возможность путем введения новых переменных и ограничений понизить степень полиномиальных функций, входящих в условия преобразованной задачи, до второй степени. Поэтому достаточно рассмотреть задачи минимизации квадратичного типа:

найти

$$q^* = \inf_{x \in E^n} Q_0(x) \quad (1)$$

при ограничениях

$$Q_i(x) = 0, \quad i \in I = \{1, \dots, N\}, \quad (2)$$

где $Q_\nu(x)$ – квадратичные (линейные) функции,

$$Q_\nu(x) = (K_\nu x, x) + (b_\nu, x) + c_\nu,$$

где K_ν – симметричные матрицы $n \times n$, $b_\nu \in E^n$, c_ν – числа, $\nu \in \{0 \cup I\}$.

Оценки снизу для q^* могут быть получены путем лагранжевой релаксации. Пусть $u = (u_1, \dots, u_N) \in R^N$ – вектор множителей Лагранжа задачи (1), (2). Рассмотрим функцию Лагранжа

$$L(x, u) = Q_0(x) + \sum_{i=1}^N u_i Q_i(x)$$

и функцию

$$\psi(u) = \inf_x L(x, u) = \inf_x [(K(u)x, x) + (b(u), x) + c(u)],$$

где

$$K(u) = K_0(x) + \sum_{i=1}^N u_i K_i(x),$$

$$b(u) = b_0 + \sum_{i=1}^N u_i b_i,$$

$$c(u) = c_0 + \sum_{i=1}^N u_i c_i.$$

Пусть $\Omega^+ = \{u : \lambda_n(K(u)) > 0\}$; Ω^- – множество u , для которых $K(u)$ имеет отрицательные собственные числа; Ω^0 – множество u , у которых $\lambda_n[K(u)] = 0$. (Здесь $\lambda_1(K) \geq \lambda_2(K) \geq \dots \geq \lambda_n(K)$ – собственные числа симметричной матрицы K размера $n \times n$, упорядоченные в порядке убывания).

Область определения функции $\psi(u)$ (или $\text{dom } \psi$) состоит из Ω^+ и подмножества точек $u \in \Omega^0$, для которых разрешима система уравнений

$$2K(u)x + b(u) = 0. \quad (3)$$

Для остальных точек $\psi(u) = -\infty$.

Если $\text{dom } \psi \neq \emptyset$, то существует нетривиальная оценка снизу для q^*

$$\varphi^* = \sup_{u \in \text{dom } \psi} \psi(u)$$

(условие $\varphi^* = +\infty$ означает, что система (2) несовместна).

Если φ^* достигается на $u^* \in \Omega^+$, то

$$\varphi^* = \psi(u^*) = q^* = Q_0(x(u^*)),$$

где $x(u^*)$ — решение системы (3) при $u = u^*$. В противном случае φ^* достигается на границе области Ω^+ , при этом может существовать положительный "разрыв" ("gap")

$$\Delta^* = q^* - \varphi^* > 0.$$

Один из способов уменьшения этого разрыва связан с введением функционально избыточных ограничений (при этом может увеличиться и множество переменных). Функционально-избыточными ограничениями назовем ограничения, добавление которых оставляет множество

оптимальных решений первоначальной задачи неизменным. Однако их введение меняет функцию Лагранжа, что может в некоторых случаях уменьшить разрыв между оптимальным значением q^* целевой функции и лагранжевой (двойственной) оценкой φ^* . Такой способ улучшения двойственных оценок был исследован в [1], [2], [3].

Рассмотрим несколько многоэкстремальных квадратичных и булевых комбинаторных задач, на которых такой подход проявляется особенно ярко.

1. Задача о максимальном взвешенном независимом множестве вершин графа

Пусть задан неориентированный граф $G(V, E)$ с множеством вершин $V = \{1, 2, \dots, n\}$ и множеством ребер E ; (i, j) обозначает ребро с вершинами i и j ($i \neq j$). Подмножество вершин $\bar{I} \subseteq V$ называется независимым, если все вершины из \bar{I} попарно не соединены ребрами. Подмножество $\bar{K} \subseteq V$ называется кликой, если любые пары (i, j) ($i \neq j$), $i, j \in \bar{K}$, являются ребрами.

Граф $\bar{G}(V, \bar{E})$ называется дополнительным к графу G , если $(i, j) \in \bar{E}$ тогда и только тогда, когда $(i, j) \notin E$ ($i \neq j$). Ясно, что независимое подмножество вершин графа G является кликой для дополнительного графа \bar{G} .

Пусть каждой вершине $k \in V$ сопоставлен вес $w_k > 0$, $I \subseteq V$ — произвольное независимое подмножество графа. Введем булевые переменные x_k для всех $k \in V$:

$$x_k = \begin{cases} 1, & \text{если } k \in I, \\ 0, & \text{если } k \notin I, \end{cases} \quad x = \{x_1, \dots, x_n\}.$$

Тогда задача о максимальном взвешенном независимом подмножестве вершин графа $G = (V, E)$ может быть сформулирована как задача линейного булева программирования следующим образом: найти

$$\alpha_w(G) = \max_x \sum_{k=1}^n w_k x_k; \quad x \in \{0, 1\}^n \quad (4)$$

при ограничениях

$$x_i + x_j \leq 1 \quad \forall (i, j) \in E, \quad (5)$$

$$0 \leq x_k \leq 1 \quad \forall k \in V. \quad (6)$$

Обозначим $M_I(G)$ выпуклую оболочку булевых векторов x , удовлетворяющих ограничениям (5); M — многогранник допустимых векторов, удовлетворяющих системе ограничений (5), (6). Можно доказать, что $M_I(G) = M$ тогда и только тогда, когда граф G является двудольным. В этом случае поиск $\alpha_w(G)$ сводится при любом $w > 0$ к задаче линейного программирования.

С каждым нечетным циклом C_{2k+1} графа G можно связать ограничение

$$\sum_{i \in C_{2k+1}} x_i \leq k \quad (k - \text{натуральное число}). \quad (7)$$

Пусть Z_{2k+1} — множество нечетных циклов длины $2k + 1$, являющихся подграфами графа G . Граф называется t -графом, если для него многогранник $M_I(G)$ совпадает с многогранником, задаваемым линейными ограничениями (5), (6) и следующими ограничениями:

$$\sum_{i \in C} x_i \leq k \quad \text{для всех } C \in Z_{2k+1}; \quad k = 1, 2, \dots, (n-1)/2. \quad (8)$$

Для t -графов существует полиномиальный алгоритм определения оптимального взвешенного независимого подмножества $X_w^*(G)$ при целочисленных w , основанный на методе эллипсоидов, при этом используется полиномиальный алгоритм определения минимального нечетного цикла при заданных \bar{x} для проверки ограничений (8) [4].

Задачу линейного программирования с целевой функцией (4) и ограничениями (5), (6), (8) можно свести к задаче минимизации выпуклой негладкой функции, используя метод негладких штрафных функций в форме функции максимума невязок в ограничениях, при этом алгоритм нахождения минимального нечетного цикла может быть использован для нахождения максимальной невязки в ограничениях вида (8). Минимизация штрафной функции может быть произведена с использованием r -алгоритма.

Задача о максимальном взвешенном независимом множестве может быть представлена также в виде следующей полиномиальной задачи квадратичного типа: найти

$$\alpha_w(G) = \max \sum_{i=1}^n w_i x_i \quad (9)$$

при ограничениях

$$x_i x_j = 0 \quad \forall (i, j) \in E, \quad (10)$$

$$x_k^2 - x_k = 0 \quad \forall k \in V. \quad (11)$$

Л. Ловас предложил ряд совпадающих по величине оценок сверху для $\alpha_w(G)$ [5]. Эти оценки являются результатом решения определенных задач матричной оптимизации. Рассмотрим некоторые из этих оценок. Пусть $x = \{x_{ij}\}_{(i,j) \in E}$. Элементы $a_{ij}(x)$ симметричной матрицы $A(x)$ размера $n \times n$ строятся следующим образом:

$$a_{ij}(x) = \begin{cases} x_{ij} & \text{для } (i, j) \in E, \\ \sqrt{w_i w_j} & \text{в остальных случаях,} \end{cases}$$

$\lambda_1(A)$ означает максимальное собственное число матрицы A . Первая оценка Ловаса строится по формуле

$$\nu_w^{(1)}(G) = \min_x \lambda_1(A(x)) \geq \alpha_w(G).$$

Следующая в определенном смысле двойственная к предыдущей задаче также служит для получения оценки Ловаса: рассматривается класс неотрицательно определенных симметричных матриц $B(y)$ размера $n \times n$ с элементами

$$b_{ij} = \begin{cases} 0 & \text{для } (i, j) \in E, \\ y_{ij} & \text{для остальных пар } (i, j). \end{cases}$$

Требуется найти

$$\nu_w^{(2)}(G) = \max_y \sum_{(i,j)} \sqrt{w_i w_j} b_{ij}(y)$$

при ограничениях

$$B(y) \geq 0; \quad \sum_{i=1}^n b_{ii}(y) \leq 1.$$

В работах [5, 6] показано, что $\nu_w^{(1)}(G) = \nu_w^{(2)}(G)$.

Как известно, для любого графа G $\alpha(G) \leq w(G)$, где $w(G)$ – минимальное число клик, которыми можно покрыть граф. Граф называется совершенным, если для графа G и любого его индуцированного подграфа G' числа $\alpha(G)$ и $w(G')$ совпадают.

Пусть q – произвольная клика графа G , Q – множество всех клик. Запишем набор так называемых кликовых ограничений для задачи (4)–(6)

$$\sum_{i \in q} x_i \leq 1 \quad \text{для всех } q \in Q. \quad (12)$$

Доказано [4], что граф G является совершенным тогда и только тогда, когда $M_I(G)$ определяется ограничениями (5) и (12); при этом оценка Ловаса будет точной и совпадет с оценкой линейного программирования, т.е. с оптимальным значением задачи (4)–(6) для любого $w \geq 0$. Отметим, что задача получения оценок Ловаса является полиномиально разрешимой. Алгоритм, полиномиальный по времени, использующий модель получения $\nu^{(2)}(G)$, был предложен в статье [7].

Как показано в [1] (см. также [3]), двойственная лагранжевая оценка для задачи (9)–(11) совпадает по величине с оценкой Ловаса.

Если добавить к ограничениям вида (10), (11) тривиальные избыточные ограничения вида $x_i x_j \geq 0$ для всех $(i, j) \notin E$, $i \neq j$, то лагранжевая двойственная оценка для модифицированной задачи может оказаться существенно лучше, чем φ^* . М.Р.Бест рассмотрел следующий граф $G(V, E)$: множество вершин V – это булевы шестимерные точки $\{0, 1\}^6$, и две вершины считаются смежными, если расстояние Хэмминга между ними не больше трех. Максимальные независимые подмножества вершин содержат по четыре вершины (например, (000000), (110000), (001100), (000011)). Двойственная оценка, использующая ограничения вида (10), (11), равна $16/3$, добавление ограничений вида $x_i x_j \geq 0$ для всех $(i, j) \notin E$, $i \neq j$, приводит к оценке 4 [6].

Для любого $k \in V$ и $(i, j) \in E$ справедливы квадратичные ограничения

$$x_k(x_i + x_j) \leq x_k.$$

С целью улучшения двойственных оценок можно рассматривать расширенное множество ограничений для задачи о максимальном взвешенном независимом множестве графа $G(V, E)$, собрав, например, все ранее выписанные квадратичные ограничения, при этом получается следующая модель: найти

$$\max \sum_{i=1}^n w_i x_i \quad (13)$$

при ограничениях

$$x_i x_j = 0; \quad (i, j) \in E; \quad (14)$$

$$x_i x_j \geq 0; \quad (i, j) \notin E; \quad (15)$$

$$x_k^2 - x_k = 0; \quad k = 1, \dots, n; \quad (16)$$

$$x_k(x_i + x_j) \leq x_k; \quad k = 1, \dots, n; \quad (i, j) \in E. \quad (17)$$

Отметим, что ограничения (15), (17) являются избыточными, причем ограничения для нечетных циклов вида (7) являются следствиями ограничений (17).

Мы можем образовать расширенную функцию Лагранжа $L^p(x, u)$, сопоставляя ограничениям (14), (16) и всем или части ограничений (15), (17) множители Лагранжа, и получить, используя методы выпуклой негладкой оптимизации (например, r -алгоритм) оценку

$$\varphi^* = \inf_{\{u \in u_{\text{дон}}, u \in \text{dom } \psi\}} \psi(u),$$

где $\psi(u) = \sup_x L^p(x, u)$. Соответствующие вычислительные эксперименты описаны в [8].

2. Задача о максимальном разрезе графа

Задача о максимальном разрезе графа $G = (V, E)$ заключается в разбиении множества его вершины V на два непустых и непересекающихся подмножества V_1 и V_2 , $V_1 \cup V_2 = V$ таким образом, чтобы вес разреза (т.е. совокупности ребер, концы которых лежат в разных подмножествах) был максимален. Вес разреза определяется как сумма весов его ребер. Вес каждого ребра $e \in E$ является произвольным вещественным числом q_e .

Пусть $|V| = n$ и $\{1, 2, \dots, n\}$ – номера вершин; ребро e , соединяющее вершины i и j , обозначим (i, j) или (j, i) , $i, j \in V$.

Без ограничения общности можно считать, что граф G является полным, так как ребрам дополнительного графа \overline{G} можно придать нулевой вес. Сопоставим каждой вершине k , $k \in \{1, 2, \dots, n\}$, переменную x_k . Для заданного разбиения $V_1 \cup V_2 = V$ будем считать, что

$$x_k = \begin{cases} +1, & \text{если } k \in V_1, \\ -1, & \text{если } k \in V_2. \end{cases}$$

В этих обозначениях задача о максимальном разрезе может быть записана как задача нелинейного программирования квадратичного типа: найти $\max f(x)$, где

$$f(x) = \frac{1}{8} \sum_{(i,j)} q_{ij} (x_i - x_j)^2 \quad (18)$$

при ограничениях

$$x_k^2 = 1 \quad \forall k \in V. \quad (19)$$

С учетом (19)

$$f(x) = \frac{1}{4} \left(\sum_{(i,j)} q_{ij} - \sum_{(i,j)} q_{ij} x_i x_j \right). \quad (20)$$

Таким образом, максимизация $f(x)$ сводится к нахождению минимума $f_1(x)$, где

$$f_1(x) = \sum_{(i,j)} q_{ij} x_i x_j$$

при ограничениях (19).

Пусть $\min f_1(x) = f_1^*$. Для получения нижних оценок для f_1^* применим метод множителей Лагранжа. Пусть $u = \{u_k\}_{k=1}^n$ — вектор множителей Лагранжа, соответствующий системе ограничений (19). Составим функцию Лагранжа

$$L_1(x, u) = f_1(x) + \sum_{k=1}^n u_k (x_k^2 - 1).$$

Пусть $Q = \{q_{ij}\}_{i,j=1}^n$ — симметричная матрица с нулевой диагональю. Тогда

$$L_1(x, u) = ((Q + \text{diag}(u))x, x) - \sum_{k=1}^n u_k \equiv (Q(u)x, x) - \sum_{k=1}^n u_k,$$

где $Q(u) = Q + \text{diag}(u)$.

Двойственная оценка для f_1^* находится как

$$\bar{f}_1(x) = \sup_u \inf_x L_1(x, u).$$

Легко видеть, что

$$\inf_x L_1(x, u) = \begin{cases} -\sum_{k=1}^n u_k & \text{для неотрицательно определенных } Q(u), \\ -\infty & \text{в остальных случаях.} \end{cases}$$

Отсюда

$$-\bar{f}_1 = \min \sum_{k=1}^n u_k \quad \text{при ограничении} \quad \lambda_n[Q(u)] \geq 0, \quad (21)$$

где $\lambda_n(A)$ означает минимальное собственное число симметричной матрицы A размерности $n \times n$. Для решения задачи (21) можно применить метод негладких штрафных функций, сводя ее к задаче безусловной минимизации: найти

$$\rho(N) = \min_u \left(\sum_{k=1}^n u_k - N \min\{0, \lambda_n(A(u))\} \right),$$

где N – штрафной коэффициент. Доказано, что при $N \geq n$ справедливо $\rho(N) = -\bar{f}_1$ [9].

После получения $-\bar{f}_1$ верхняя оценка \bar{f} для максимального разреза получается из (20) по формуле

$$\bar{f} = \frac{1}{4} \left(\sum_{(i,j)} q_{ij} - \bar{f}_1 \right).$$

Другой способ получения оценок для задачи о максимальном разрезе связан с использованием реберных переменных.

Для каждого ребра $(i, j) \in E$ введем булеву переменную y_{ij} , которая принимает значение единицы, если (i, j) принадлежит разрезу, и нулю в противном случае. Каждому разбиению множества вершин $V = V_1 \cup V_2$ множество ребер, соответствующих разрезу, образует двудольный частичный подграф графа G . Таким образом, множеству разрезов можно сопоставить в пространстве R^E совокупность инцидентных разрезам булевых векторов $y\{R_\alpha\}$ (α – индекс разреза), т.е.

$$y(R_\alpha) = \{y_{ij}^\alpha\}_{(i,j) \in E}; \quad y_{ij}^\alpha = \begin{cases} 1, & \text{если } (i, j) \in R_\alpha, \\ 0, & \text{если } (i, j) \notin R_\alpha. \end{cases}$$

Отметим, что координаты (i, j) и (j, i) отождествляются. Пусть M_R^G – выпуклая оболочка векторов $y(R_\alpha)$. Тогда задача о максимальном разрезе эквивалентна нахождению

$$mc_Q(G) = \max_{y \in M_R^G} \sum_{(i,j) \in E} q_{ij} y_{ij}. \quad (22)$$

Выпуклый многогранник M_R^G с булевыми вершинами для нетривиальных задач о максимальном разрезе устроен очень сложно. Однако, имеется подмножество его граней, которое описывается сравнительно просто. Пусть $F \subseteq E$ – некоторое подмножество ребер, $|F|$ – число ребер в F ; $C \subseteq E$ – произвольная циклическая цепь (цикл) из ребер, $q(M) = \sum_{(i,j) \in M} q_{ij}$; M – произвольное подмножество ребер. Легко показать, что для любого разреза R , цикла C и нечетного подмножества $F \subseteq C$ выполняется неравенство

$$\sum_{(i,j) \in F} y_{ij} - \sum_{(i,j) \in (C \setminus F)} y_{ij} \leq |F| - 1. \quad (23)$$

Система ограничений вида (23) определяет в R^E некоторый выпуклый многогранник, который содержит M_R^G , но в некоторых случаях,

например для так называемых слабо двудольных графов, совпадает с M_R^G .

Определение. Граф $G = (V, E)$ называется слабо двудольным, если система ограничений

$$\sum_{(i,j) \in C} y_{ij} = |C| - 1,$$

где C – множество произвольных нечетных циклов, задает в пространстве R^E многогранник, совпадающий с M_R^G . Эта система неравенств соответствует случаю $C \equiv F$ в (23), т.е. неравенства (23) выписываются для каждого нечетного цикла графа G .

Задача о нахождении максимального разреза графа после замены условия булевости переменных y_{ij} на двухсторонние ограничения на переменные

$$0 \leq y_{ij} \leq 1 \quad \forall (i, j) \in E \quad (24)$$

может быть релаксирована к задаче линейного программирования с целевой функцией (22) и ограничениями вида (23) и (24). Отметим, что как показано в [10], множество решений такой задачи линейного программирования совпадает с множеством решений задачи о максимальном разрезе тогда и только тогда, когда G не имеет подграфов, стягиваемых к K_5 (см. также [11]).

В общем случае даже для небольших задач количество ограничений вида (23) чрезвычайно велико. Однако для проверки выполнимости ограничений вида (23) для заданных $\{\bar{y}_{ij}\}$, удовлетворяющих условиям (24), можно сформулировать специальную задачу о кратчайшем пути для некоторого графа $\bar{G}(G, \bar{y})$, который легко получается из графа G следующим образом ([4], стр. 250).

Построим новый граф $H = (V' \cup V'', E' \cup E'' \cup E''')$, состоящий из двух разделенных копий графа G : $G' = (V', E')$, $G'' = (V'', E'')$ и дополнительного множества ребер E''' , которое содержит для каждого $(i, j) \in E$ два ребра: i', j'' и i'', j' . Вес ребер $(i', j') \in E'$ и $(i'', j'') \in E''$ принимается равным \bar{y}_{ij} , а вес ребер i', j'' и i'', j' равен $1 - \bar{y}_{ij}$. Для каждой вершины $i \in V$ вычисляется кратчайший путь в графе H из вершины $i' \in V'$ к вершине $i'' \in V''$, при этом длины ребер берутся в соответствии с указанным выше правилом. Такой путь содержит нечетное число ребер в E''' и соответствует замкнутому пути (циклу) в G . Ясно, что если кратчайший из этих путей имеет длину по крайней мере 1, то \bar{y} удовлетворяет условиям (24), в противном случае существуют цикл C и множество $F \subseteq C$ с нечетным $|F|$, для которого \bar{y}

не удовлетворяет соответствующему неравенству. В этом случае градиент к такому неравенству может быть использован для построения отсекающей гиперплоскости в методах отсечения (например, в методе эллипсоидов или в методе вписанных эллипсоидов) либо для построения направления движения в методах субградиентного типа (например, в r -алгоритме).

Реберные переменные y_{ij} (i, j) $\in E$ легко выражаются через вершинные переменные x_k , $k \in V$:

$$y_{ij} = \frac{1 - x_i x_j}{2}. \quad (25)$$

Действительно, если (i, j) принадлежит разрезу, то $x_i x_j = -1$ и $y_{ij} = 1$. Если (i, j) не принадлежит разрезу, то $x_i x_j = 1$ и $y_{ij} = 0$.

Пусть множество ребер $F \subseteq C$, где C — некоторый цикл и $|F|$ нечетно. Используя соотношения (25), запишем ограничения (23) через вершинные переменные

$$\sum_{(i,j) \in F} \frac{1 - x_i x_j}{2} - \sum_{(i,j) \in C \setminus F} \frac{1 - x_i x_j}{2} \leq |F| - 1,$$

$$\frac{1}{2}|F| - \frac{1}{2}(|C| - |F|) - |F| + 1 \leq \frac{1}{2} \left(\sum_{(i,j) \in F} x_i x_j - \sum_{(i,j) \in C \setminus F} x_i x_j \right)$$

или

$$\sum_{(i,j) \in F} x_i x_j - \sum_{(i,j) \in C \setminus F} x_i x_j \geq 2 - |C|. \quad (26)$$

Таким образом, в вершинных переменных $x = \{x_k\}_{k=1}^n$ ($x_k \in \{1, -1\}$ при всех k) для любого нечетного множества F , входящего в цикл C , справедливо неравенство (26).

Выше описан алгоритм проверки ограничений (23), заданных в реберных переменных, эквивалентных ограничениям (26). Сопоставим каждому ограничению вида (26) двойственные переменные $u(F, C)$ и рассмотрим новую задачу получения оценки сверху для $mc_Q(G)$. Найти

$$\max f(x) = \max \frac{1}{8} \sum_{(i,j) \in E} (x_i - x_j)^2 q_{ij} \quad (27)$$

при ограничениях

$$x_k^2 - 1 = 0 \quad \forall k \in V; \quad (28)$$

$$\sum_{(i,j) \in C} x_i x_j - 2 \sum_{(i,j) \in F} x_i x_j + (2 - |C|) \leq 0 \quad (29)$$

для произвольного цикла C и произвольного множества ребер F , $F \subseteq C$, $|F|$ нечетно.

В задачу (27)–(29) не обязательно включать все ограничения вида (29); таких ограничений может быть достаточно много даже для сравнительно небольших графов. Обозначим через \bar{M} подмножество ограничений вида (29), учитываемых на текущей стадии решения задачи. Присоединяя к ограничениям (28) функционально избыточные ограничения вида (29), можно существенно улучшить двойственные оценки, сформулировав эквивалентную задачу в новой форме.

Рассмотрим функцию Ланранжа $\Phi(x, u)$ задачи (27)–(29). Пусть вектор множителей Лагранжа

$$u = (\{u_k\}_{k=1}^n, \{u[F, C]\}_{[F, C] \in \bar{M}}).$$

Тогда

$$\begin{aligned} \Phi(x, u) = & f(x) + \sum_{k=1}^n u_k (x_k^2 - 1) + \sum_{[F, C] \in \bar{M}} u[F, C] \times \\ & \times \left(\sum_{(i,j) \in C} \frac{x_i x_j + x_j x_i}{2} - 2 \sum_{(i,j) \in F} \frac{x_i x_j + x_j x_i}{2} + (2 - |C|) \right) \end{aligned}$$

(неравенства вида (29) мы рассмотрим в симметричной форме относительно пар индексов (i, j) и (j, i)).

Пусть $K(u)$ – матрица квадратичной по x формы функции Лагранжа $\Phi(x, u)$ при фиксированном u . Тогда $\Phi(x, u)$ можно записать в виде

$$\Phi(x, u) = (K(u)x, x) - \left(\sum_{k=1}^n u_k + \sum_{[F, C] \in \bar{M}} u[F, C](2 - |C|) \right).$$

При любом $\bar{u} = (\{\bar{u}_k\}_{k=1}^n, \{\bar{u}[F, C] \geq 0\}_{[F, C] \in \bar{M}})$ функция $\varphi(\bar{u}) = \sup_x \Phi(x, \bar{u})$ дает оценку сверху для $mc_Q(G)$. Если $K(\bar{u})$ не является отрицательно полуопределенной, то $\varphi(\bar{u}) = +\infty$, и в результате получим тривиальную оценку. Поэтому мы должны выделить область значений \bar{u} , удовлетворяющих $K(\bar{u}) \leq 0$, на которой функция φ определена. При этом $\sup_x \Phi(x, \bar{u})$ достигается в точке $x = 0$ и

$$\varphi(\bar{u}) = - \sum_{k=1}^n \bar{u}_k - \sum_{[F, C] \in \bar{M}} \bar{u}[F, C](2 - |C|).$$

Таким образом, задача нахождения оценки Лагранжа ψ^* для $mc_Q(G)$ сводится к минимизации $\varphi(\bar{u})$: $\min \varphi(u)$ при ограничениях: $\lambda_1(K(u)) \leq 0$ и $u[F, C] \geq 0$ для всех циклов C и $F \subseteq C$, $|F|$ нечетно.

Алгоритм отсеечения [10] дает возможность использовать метод эллипсоидов и построить полиномиальный алгоритм решения задачи линейного программирования с целевой функцией (22) и ограничениями (23), (24), в частности, решить задачу о максимальном разрезе в случае графов, не стягиваемых к K_5 , за полиномиальное время. Для плоских графов полиномиальные алгоритмы, использующие алгоритм паросочетаний, были предложены в работах [12, 13]; для графов, не стягиваемых к K_5 , в [14] предложен комбинаторный полиномиальный алгоритм, основанный на идеях декомпозиции. Упрощенная версия метода отсеечения использована в [15] для построения полиномиального алгоритма для случая слабо двудольных графов. Отметим, что этот класс содержит графы, не стягиваемые к K_5 [16, 17].

Релаксация задач о максимальном разрезе к задачам линейного программирования с ограничениями вида (23), как показано в [18], дает хорошие результаты при решении прикладных задач значительной размерности из области физики и проектирования электронных схем.

Задача линейного программирования с целевой функцией (22) и ограничениями (23),(24) может иметь большое количество ограничений даже для небольших значений n , число переменных равно $|E|$ и не превышает $\frac{n(n-1)}{2}$. В ряде работ описано успешное использование при решении таких задач симплекс-метода с генерацией строк (23) для получения оценок сверху для целевой функции максимального разреза графа и нахождения близкого к оптимальному максимального разреза (см., например, [19]). Схему генерации строк-ограничений вида (23) можно также использовать в сочетании с r -алгоритмом для минимизации негладкой штрафной функции вида

$$S_N(y) = -(q, y) + N \max\{0; \max_i \{-y_i, y_i - 1\}, \max_{[F, C] \in M} [\sum_{k \in F} y_k - \sum_{k \in C \setminus F} y_k - |F| + 1]\},$$

где N —достаточно большой штрафной множитель, при этом определение

$$\varphi(y) = \max_{[F, C] \in M} [\sum_{k \in F} y_k - \sum_{k \in C \setminus F} y_k - |F| + 1]$$

проводится с помощью описанной выше процедуры нахождения кратчайших путей.

Указанный способ нахождения оценок в задаче о максимальном разрезе графа был проверен для тестовых примеров из [9] и для графа в форме икосаэдра, весовые коэффициенты ребер которого приведены в табл. 1.

Рассмотрим общую схему алгоритма решения двойственной задачи к задаче (27)–(29) с использованием r -алгоритма. Пусть на очередном шаге мы имеем множество \overline{M} ограничений вида (29) и вектор \overline{x} вершинных переменных; по формулам (25) находим соответствующие значения реберных переменных $\overline{y} = \{\overline{y}_{ij}\}_{(i,j) \in E}$. Используя алгоритм кратчайших путей, проверим, удовлетворяет ли \overline{y} всем ограничениям вида (23). При положительном ответе переходим к очередному шагу r -алгоритма, не применяя функции Лагранжа. В противном случае находим ограничение вида (23) с максимальной невязкой и строим соответствующее квадратичное ограничение вида (26) в вершинных переменных, добавляем его к системе ограничений вида (29), при этом появляется новый множитель Лагранжа $\overline{u}[F, C]$. Присвоим ему значение 0 и перейдем к очередному шагу r -алгоритма в пространстве двойственных переменных.

Для ряда графов были проведены вычислительные эксперименты. В табл. 2 приведены результаты экспериментов для графа в виде икосаэдра, который содержит 12 вершин и 30 ребер (веса его ребер см. в табл. 1). Тестовый эксперимент для этого графа состоит в следующем. Сначала находится двойственная оценка сверху для задачи (18)–(19), которая равна 665,53. Далее по эвристическому алгоритму из [9] получено допустимое решение с оценкой 642, при этом был найден следующий разрез: $V_1 = \{1, 2, 9, 10, 11, 12\}$, $V_2 = \{3, 4, 5, 6, 7, 8\}$. Затем последовательно добавлялись неравенства вида (29) при $F = C$ и $|C| = 3$. В табл. 2 в строке i, j, k представлены номера вершин треугольных циклов. Рассмотрены результаты расчетов, когда добавляется один цикл, два цикла и так далее до десяти циклов. Поскольку все ограничения вида (29) являются избыточными по отношению к модели (18)–(19), то следует ожидать уменьшения двойственной оценки, которая представлена в графе φ^* таблицы 2. Действительно, оценка при десяти циклах отличается от достигнутого рекорда меньше, чем на единицу, что доказывает оптимальность полученного допустимого решения. Эксперимент показал эффективность использования избыточных ограничений для улучшения двойственных оценок.

Таблица 1.

номер ребра	i	j	w_{ij}	номер ребра	i	j	w_{ij}	номер ребра	i	j	w_{ij}
1	1	2	20.0	11	3	9	18.0	21	7	8	18.0
2	1	3	30.0	12	3	10	15.0	22	7	9	27.0
3	1	4	40.0	13	4	5	32.0	23	7	10	36.0
4	1	5	50.0	14	4	10	24.0	24	7	11	45.0
5	1	6	60.0	15	4	11	20.0	25	7	12	54.0
6	2	3	16.0	16	5	6	40.0	26	8	9	14.0
7	2	6	48.0	17	5	11	30.0	27	8	12	42.0
8	2	8	12.0	18	5	12	25.0	28	9	10	21.0
9	2	9	10.0	19	6	8	30.0	29	10	11	28.0
10	3	4	24.0	20	6	12	36.0	30	11	12	35.0

Таблица 2.

номер цикла	i, j, k	φ^*	номер цикла	i, j, k	φ^*
1	6,8,12	665.53	6	5,6,12	647.06
2	1,5,6	661.94	7	2,6,8	646.72
3	1,4,5	657.13	8	5,11,12	646.41
4	7,11,12	652.82	9	7,8,12	645.14
5	7,9,10	650.66	10	7,10,11	642.90

3. Оценки для глобального минимума полиномиальных функций

Ярким примером, где большую роль в улучшении двойственных оценок играют избыточные ограничения, служит задача нахождения глобального минимума полиномиальной функции от одной или нескольких переменных. Рассмотрим сначала случай полинома степени $n = 2k$ от одной переменной x_1

$$P(x_1) = x_1^{2k} + \sum_{i=1}^{2k} a_i x_1^{2k-i}.$$

Введем новые переменные x_2, \dots, x_k , где $x_{i+1} = x_1 x_i$, $i = 1, \dots, k-1$. В новых переменных $x = \{x_i\}_{i=1}^k$ полином $P(x_1)$ может быть представлен (неоднозначно) как квадратичная функция

$$P(x_1) \sim K_0(x) = x_k^2 + \sum_{r=1}^k a_r^{2k-r} x_k x_{k-r} + \sum_{r=k+1}^{2k} a_r^{2k-r} x_{2k-r}.$$

Таким образом, задача нахождения $P^* = \min_{x_1} P(x_1)$ свелась к нахождению минимального значения задачи квадратичного типа

$$\min K_0(x) \quad \text{при ограничениях} \quad x_1 x_i - x_{i+1} = 0; \quad i = 1, \dots, k-1. \quad (30)$$

Пусть $L_0(x, u) = K_0(x) + \sum_{i=1}^{k-1} u_i (x_1 x_i - x_{i+1})$ — функция Лагранжа задачи и $\psi_0(u) = \inf_x L_0(x, u)$.

Во многих случаях $\text{dom } \psi_0 = \emptyset$ и, казалось бы, лагранжева релаксация в данном случае не применима. Однако, если ввести избыточные ограничения вида

$$x_d - x_{d-r} x_r \leq 0 \quad \forall r: \quad 1 \leq r \leq \frac{d}{2} \quad \text{при} \quad d \leq k, \quad (31)$$

$$x_k x_{d-k} - x_{d-r} x_r \quad \forall r: \quad r \leq \frac{d}{2} \quad \text{при} \quad k < d \leq 2k-1, \quad (32)$$

то, построив соответствующую функцию Лагранжа $L(x, \bar{u})$ (\bar{u} — расширенный вектор множителей Лагранжа, компоненты которого соответствуют как первоначальным ограничениям (30), так и избыточным

ограничениям (31),(32)), получим, что область определения $\psi(\bar{u})$ не пуста и

$$\varphi^* = \sup_{\bar{u} \in \text{dom } \psi} \psi(\bar{u}) = P^*,$$

т.е. двойственная оценка является точной [20, 21].

Во многомерном случае основной результат формулируется более сложно.

Пусть $y = \{y_1, \dots, y_n\}$ – n -мерный вектор. Введем обозначение для одночлена $y_1^{\alpha_1} y_2^{\alpha_2} \dots y_n^{\alpha_n} = R[\alpha](y)$, где $\alpha = \{\alpha_1, \dots, \alpha_n\}$ – целочисленный n -мерный вектор, $\alpha_i \geq 0$, $i = 1, \dots, n$, $R[\alpha]$ – соответствующая одночленная функция.

Любой вещественный многочлен $P(y_1, \dots, y_n)$ от n переменных можно представить в виде

$$P(y) = P(y_1, \dots, y_n) = \sum_{\alpha \in A(P)} c_\alpha R[\alpha](y),$$

где $A(P)$ – множество одночленов $R[\alpha]y$, входящих в полином $P(y)$ с ненулевыми коэффициентами c_α .

Пусть M_k – множество многочленов тотальной степени, не превышающей k . Для четного $k = 2s$ каждый одночлен представим в виде произведения двух одночленов тотальной степени, не превышающей s .

Пусть также $M[k] = \{\alpha : \sum_{i=1}^n \alpha_i \leq k\}$; α_i – целые неотрицательные числа.

Тогда любой вектор $\alpha \in M[2s]$ может быть представлен, вообще говоря, неоднозначно, в виде

$$\alpha = \alpha^{(1)} + \alpha^{(2)}, \quad \alpha^{(1)}, \alpha^{(2)} \in M[s]. \quad (33)$$

Представления вида (33) назовем допустимыми. Для каждого $\alpha \in M[2s]$ выделим одно из допустимых представлений и назовем его каноническим:

$$\alpha = \alpha^{(1)}(\alpha) + (\alpha - \alpha^{(1)}(\alpha)),$$

где $\alpha^{(1)}(\alpha)$ – первая часть канонического разложения вектора α .

Рассмотрим систему тождественных соотношений вида:

$$\Delta(\alpha, \alpha^{(1)}) = R[\alpha^{(1)}(\alpha)] * R[\alpha - \alpha^{(1)}(\alpha)] - R[\alpha^{(1)}] * R[\alpha - \alpha^{(1)}] = 0, \quad (34)$$

которые выписаны для всех $\alpha \in M[2s]$ и для всех $\alpha^{(1)} \leq \alpha$, $\alpha^{(1)} \in M[s]$, $\alpha^{(1)} \neq \alpha^{(1)}(\alpha)$.

Назовем допустимым квадратичным представлением полинома $P(y) = \sum_{\alpha \in A(P)} c_\alpha R[\alpha](y)$ тотальной степени, не превышающей $2s$, квадратичную функцию вида

$$\overline{K}_P[R] = \sum_{\alpha \in A(P)} \left(\sum_i c_\alpha^{(i)} R[\overline{\alpha}^{(i)}] * R[\alpha - \overline{\alpha}^{(i)}] \right),$$

где $R[\overline{\alpha}^{(i)}] * R[\alpha - \overline{\alpha}^{(i)}]$ – допустимые разложения одночлена $R[\alpha]$, $\forall \alpha \in A(P)$, $\sum_{i \in I(\alpha)} c_\alpha^{(i)} = c_\alpha$, $I(\alpha)$ – множество допустимых разложений $R(\alpha)$.

Легко доказывается следующая лемма.

Лемма. Пусть $\overline{K}_P(R)$ – произвольное квадратичное представление полинома $P(y)$ тотальной степени, не превышающей $2s$. Тогда существует вектор $\overline{u} = \{\overline{u}(\alpha, \alpha^{(1)})\}$ множителей $\overline{u}(\alpha, \alpha^{(1)})$ такой, что

$$\overline{K}_P(R) = K_P^{(0)}(R) + \sum_{(\alpha, \alpha^{(1)})} \overline{u}(\alpha, \alpha^{(1)}) \Delta(\alpha, \alpha^{(1)}), \quad (35)$$

где $K_P^{(0)}(R)$ – каноническое квадратичное представление P .

Пусть $\inf_y P(y) = P^* > -\infty$. Тогда $\overline{P}(y) = P(y) - P^* \geq 0$ для всех y . Рассмотрим квадратичную задачу:

$$\text{найти } \inf K_P^{(0)}(R) \text{ при ограничениях вида} \quad (34). \quad (36)$$

Выражение вида (35) можно рассматривать как функцию Лагранжа задачи (36), при этом вектор \overline{u} играет роль вектора множителей Лагранжа. Отметим, что большая часть ограничений (34) является функционально избыточной.

Рассмотрим функцию Лагранжа $L(R, u)$ задачи (36). Пусть $\psi(u) = \inf_R L(R, u)$. Если $\text{dom } \psi \neq \emptyset$, то $\psi(u)$ – собственная вогнутая функция от u . Рассмотрим $\psi^* = \sup_u \psi(u)$.

Доказана следующая основная теорема.

Теорема. Оценка $\psi^* = P^*$ тогда и только тогда, когда неотрицательный полином $\overline{P}(y)$ может быть представлен в виде суммы квадратов вещественных полиномов тотальной степени, не превышающей s [20, 21, 8].

В работе [22] для нахождения ψ^* применена модификация г-алгоритма и проведены вычислительные эксперименты, подтверждающие эффективность такой методики.

Заметим, что проблемой разложения неотрицательных полиномов четной степени в сумму квадратов других вещественных полиномов занимался еще Д. Гильберт [23]. Близкая проблема представления неотрицательных рациональных функций в виде суммы квадратов рациональных функций (17-я проблема Гильберта) была положительно решена в работе Э. Артина [24]. Наш подход дает возможность получить разложение $\overline{P}(y)$ в сумму квадратов, если оно существует. Автору неизвестны другие алгоритмы, осуществляющие указанное разложение.

В заключение отметим, что проблема получения улучшенных оценок в невыпуклых задачах квадратичного типа находится в центре внимания специалистов в области дискретного и нелинейного программирования. Определение таких оценок связано с быстро развивающейся областью матричной оптимизации - полуопределенным программированием (semidefinite programming). Для численного решения соответствующих задач в большинстве работ предлагается использовать методы внутренних точек [25]–[29]. В отличие от такого подхода мы рассматриваем вопросы получения оценок для невыпуклых квадратичных задач в рамках теории двойственных (лагранжевых) оценок. Вычисление этих оценок приводит к необходимости решения негладких выпуклых оптимизационных матричных задач специального вида с ограничениями на величину максимальных (минимальных) собственных чисел. Соответствующие задачи рекомендуем решать с помощью эффективных алгоритмов субградиентного типа с преобразованием пространства аргументов. Как показали численные эксперименты, предложенный нами подход вполне конкурентноспособен по сравнению с другими методами получения оценок в квадратичных задачах. Для улучшения двойственных оценок можно успешно применять модели с расширенным множеством ограничений, часть которых является функционально избыточными.

Литература

1. SHOR N.Z., STETSENKO S.I. Quadratic boolean problems and Lovasz bounds. – 12th IFIP Conf. on System modelling and optimization.: Abstr. (Budapest, September 2–6, 1985). – Budapest, 1995. – P. 302–304.
2. ШОР Н.З. Задачи квадратичного типа // Изв. АН СССР. Техн. кибернетика. – 1987. – № 1. – С. 128–139.

3. ШОР Н.З., СТЕЦЕНКО С.И. Квадратичные экстремальные задачи и недифференцируемая оптимизация. – Киев: Наук. думка, 1989. – 208 с.
4. GRÖTSCHEL M., LOVASZ L., SCHRIJVER A. Geometric Algorithms and Combinatorial Optimization. – Springer-Verlag, Berlin. – 1988. – 362 p.
5. LOVASZ L. On the Shannon capacity of a graph // IEEE Trans. Inform. Theory. – 1979. – **25**. – № 1. – P. 1–7.
6. SCHRIJVER A. A comparison of the Delsart and Lovász Bounds // Ibid. – 1979. – **25**. – № 4. – P. 425–429.
7. GRÖTSCHEL M., LOVASZ L., SCHRIJVER A. The ellipsoid methods and its consequences in combinatorial optimization // Combinatorica. – 1981. – № 1. – P. 169–197.
8. SHOR N.Z. Nondifferentiable optimization and polynomial problems. – Kluwer Academic Publishers, 1998. – 394 p.
9. ШОР Н.З., БЕРЕЗОВСКИЙ О.А. Новые алгоритмы решения задачи о максимальном разрезе графа // Кибернетика и систем. анализ. – 1995. – № 2. – С. 100–106.
10. VARAHONA F., MAHJOUR A.R. On the cut polytope // Mathematical Programming. – 1986. – **36**. – P. 157–173.
11. SEYMOUR P.D. On odd cuts and plane multicommodity flows // Proceedings of the London Mathematical Society. – 1981. – **42**. – № 3. – P. 178–192.
12. ОРЛОВА Г.И., ДОРФМАН Ю.Г. Нахождение максимального разреза в графе // Известия АН СССР, Техническая кибернетика. – 1972. – № 3. – С. 155–159.
13. HADLOCK F. Finding a maximum cut of a planar graph in polynomial time // SIAM Journal on Computing. – 1975. – **4**. – P. 221–225.
14. VARAHONA F. The max-cut problem in graphs not contractible to K_5 // Operations Research Letters. – 1983. – **2**. – P. 107–111.
15. GRÖTSCHEL M., PULLEYBLANC W.R. Weakly bipartite graphs and the max-cut problem // Operations Research letter. – 1981. – **1**. – P. 23–27.

16. FONLUPT Z., MAHJOUR A.R., UHRY J.-P. Composition of graphs and the bipartite subgraph polytope. – Research report No. 459. – Laboratoire ARTEMIS (IMAG). Universite de Grenoble. – Grenoble. – 1984.
17. BARAHONA F., GRÖTSCHEL M., MAHJOUR A.R. Facets of bipartite subgraph polytope. – Mathematics of Operations Research. – 1985. – **10**. – P. 340–358.
18. BARAHONA F., GRÖTSCHEL M., JUNGER M., REINELT G. An application of combinatorial optimization to statistical physics and circuit layout design. – Preprint No. 95. – Augsburg: Institut für Mathematik, Universität Augsburg. – 1986.
19. GOEMANS M.X., WILLIAMSON D.P. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming // Journal of the ACM. – 1995. – **6**. – P. 1115–1145.
20. ШОР Н.З. Об одном классе оценок глобального минимума полиномиальных функций // Кибернетика. – 1987. – № 6. – С. 9–11.
21. ШОР Н.З. Об одном подходе к получению глобальных экстремумов в полиномиальных задачах математического программирования // Кибернетика. – 1987. – № 5. – С. 102–106.
22. ШОР Н.З., СТЕЦЮК П.И. Использование модификации г-алгоритма для нахождения глобального минимума полиномиальных функций // Кибернетика и систем. анализ. – 1987. – № 4. – С. 28–49.
23. HILBERT D. Über die Darstellung definiter Formen als Summen von Formen quadraten // Math. Ann. – Leipzig, 1888. – Bd. 22. – P. 342–350.
24. ARTIN E. Über die Zerlegung definiter Functionen in Quadrate // Hamburg Abh. – 1927. – № 5. – P. 100–115.
25. LOVASZ L., SCHRIJVER A. Cones of matrices and set functions and 0-1 optimization // SIAM Journal on Optimization. – 1991. – **1**. – P. 166–190.
26. ALIZADEH F. Interior point methods in semidefinite programming with applications to combinatorial optimization // SIAM Journal on Optimization. – 1995. – **5**. – P. 13–51.

27. NESTEROV Y., NEMIROVSKY A. Interior point polynomial algorithms in convex programming. – SIAM Studies in Applied Mathematics, **13**. – 1993. – 405 p.
28. NESTEROV Y. Semidefinite relaxation and nonconvex quadratic optimization // Optimization methods and Software, Special Issue. – Febr., 1998. – P. 141–160.
29. Quist A.J., Klerk E., Roos C., Terlaky T. Copositive relaxation for general quadratic programming // Optimization methods and Software, Special Issue. – Febr., 1998. – P. 185–208.

Lagrangian bounds in multiextremal polynomial and discrete optimization problems

Naum Z. Shor and Petro I. Stetsyuk

Journal of Global Optimization. – 2002. – 23. – P. 1–41.

Many polynomial and discrete optimization problems can be reduced to multiextremal quadratic type models of nonlinear programming. For solving these problems one may use Lagrangian bounds in combination with branch and bounds techniques. The Lagrangian bounds may be improved for some important examples by adding in a model so-called superfluous quadratic constraints which modify Lagrangian bounds.

Problems of finding Lagrangian bounds as a rule can be reduced to minimization of nonsmooth convex functions and may be successively solved by modern methods of nondifferentiable optimization. This approach is illustrated by examples of solving polynomial-type problems and some discrete optimization problems on graphs.

1. Lagrangian bounds, nondifferentiable optimization and nonsmooth matrix functions

Consider the problem of nonlinear programming in general form: to find

$$f^* = \inf_{x \in X \subseteq E^n} f_0(x), \quad \text{subject to } f_i(x) = 0, \quad i = 1, \dots, m, \quad (1)$$

where E^n is n -dimensional Euclidean space, X is a closed set in this space, f_0, f_1, \dots, f_m are continuous functions, defined on E^n . We set $f^* = +\infty$, if the problem (1) has no feasible solution. Let us form the usual Lagrange function:

$$L(x, u) = f_0(x) + \sum_{i=1}^m u_i f_i(x),$$

where $u = (u_1, \dots, u_m)$ is a vector of Lagrange multipliers.

For each $\bar{u} \in R^m$ we obtain the local problem: to find

$$\psi(\bar{u}) = \inf_{x \in X} L(x, \bar{u}). \quad (2)$$

Function $\psi(u)$ is a concave function with respect to u as a result of minimization of the family of linear in u functions $L_x(u) = L(x, u)$.

Let $\psi(u)$ has a nonempty domain of full dimension m and $x(\bar{u})$ is a solution of local problem (2). It is easy to verify that $\psi(\bar{u}) \leq f^*$ for an arbitrary $\bar{u} \in \text{dom}\psi$. The supergradient of ψ in the point \bar{u} can be calculated by formula:

$$g_\psi(\bar{u}) = \{f_i(x(\bar{u}))\}_{i=1}^m \in R^m. \quad (3)$$

If the set of vectors generated by (3) is not single, then \bar{u} is the point of nondifferentiability of function ψ .

We try to find the best lower bound for f^* in this class of Lagrangian estimates and obtain the coordinating problem: to find

$$\psi^* = \sup_{u \in R^m} \psi(u).$$

Note that the problem of finding the best Lagrangian bounds ψ^* is one of the main sources of generating the nonsmooth optimization models.

Many combinatorial optimization problems can be formulated as Boolean LP problems, and the corresponding dual bounds may be obtained by LP relaxations of such models. But in some cases the nonlinear quadratic-type formulation of a combinatorial problem is more convenient and may give more exact dual bounds. In these cases, as a rule, the problem of obtaining dual bounds may be reduced to the convex programming problems with nonsmooth matrix function (or to the equivalent problems of semidefinite programming).

Now there exist many methods of nondifferentiable optimization, for example, simple subgradient method, ε -subgradient methods, methods with space transformation. One of the most practically effective modern methods is the algorithm with space dilation in the direction of difference of two successive subgradients (so-called r -algorithms).

This method was proposed by N.Z. Shor in 1970 [32] for acceleration of convergence of subgradient methods. A family of r -algorithms contains different realizations of subgradient-type methods with space dilations in the direction of difference of two successive subgradients. Below we give general scheme of r -algorithm.

Denote by $R_\alpha(\xi)$ the linear operator of space dilation in the direction ξ , $\|\xi\| = 1$, with coefficient α , $\alpha \geq 0$, specifying for each $x \in E^n$ the vector

$y \in E^n$ due to the formula:

$$y = R_\alpha(\xi)y = x + (\alpha - 1)(x, \xi)\xi.$$

Let $f(x)$ be a minimized convex function denoted on E^n , $x_0 \in E^n$ be a given point. Denote by $g_f(\bar{x})$ a subgradient of function f in the point \bar{x} . The general scheme of r -algorithm is following:

The first step:

$$x_1 = x_0 - h_0 \frac{g_f(x_0)}{\|g_f(x_0)\|},$$

where h_0 is the step multiplier such, that $(g_f(x_0), g_f(x_1)) \leq 0$. Fix $g_f(x_0)$, $g_f(x_1)$, $B_0 = I_n$ ($n \times n$ identity matrix). After k steps we have x_1, x_2, \dots, x_k and fix x_k , $g_f(x_{k-1})$ and $n \times n$ matrix B_k .

$(k + 1)$ -st step.

Calculate:

- (a) $g_f(x_k)$;
- (b) $r_k = g_f(x_k) - g_f(x_{k-1})$;
- (c) $\xi_k = \frac{B_{k-1}r_k}{\|B_{k-1}r_{k-1}\|}$;
- (d) $B_k = B_{k-1}R_{\beta_k}(\xi_k)$, $0 < \beta_k < 1$;
- (e) $x_{k+1} = x_k - h_k B_k \frac{B_k^T g_f(x_k)}{\|B_k^T g_f(x_k)\|}$.

If the stopping criteria is not fulfilled, fix x_{k+1} , $g_f(x_{k-1})$, B_k and go to the next step.

Comments. After k steps of r -algorithm let A_{k-1} be the resulting matrix of space transformation: $y = A_{k-1}x$, or $x = B_{k-1}y$, where $B_{k-1} = A_{k-1}^{-1}$, and $\varphi_k(y) = f(B_{k-1}y)$. Since $g_{\varphi_k}(y) = B_{k-1}^T g_f(x)$, r_k is the difference of two subgradients of function φ_k taken in the points $y_k = A_{k-1}x_k$ and $\tilde{y}_{k-1} = A_{k-1}x_{k-1}$. So ξ_k is the normalized direction of the difference of two successive subgradients of transformed function $\varphi_k(y)$. In this direction we make a current dilation of transformed space and obtain resulting matrix $A_k = R_{\alpha_k}(\xi_k)A_{k-1}$. The inverse matrix $B_k = B_{k-1}R_{\beta_k}(\xi_k)$, $\beta_k = 1/\alpha_k$. Consider $\varphi_{k+1}(y) = f(B_k y)$, and use subgradient step for function $\varphi_{k+1}(y)$ from $\tilde{y}_k = B_k x_k$:

$$y_{k+1} = \tilde{y}_k - h_k \frac{B_k^T g_f(x_k)}{\|B_k^T g_f(x_k)\|}.$$

In original space the point $x_{k+1} = B_k y_{k+1}$ corresponds to the point x_{k+1} , and we obtain

$$x_{k+1} = x_k - h_k B_k \frac{B_k^T g_f(x_k)}{\|B_k^T g_f(x_k)\|}.$$

Thus we calculate x_{k+1} , $g_f(x_k)$, B_k , and are ready to make the next iteration.

The family of r -algorithms has two sequences of parameters: $\{\beta_k\}_{k=1}^{\infty}$ and $\{h_k\}_{k=1}^{\infty}$. Naturally, we must think about rational choice of these sequences to obtain "good" convergence to the optimal point of $f(x)$.

For minimization of nonsmooth convex functions defined on E^n we recommend to use the following specifications of r -algorithm:

The space dilation coefficients α_k equal to α , where $\alpha \in [2, 4]$. To determine a step multiplier h_k we use adaptive technique of step length regulation (see [35]) determined by parameters: $h_0^{(0)}$ (initial step-length), integer number $\bar{m} > 1$, and coefficients $q_1 < 1$ and $q_2 > 1$ for decreasing (increasing) of step-multiplier. After k iterations of r -algorithm we obtain step constant h_k^0 . On $(k+1)$ -st iteration we choose the direction of descent due to r -algorithm and move in this direction with a step multiplier h_k^0 until the condition of stopping the search along the direction is fulfilled or the number of steps would be equal to \bar{m} . In the last case we continue descent along the same direction with a new step constant $h_k^1 = q_2 h_k^0$. If after \bar{m} steps the condition of interrupting of search direction is not fulfilled, we set $h_k^2 = q_2 h_k^1$ and so on.

We suppose that $\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$, so after finite number of steps the stopping condition for directional search will be fulfilled.

The details of such way of regulating step-multiplier one may find in Section 3 of [33].

The results of testing of r -algorithms show that if the errors of rounding are not essential, the objective function values as a rule may be majored by a geometrical progression of the form $Cq^{\frac{k}{n}}$, where k is the number of current step and $q = \frac{1}{2}$. So, as a rule the convergence of r -algorithm approximately $2n$ times faster than of well known ellipsoid methods.

The most typical examples of nonsmooth functions are maximal and minimal eigenvalues of symmetric matrices and sums of k largest eigenvalues (for example, see [23]).

Let Σ_n be the class of real $n \times n$ symmetric matrices. Any matrix $A \in \Sigma_n$ has n real eigenvalues (with account of their multiplicity) and a pair of eigenvectors associated with two different eigenvalues are orthogonal. Let

$$A \in \Sigma_n, A = \{a_{ij}\}_{i,j=1}^n,$$

$$\lambda_1(A) \geq \lambda_2(A) \geq \dots \geq \lambda_n(A)$$

be the eigenvalues of A , ordered in nonincreasingly.

A symmetric real matrix is called positive definite (semidefinite) if $\lambda_n(A) > 0$ ($\lambda_n(A) \geq 0$). We shall write $A \succ 0$ ($A \succeq 0$), if $A \in \Sigma_n$ and A is positive definite (semidefinite).

The Rayleigh–Ritz formula is known for maximal eigenvalue $\lambda_1(A)$:

$$\lambda_1(A) = \max_{\|y\|=1} (Ay, y) = \max_{\|y\|=1} \sum_{i,j=1}^n a_{ij} y_i y_j, \quad (4)$$

where $y = (y_1, \dots, y_n) \in E^n$. $\lambda_1(A)$ is a convex function defined on Σ_n since formula (4) gives representation of this function as a maximum of a family of linear functions in entries $\{a_{ij}\}_{i,j=1}^n$.

Denote by $Y^*(A)$ a set of normalized vectors y , which get maximum in (4), i.e. $\lambda_1(A) = (Ay^*, y^*)$ for all $y^* \in Y^*(A)$. From (4), one may obtain the subgradient set $G_{\lambda_1}(\bar{A})$ of function $\lambda_1(\cdot)$ in the point \bar{A} :

$$G_{\lambda_1}(\bar{A}) = \text{conv} \left\{ \bigcup_{y \in Y^*(\bar{A})} yy^T \right\}.$$

The calculation of subgradient $g_{\lambda_1}(\bar{A}) \in G_{\lambda_1}(\bar{A})$ may be reduced to finding an arbitrary $y^*(\bar{A}) \in Y^*(\bar{A})$ and applying the next formula:

$$g_{\lambda_1}(\bar{A}) = \{y^*(\bar{A})[y^*(\bar{A})]^T\}.$$

(Note that $yy^T, y \in E^n$, is the symmetric matrix of rank 1 with entries $\{y_i y_j\}_{i,j=1}^n$). If $\lambda_1(\bar{A})$ has multiplicity 1, then $g_{\lambda_1}(\bar{A})$ is unique and function $\lambda_1(A)$ is differentiable at the point \bar{A} . When multiplicity of $\lambda_1(\bar{A})$ is more than 1, the function $\lambda_1(A)$ is nondifferentiable in \bar{A} .

Now we introduce a very interesting class of convex matrix functions defined on symmetric matrices $A \in \Sigma_n$, namely, the sums of the k largest eigenvalues:

$$S_{n,k}(A) = \sum_{r=1}^k \lambda_r(A), \quad 1 \leq k \leq n.$$

Famous mathematician Fan Ky gave in 1949 variational description of $S_{n,k}(A)$ that is a far going generalization of Rayleigh–Ritz formula (4). Let M_n^k be the class of rectangular $n \times k$ matrices Y , the columns $y_i, i = 1, \dots, k, k \leq n$ form an orthonormal system of n -dimensional vectors, i.e. $Y^T Y = I_k$ (I_k is $k \times k$ identity matrix).

Theorem 1 (Fan Ky [17]).

$$S_{n,k}(A) = \max_{Y \in M_n^k} \{tr(AYY^T)\}, \forall A \in \Sigma_n. \quad (5)$$

The maximum in formula (5) is reached at orthonormal system of eigenvectors y_1^*, \dots, y_k^* corresponding to eigenvalues $\lambda_1(A), \dots, \lambda_k(A)$. Indeed,

$$\begin{aligned} Ay_i^* &= \lambda_i(A)y_i^*, \quad i = 1, \dots, k, \\ tr(AYY^T) &= (A, YY^T) = \sum_{i=1}^k (\lambda_i(A)y_i^*, y_i^*) = \\ &= \sum_{i=1}^k \lambda_i(A) = S_{n,k}(A). \end{aligned}$$

When $k = 1$, formula (5) is reduced to the expression (4).

Consider the class of $n \times n$ matrices

$$C_n^k = \text{conv}\{YY^T : Y \in M_n^k\}.$$

Theorem 2. C_n^k coincides with the class of all positive semidefinite matrices C with $\lambda_1(C) \leq 1$ and with trace, equal to k [17].

Due to this theorem, we obtain a new variational formula for $S_{n,k}(A)$:

$$S_{n,k}(A) = \max_{C \in C_n^k} (A, C), \quad (6)$$

where $(A, C) = \sum_{i,j=1}^n a_{ij}c_{ij}$ ($A = \{a_{ij}\}_{i,j=1}^n$, $C = \{c_{ij}\}_{i,j=1}^n$).

Formulas (5) and (6) give us the representation of the function $S_{n,k}(A)$ as a pointwise maximum function on infinite family of linear (in matrix variable $A \in \Sigma_n$) functions. So, $S_{n,k}(A)$ is a convex function for any $n, k \leq n$. The structure of subgradient set $G_{S_{n,k}}(A)$ is determined by (6):

$$G_{S_{n,k}}(A) = \{C^* \in C_n^k : S_{n,k}(A) = (A, C^*)\}$$

For almost all A , C^* is unique and gives us the gradient $g_{S_{n,k}}(A)$ of $S_{n,k}$ in the point A .

The most strict way for obtaining a subgradient $g_{S_{n,k}}(A)$ is the following:

- (i) solve the eigenvalue problem for matrix A and find the eigenvalues $\lambda_1(A), \dots, \lambda_k(A)$ and the corresponding orthonormal system of eigenvectors $Y_j(A) = \{y_i^j\}_{i=1}^n$, $j = 1, \dots, k$;
- (ii) construct the $n \times k$ matrix $Y = \{y_i^j\}_{i=1, \dots, n}^{j=1, \dots, k}$;
- (iii) set $\bar{g}_{S_{n,k}}(A) = YY^T$.

In general, when not all k largest eigenvalues of A have multiplicity 1, the subgradient $g_{S_{n,k}}$ is not unique, because the system of eigenvectors $\{Y_i(A)\}_{i=1}^k$ is determined nonuniquely in this case. But if one is interested in calculating any subgradient from $G_{S_{n,k}}(A)$ one can use the procedure described above for arbitrary orthonormal system $Y(A)$ of eigenvectors associated with the k largest eigenvalues $\lambda_1(A) \geq \lambda_2(A) \geq \dots \geq \lambda_k(A)$.

Let A be a diagonal matrix and $a_{11} \geq a_{22} \geq \dots \geq a_{kk} \geq \dots \geq a_{nn}$, $1 \leq k \leq n$. Consider two cases:

(I) $a_{kk} > a_{(k+1)(k+1)}$. In this case in formula:

$$S_{n,k}(A) = \max_{C \in C_n^k} (A, C) = (A, C^*).$$

C^* is determined uniquely (C^* is a diagonal matrix with $c_{ii} = 1$ for $i \leq k$ and $c_{ii} = 0$ for $i > k$). The subgradient of the function $S_{n,k}$ in the point A , $g_{S_{n,k}}(A)$ is equal to C^* , and the function $S_{n,k}$ is differentiable at A ;

(II) $a_{kk} = a_{(k+1)(k+1)}$. In this case the subgradient set $G_{S_{n,k}}(A)$ contains more than one extremal point. For example, if $a_{ii} = a_{kk}$ for all i , $k - s \leq i \leq k + p$ ($s \geq 0$, $p \geq 1$), then an arbitrary diagonal matrix A with properties:

(1) $a_{ii} = 1$ for $i < k - s$;

(2) $a_{ii} = 0$ for $i > k + p$;

(3) the set of values $\{a_{ii}\}$, $k - s \leq i \leq k + p$ contains exactly s ones and p zeroes;

is an extreme point of $G_{S_{n,k}}(A)$. So, in the case (II) the function $S_{n,k}$ is nondifferentiable at A .

In general, if for a symmetric matrix A , $\lambda_k(A) > \lambda_{k+1}(A)$ then $S_{n,k}$ is differentiable at A ; otherwise (i.e. $\lambda_k(A) = \lambda_{k+1}(A)$) the function $S_{n,k}$ is nondifferentiable at A .

In many applications we meet with a weighted sum of k largest eigenvalues:

$$S_{n,k}(A, w) = \sum_{i=1}^k w_i \lambda_i(A), \text{ where } w = (w_1, \dots, w_k) \geq 0.$$

Lemma 1. *If $w_1 \geq w_2 \geq \dots \geq w_k$, then $S_{n,k}(A, w)$ is a convex function defined on Σ_n .*

The weighted sum of the largest eigenvalues $S_{n,k}(A, w)$ can be represented also by the variational formula similar to formula (5):

$$S_{n,k}(A, w) = \max_{Z \in M_n^k(w)} \{\text{tr}(AZZ^T)\},$$

where $M_n^k(w)$ is the class of rectangular $n \times k$ matrices Z , the columns Z_i forming an orthogonal system of n -dimensional vectors, and $\|Z_i\|^2 = w_i$, $i = 1, \dots, k$, $w_1 \geq w_2 \geq \dots \geq w_k$.

Due to expression (5) of Theorem 1 the subgradient set for the function $S_{n,k}(\cdot)$ at point X is given by the following expression:

$$G_{S_{n,k}}(X) = \text{conv}\left(\sum_{i=1}^k y_i y_i^T\right),$$

where y_i , $i = 1, \dots, k$, form an arbitrary orthonormal system of vectors, associated with eigenvalues $\lambda_1(X), \dots, \lambda_k(X)$. If multiplicity of all eigenvalues $\lambda_1(X), \dots, \lambda_k(X)$ is equal to one, then the matrix $\sum_{i=1}^k y_i y_i^T$ is determined uniquely and coincides with gradient $S_{n,k}(X)$ at point X .

Functions $\lambda_m(X)$, $1 < m < n$ are quasi-differentiable functions (in the sense of Demjanov and Rubinov [11]). They may be considered as a difference of two convex functions

$$\lambda_m(X) = S_{n,m-1}(X) - S_{n,m}(X).$$

If $w = \{w_1, \dots, w_n\} \geq 0$ and $w_k \geq w_{k+1}$ for $k = 1, \dots, n-1$, then subgradient set of convex function $S_n^w(X) = \sum_{i=1}^n w_i \lambda_i(X)$ can be represented by the following expression:

$$G_{S_n^w}(X) = \text{conv}\left\{\sum_{i=1}^n w_i y_i y_i^T\right\},$$

where $\{y_i\}_{i=1}^n$ is any orthonormal system of eigenvectors of matrix X (each y_i is associated with $\lambda_i(X)$). If all eigenvectors $y_i(X)$ with $w_i > 0$ have multiplicity 1, then S_n^w is differentiable at X .

2. Quadratic-type minimization problems, Lagrangian lower bounds and superfluous constraints

Consider now the problems of finding Lagrangian lower bounds for quadratic-type optimization models:

$$\text{Find } Q^* = \inf_{x \in E^n} Q_0(x) \quad \text{subject to} \quad Q_i(x) = 0, \quad i = 1, \dots, m, \quad (7)$$

where $Q_\nu(x)$, $\nu = 0, \dots, m$ are quadratic or linear functions, determined on n -dimensional Euclidean space E^n . If the problem (7) has no feasible solution, we set $Q^* = +\infty$. Using usual Lagrange function $L(x, u)$, $u = \{u_1, \dots, u_m\}$, one can obtain Lagrangian lower bounds for such problems by finding $\psi^* = \sup \psi(u)$, where $\psi(u) = \inf_{x \in E^n} L(x, u)$. If $\text{dom } \psi$ is nonempty, then $\psi(u)$ is a proper concave function. In opposite case we obtain a trivial bound $\psi^* = -\infty$.

Consider the problem of finding dual (Lagrangian) estimates for quadratic-type problems of the form (7) in more detail. Let quadratic functions $Q_\nu(x)$, $\nu = 0, \dots, m$, have the following description:

$$Q_\nu(x) = (K_\nu x, x) + (c_\nu, x) + d_\nu,$$

where K_ν are symmetric quadratic $n \times n$ matrices, c_ν are n -dimensional vectors, d_ν are numbers. So, $(K_\nu x, x)$ is a quadratic part of $Q_\nu(x)$ and (c_ν, x) is a linear part of $Q_\nu(x)$, $\nu = 0, \dots, m$.

Usual Lagrange function $L(x, u)$ can be represented as

$$L(x, u) = Q_0(x) + \sum_{i=1}^m u_i Q_i(x) = (K(u)x, x) + (c(u), x) + d(u),$$

where

$$\begin{aligned} K(u) &= K_0 + \sum_{i=1}^m u_i K_i; \\ c(u) &= c_0 + \sum_{i=1}^m u_i c_i; \\ d(u) &= d_0 + \sum_{i=1}^m u_i d_i; \end{aligned}$$

(here $u = \{u_1, \dots, u_m\}$ is m -dimensional vector of Lagrange multipliers).

Consider $\psi(u) = \inf_x L(x, u)$. If $K(u)$ is a positive definite matrix then $\psi(u) = L(x(u), u)$, where $x(u)$ is a solution of linear system of equations

$$2K(u)x + c(u) = 0,$$

i.e.

$$x = -\frac{1}{2} \left(K(u) \right)^{-1} c(u).$$

If the minimal eigenvalue of $K(u)$ $\lambda_n \left(K(u) \right) < 0$, then $\psi(u) = -\infty$.

In the case $\lambda_n \left(K(u) \right) = 0$ matrix $K(u)$ is positive semidefinite but singular. Let s_1, \dots, s_n be the orthonormal basis in E^n , corresponding to eigenvectors ordered in decreasing order of their eigenvalues (with taking in account their multiplicity). Then $c(u)$ can be represented in the form:

$$c(u) = \sum_{i=1}^{n-r} \alpha_i(u) s_i + \sum_{i=n-r}^r \alpha_i(u) s_i;$$

where r is the multiplicity of the minimal eigenvalue equal to zero. If all $\alpha_k(u) = 0$ for $n-r+1 \leq k \leq n$, then $x(u)$ exists, i.e. $u \in \text{dom } \psi$. Otherwise $\psi(u) = -\infty$.

Let $\text{dom } \psi$ be non-empty. Then ψ is a proper concave function as a result of minimization with respect to x of functions

$$\psi_u(x) = Q_0(x) + \sum_{i=1}^m u_i Q_i(x),$$

which are linear by u for all x .

For each $u \in \text{dom } \psi$ $\psi(u) \leq Q^*$ (Q^* is an optimal value of initial problem (7)).

Let $\bar{u} \in E^m$ and $K(\bar{u}) \succ 0$. Then $Q(\bar{u})$ is an interior point of $\text{dom } \psi$. Denote by Ω^+ the set

$$\{u \in E^m / K(u) \succ 0\}.$$

Boundary $\bar{\Omega}$ of Ω^+ consists of u for which $K(u)$ has the minimal eigenvalue equal to zero. It is easy to prove that all $u \in \bar{\Omega}$ are limit points of the set Ω^+ . So if Ω^+ is nonempty then $\text{dom } \psi$ is closed in E^m .

Let $\psi^* = \sup_{u \in E^m} \psi(u)$ and there exists a point u^* such that $\psi^* = \psi(u^*)$. If $u^* \in \Omega^+$ then $g_\psi(u^*) = Q_i(x^*) = 0$, $i = 1, \dots, m$. In this case $\psi^* = Q^*$. Otherwise, when $x^* \in \bar{\Omega}$, it may be a positive defect of duality:

$$\Delta = Q^* - \psi^* > 0.$$

For many interesting problems we may improve the dual bounds by using the so-called functionally superfluous constraints in the form of quadratic equations (inequalities) which do not change the optimal value of initial polynomial problem but lead to modification of Lagrange function of corresponding quadratic problem. This modification may give substantial increasing of new Lagrangian bound for a modified quadratic type problem in comparison with ψ^* for the old one.

Note that if we add to initial problem (7) new quadratic superfluous constraints $Q_{m+1}(x) = 0, \dots, Q_{m+r}(x) = 0, r \geq 1$ and form a longer vector of Lagrange multipliers $U = \left\{ \{u\}, u_{m+1}, \dots, u_{m+r} \right\}$ then for the problem:

$$Q^* = \inf_{x \in E^n} Q_0(x) \quad \text{subject to} \quad Q_i(x) = 0, \quad i = 1, \dots, m, m+1, \dots, m+r$$

the corresponding Lagrange function will be

$$L_1(x, U) = Q_0(x) + \sum_{i=1}^{m+r} u_i Q_i(x) = L(x, u) + \sum_{i=m+1}^{m+r} u_i Q_i(x).$$

So

$$L(x, u) = L_1 \left(x, (\{u\}, 0, \dots, 0) \right),$$

$$\psi_1(U) = \inf_x L_1(x, U) \geq \inf L(x, u) = \psi(u),$$

and

$$\psi_1^* = \sup \psi_1(U) \geq \psi^*.$$

We shall demonstrate the possible improving of dual bounds by introducing superfluous constraints on a simple example.

Example. Let $P_6(x_1)$ be a sixth-degree polynomial of one variable x_1 :

$$p_6(x_1) = x_1^6 + a_5 x_1^5 + a_4 x_1^4 + a_3 x_1^3 + a_2 x_1^2 + a_1 x_1 + a_0.$$

The problem is to find value p^* of (global) minimum of $p_6(x_1)$. One may transform this problem in a quadratic-type problem by introducing new variables $x_2 = x_1^2$; $x_3 = x_1 x_2$.

Consider a vector of variables $x = \{x_1, x_2, x_3\}$ and obtain the equivalent problem:

to minimize

$$Q_0(x) = x_3^2 + a_5 x_2 x_3 + a_4 x_3 x_1 + a_3 x_3 + a_2 x_2 + a_1 x_1 + a_0 \quad (8)$$

subject to constraints:

$$Q_1(x) = x_1^2 - x_2 = 0, \quad (9)$$

$$Q_2(x) = x_1 x_2 - x_3 = 0. \quad (10)$$

The Lagrange function $L(x, u)$ of this problem has the following form ($u = \{u_1, u_2\}$ is a vector of Lagrange multipliers):

$$L(x, u) = Q_0(x) + u_1 Q_1(x) + u_2 Q_2(x).$$

Consider the matrix $K(u)$ which defines the quadratic part of $L(x, u)$.

$$K(u) = \begin{pmatrix} u_1 & \frac{u_2}{2} & \frac{a_4}{2} \\ \frac{u_2}{2} & 0 & \frac{a_5}{2} \\ \frac{a_4}{2} & \frac{a_5}{2} & 1 \end{pmatrix}.$$

One may see that if $a_5 \neq 0$, for all $u = \{u_1, u_2\}$, $K(u)$ cannot be positive semidefinite, and we obtain a trivial lower bound $\psi^* = -\infty$. So, it seems an attempt to use dual quadratic bounds for our example failed. But if we add to our model a superfluous constraint

$$Q_3(x) = x_2^2 - x_1 x_3 = 0 \quad (11)$$

and modify respectively the Lagrange function, we radically change the situation.

New Lagrange function L_1 has 3 Lagrange multipliers

$$u^{(1)} = \{u_1, u_2, u_3\}$$

and

$$L_1(x, u^{(1)}) = L(x, u) + u_3(x_2^2 - x_1 x_3)$$

and $K(u)$ changes for

$$K_1(u^{(1)}) = \begin{pmatrix} u_1 & \frac{u_2}{2} & \frac{a_4 - u_3}{2} \\ \frac{u_2}{2} & u_3 & \frac{a_5}{2} \\ \frac{a_4 - u_3}{2} & \frac{a_5}{2} & 1 \end{pmatrix}.$$

It is easy to show that if we choose $u_3 > \frac{a_4^2}{4}$ and u_1 large enough to make $\det(K_1(u^{(1)})) > 0$, the matrix $K_1(u^{(1)})$ becomes positive definite so the function $\psi_1(u^{(1)}) = \inf_x L_1(x, u^{(1)})$ has nonempty domain, and we obtain nontrivial Lagrangian bound ψ_1^* . Moreover, we show later that this bound is exact, i.e. $\psi_1^* = p_6^*$.

3. Quadratic-type problems for finding global minimum of polynomials

Consider a more general problem.

Let $P_{2m}(x_1)$ be a polynomial of one variable x_1 of even degree $2m$ with the eldest coefficient 1:

$$P_{2m}(x_1) = x_1^{2m} + \sum_{i=1}^{2m} a_{2m-i} x_1^{2m-i}.$$

Introduce variables $x_r = x_1^r$, $r = 1, \dots, m$, and represent all monomials x_1^k , $k = 1, \dots, 2m$, as a product of no more than two monomials x_1^r , $0 < r \leq m$. We call these representations as feasible. For some k such representations may be nonunique. So we use so-called "standard" representation:

$$x_1^k = \begin{cases} x_k & \text{for } k \leq m, \\ x_{k-m} x_m & \text{for } m \leq k \leq 2m. \end{cases} \quad (12)$$

By using (12) we obtain the "standard" representation of polynomial $P_{2m}(x_1)$ as a quadratic function in variables x_r , $r = 1, \dots, m$:

$$P_{2m}(x_1) = K_0(x_1, \dots, x_m) = x_m^2 + \sum_{k=1}^{m-1} a_{2m-k} x_{m-k} x_m + \sum_{k=m}^{2m} a_{2m-k} x_{2m-k}. \quad (13)$$

Consider all possible nonstandard representations of monomials x_1^k , $1 < k \leq 2m$. One may obtain the full system of quadratic-type equalities in the form:

$$\pi_{ks} = x_k - x_s x_{k-s} = 0, \quad 1 \leq s \leq \frac{k}{2}, \quad 2 \leq k \leq m. \quad (14)$$

$$\pi_{kt} = x_{k-m} x_m - x_t x_{k-t} = 0, \quad t \leq k - t < m, \quad 2m - 2 > k > m. \quad (15)$$

For example, if $m = 3$, we obtain the constraints $x_2 - x_1^2 = 0$; $x_3 = x_1 x_2$; $x_1 x_3 - x_2^2 = 0$.

Thus, we have reduced the problem of finding a global minimum value P_{2m}^* of polynomial $P_{2m}(x_1)$ to quadratic-type problem: minimize $K_0(x_1, \dots, x_m)$ subject to constraints of the form (14), (15). It is easy to determine that the number of basic equations necessary to convert the initial problem of finding P_{2m}^* in equivalent quadratic-type problem is $(m - 1)$:

$$x_k = x_1 x_{k-1}; \quad 2 \leq k \leq m.$$

The other constraints from (14), (15) are superfluous. The number of all constraints from (14), (15) equals to $\frac{m(m-1)}{2}$.

Consider the Lagrange function for the quadratic-type problem (13), (14), (15). Denote by u_{ks} the Lagrange multipliers corresponding to equations $\pi_{ks} = 0$ in (14) and by u_{kt} to the equations $\pi_{kt} = 0$, from (15). Let us form the usual Lagrange function $L_{2m}(x_m, u_m)$ for the quadratic-type problem (13), (14) and (15), where x_m denotes the vector (x_1, \dots, x_m) , u_m denotes the vector of Lagrange multipliers corresponding to constraints from (13), (14). Let formulate the main Theorem.

Theorem 3. *The best Lagrangian bound $\psi_m^* = \sup_{u_m} \psi(u_m)$, where $\psi(u_m) = \inf_{x_m} L_m(x_m, u_m)$ is equal to P_{2m}^* , i.e. this bound is exact.*

To prove this fact we use following Lemma.

Lemma 2. *The nonnegative polynomial*

$$\overline{P}_{2m}(x_1) = P_{2m}(x_1) - P_{2m}^*$$

can be represented as a sum of squares of real polynomials of degrees not exceeding m .

Proof. All complex roots of equation $\overline{P}_{2m}(x_1) = 0$, if they exist, form pairs of the form $a_\alpha + ib_\alpha$ and $a_\alpha - ib_\alpha$, here a_α and b_α are real numbers. The real roots β_r , $r = 1, \dots, k$, must have the even multiplicity $2\delta_r$, otherwise $\overline{P}_{2m}(x_1)$ cannot be positive.

Let

$$f_1(x_1) = \prod_{\alpha} [x_1 - (a_\alpha + ib_\alpha)] = R(x_1) + iQ(x_1),$$

$$f_2(x_1) = \prod_{\alpha} [x_1 + (a_\alpha - ib_\alpha)] = R(x_1) - iQ(x_1),$$

where R and Q are real polynomials.

Then

$$\begin{aligned} \overline{P}_{2m}(x_1) &= f_1(x_1)f_2(x_1) \left[\prod_r (x_1 - \beta_r)^{2\delta_r} \right] = \\ &= \left[R^2(x_1) + Q^2(x_1) \right] \left[\prod_r (x_1 - \beta_r)^{\delta_r} \right]^2. \end{aligned}$$

So the nonnegative polynomial $\overline{P}_{2m}(x_1)$ can be represented as a sum of squares of real polynomials of degree not exceeding m . The proof is over.

Instead of monomials $x_i = x_1^i, i = 1, \dots, m$, consider another system of basic polynomial functions which corresponds to moving of origin of x_1 by constant h :

$$z_i = (x_1 + h)^i; \quad i = 1, \dots, m; \quad z_0 = 1.$$

We want to expose $x_1^r, i = 1, \dots, 2m$, as quadratic (or linear) functions of z_1, \dots, z_m . Let us use induction in r . For $r = 2$

$$x_1^2 = (x_1 + h)^2 - 2h(x_1 + h) + h^2 = z_2 - 2hz_1;$$

Denote by $Q_k(z_1, \dots, z_m)$ the representations of $x_1^k (k = 1, \dots, 2m - 1)$ as a quadratic function of variables x_1, \dots, x_m . Then

$$Q_{k+1}(x_1, \dots, x_m) = x_1 Q_k(z_1, \dots, z_m) = z_1 Q_k(z_1, \dots, z_m) - h Q_k(z_1, \dots, z_m).$$

One may generate the full system of quadratic equalities for monomials $z_i^r, 2 \leq r \leq 2m$, similar to (14), (15):

$$z_k - z_r z_{k-r} = 0 \text{ for } k \leq m, \quad r \leq \frac{k}{2}; \tag{16}$$

$$z_{k-m} z_m - z_r z_{k-r} = 0 \text{ for } m < k \leq 2m; \quad r \leq \frac{k}{2}; \quad k - r \neq m. \tag{17}$$

When we use the recurrent formula (16) some monomials after multiplication by z_1 may possess nonstandard form and we must add the equalities of the form (16) or (17) with corresponding multipliers to obtain the expressions for $x_1^{(k)}, k = 1, \dots, 2m$, as a quadratic function of variables z_1, \dots, z_m in standard form.

If we substitute these expressions for x_1^k in polynomial $P_{2m}(x_1)$ given in standard form we obtain quadratic function $K_0^{(h)}(z_1, \dots, z_m)$ represented the polynomial $P_{2m}(x_1 + h)$.

Thus $K_0^{(h)}(z_1, \dots, z_m)$ coincides with Lagrange function $L(z_1, \dots, z_m; \overline{U}_h(P_{2m}))$, where $\overline{U}_h(P_{2m})$ is the vector of Lagrange multipliers for equalities (16) and (17), dependent on moving h and coefficients of initial polynomial $P_{2m}(x_1)$ to be minimized. From previous discussion one may obtain the following Lemma.

Lemma 3. *For an arbitrary fixed moving h and a polynomial $P_{2m}(x_1)$ there exists a vector of Lagrange multipliers $\overline{U}_h(P_{2m})$ such that*

$$L_1(x_1, \dots, x_m; U - \overline{U}_h(P_{2m})) = L(x_1, \dots, x_m; U),$$

where L_1 is Lagrange function for quadratic representation of polynomial $P_{2m}(x_1 + h)$.

Corollary 1. *The best Lagrangian quadratic bounds $\psi^*(h)$ for polynomial $P_{2m}(x_1 + h)$ is equal to ψ^* .*

We say that a polynomial $P_{2m}(x_i)$ possesses E -property if the best Lagrangian quadratic bound of it $\psi^* = \min_{x_1} P_{2m}(x_i) = P_{2m}^*$. It is obvious.

Lemma 4. *If the polynomial $P_{2m}(x_1)$ of degree $2m$ with the eldest coefficient 1 possesses E -property then for arbitrary $h \in R$ the polynomial $P_{2m}(x_1 + h)$ also possesses E -property.*

Let the point $x_1^* \in E^n$ be a point of global minimum of polynomial $P_{2m}(x_1)$ and $P_{2m}(x_1^*) = P_{2m}^*$. Consider the polynomial $P_{2m}(x_1) = P_{2m}(x_1 - x_1^*) + P_{2m}^*$. The nonnegative polynomial $\bar{P}_{2m}(x_1)$ has its global minimum value 0 in the point $x_1 = x_1^*$. By Lemma 2 polynomial $\bar{P}_{2m}(x_1)$ can be decomposed into a sum of squares of real polynomials of degree not exceeding m .

$$\bar{P}_{2m}(x_1) = \sum_{i=1}^N [P_m^{(i)}]^2. \quad (18)$$

Since $\bar{P}_{2m}(0) = 0$, each of polynomials $P_m^{(i)}(x_1)$ has no constant part.

Let $\bar{L}^*(x_1, \dots, x_m; U)$ be the Lagrange function for quadratic representation of $\bar{P}_{2m}(x_1)$. Set $U = 0$. Then $\bar{L}^*(x_1, \dots, x_m; 0) = \bar{P}_{2m}(x_1)$ and can be decomposed into a sum of squares (see (18)). Each of polynomials $P_m^{(i)}$ can be represented as a linear form in variables x_1, \dots, x_m , so $\min_x \bar{L}^*(x_1, \dots, x_m; 0) = \bar{L}^*(0, 0, \dots, 0; 0) = 0$, the best Lagrangian quadratic bound is exact.

From Lemma 4 we obtain that the Lagrangian quadratic bound for initial problem: to find a global minimum for polynomial $P_{2m}(x_1)$ is also exact. The proof of Theorem 3 is over.

For polynomials of several variables the situation is more complicated. Further we shall give a review of main results in the theory of dual quadratic bounds (with using of superfluous constraints) for polynomials of several variables.

Let R^n be n -dimensional linear space of real vectors $x = (x_1, \dots, x_n)$, $P(x) = P(x_1, \dots, x_n)$ be a polynomial real function defined on R^n . Consider the problem of finding

$$f^* = \inf_{z \in R^n} P(x_1, \dots, x_n).$$

We will be interested in nontrivial case where $f^* > -\infty$, i.e. $P(x)$ is bounded from below. Such polynomials will be called *BB*-polynomials. It is clear that if $P(x)$ belongs to *BB*-class, then for any i , $1 \leq i \leq n$, the highest degrees S_i of variables x_i must be even. Note that the problem of the *BB*-property is in general similar by its computational complexity to the problem of finding f^* .

Let $S_i = 2l_i$, $i = 1, \dots, n$, and $P(x)$ be recorded in a standard form as a sum of monomials with some real nonzero coefficients. For compact record of monomials we use a vector of degree $\alpha = (\alpha_1, \dots, \alpha_n)$ with nonnegative integer entries and symbols $R[\alpha]$ of the corresponding monomials.

$$R[\alpha] = x_1^{\alpha_1}, \dots, x_n^{\alpha_n}; \quad \alpha_i \leq S_i, \quad i = 1, \dots, n.$$

So

$$P(x) = \sum_{\alpha} c_{\alpha} R[\alpha] \text{ in new variables, } 0 \leq \alpha_i \leq 2l_i, i = 1, \dots, n. \quad (19)$$

Let all monomials of polynomial $P(x) = P(x_1, \dots, x_n)$ have maximal degree on variable x_i equal $2l_i$ ($i = 1, \dots, n$). Consider “feasible” monomials $R[\alpha] = \prod_{i=1}^n x_i^{\alpha_i}$, where $\alpha_i \leq l_i$, $i = 1, \dots, n$.

For each monomial $R[\alpha]$ choose the “standard” representation of monomial $R[\alpha]$ as a product of two feasible monomials:

$$R[\alpha] = R[\alpha_1(\alpha)] \cdot R[\alpha - \alpha_1(\alpha)],$$

$$\alpha_1(\alpha), (\alpha - \alpha_1(\alpha)) \geq 0.$$

Moreover, an integer vector

$$\alpha_1(\alpha) \geq \alpha - \alpha_1(\alpha) \geq 0,$$

or

$$\alpha_1(\alpha) \geq \frac{\alpha}{2}.$$

Consider for each monomial $R[\alpha]$, $\alpha_i \leq 2l_i$, $i = 1, \dots, n$ the full set of quadratic-type equalities in feasible R -variables of the form:

$$R[\alpha_1(\alpha)] \cdot R[\alpha - \alpha_1(\alpha)] - R[\beta] \cdot R[\alpha - \beta] = 0, \quad (20)$$

where β runs all possible values of integer feasible vectors, not equal $\alpha_1(\alpha)$, such that

$$\beta \geq \alpha - \beta \geq 0, \text{ so } \beta \geq \frac{\alpha}{2}.$$

We obtain the “full” set of constraints if we record the equations of the type (20) for all $\alpha = (\alpha_1, \dots, \alpha_n), \alpha_i \leq 2l_i, i = 1, \dots, n$.

Using the full family of equalities of the type (20) one can get all possible representations of polynomial $P(x_1, \dots, x_n)$ as quadratic function in feasible R -variables:

$$P(x_1, \dots, x_n) = L(R, u) = \sum_{\alpha} c_{\alpha} R[\alpha_1(\alpha)] \cdot R[\alpha - \alpha_1(\alpha)] + \sum_{\alpha, \beta} u_{\alpha\beta} (R[\alpha_1(\alpha)] \cdot R[\alpha - \alpha_1(\alpha)] - R[\beta] \cdot R[\alpha - \beta]), \quad (21)$$

where $u_{\alpha\beta}$ are arbitrary multipliers in the left part of equalities (20) with corresponding $\{\alpha, \beta\}$, c_{α} are coefficients in usual representation $P(x)$ in (19).

On the other hand, we can consider $L(R, u)$ as a Lagrange function of quadratic-type problem in feasible variables $R(\alpha)$:

to minimize

$$P(x_1, \dots, x_n) = \sum_{\alpha} c_{\alpha} \prod_{i=1}^n x_i^{\alpha_i} = \sum_{\alpha} c_{\alpha} (R[\alpha_1(\alpha)] \cdot R[\alpha - \alpha_1(\alpha)])$$

subject to the full set of constraints of the form (20), $u = \{u_{\alpha\beta}\}$ are Lagrange multipliers.

Our aim is to find conditions when

$$\psi^* = \sup_u \left[\inf_R L(R, u) \right] = P^* = \min_x P(x).$$

These conditions are formulated in

Main Theorem. *Let polynomial function in n variables $P(x)$ reaches the global minimum at point x^* and $P(x^*) = P^*$.*

Then a dual quadratic bound $\psi^ = P^*$ if and only if the nonnegative polynomial $\bar{P}(x) = P(x) - P^*$ can be represented as a sum of squares of polynomials, which have only feasible monomials.*

Proof of the Main Theorem. We say that a polynomial $P(x)$, $\min_{x \in E^n} P(x) = P(x^*) = P^*$, possesses E -property if Lagrangian quadratic bound is exact i.e. $\psi^* = P^*$. To prove the Main Theorem we use the next

Lemma. *If polynomial $P(x), x \in E^n$ possesses E -property then for arbitrary $a \in E^n$ the polynomial $P_a(x) = P(x + a)$ also possesses this property (see [33]).*

Due to the previous Lemma one can suppose without loss of generality that optimal point is $x^* = 0$.

Let BB -polynomial $\overline{P}(x) - P^*$ is represented as a sum of squares of real polynomials $R_i(x)$, $i = 1, \dots, k$, i.e.

$$\overline{P}(x) = \sum_{i=1}^k [R_i(x)]^2.$$

Replace each monomial M_{α^i} contained in polynomial $R_i(x)$ by the corresponding feasible variable $R[\alpha^{(i)}]$. The output of $[R_i(x)]^2$ for each i can be represented as a sum of monomials of the form $C_s^{(i)} C_t^{(i)} R[\alpha^{(it)}] R[\alpha^{(is)}]$. After summing the similar terms, one obtains for each possible $R[\alpha]$ the corresponding coefficient

$$C_\alpha = \sum_i \sum_{(s,t)} C_s^{(i)} C_t^{(i)},$$

where pairs (s, t) are such that $\alpha^{(i,s)} + \alpha^{(i,t)} = \alpha$.

Choose Lagrange multipliers in expression for $L(R, u)$ (21) equal to zero. Then the value of corresponding Lagrange function $L(x, u)$ coincides with the objective function. Since the objective function minus P^* is a sum of squares, the corresponding representation of a quadratic part in $R[\alpha]$ variables is positive semidefinite, so $\overline{u} = 0$ belongs to $\text{dom } \psi$. $x^*(0)$ is a solution of linear system of equations in variables $R[\alpha]$:

$$Q_i[R] = 0, \quad i = 1, \dots, m.$$

The optimal value of the function $\overline{P}(x)$ equals to zero, i.e. $\overline{P}(0)$. So, due to the Lemma, $P(x)$ can be represented as a sum of squares of real polynomials. $P(x)$ possesses E -property.

Continue the proof of the Main Theorem. Let polynomial $P(x)$ possesses E -property and $\min_{x \in E^n} P(x) = P(x^*) = P^*$. This polynomial $\overline{P}(x) = P(x) - P^*$ takes its minimum at the same point x^* and $\overline{P}(x^*) = 0$. Since $P(x)$ possesses E -property there exists u^* such that for the corresponding quadratic-type problems in variables $R[\alpha]$ Lagrange function $L(R, u)$, when $u = u^*$ is positive-semidefinite in R -variables, so it can be represented as a sum of squares of linear functions in feasible R -variables

$$L(R, u^*) = \sum_{i=1}^k \{l_i^{u^*}[R]\}^2.$$

Instead of each term of such defined linear functions in $R[\alpha]$ variables one can substitute corresponding possible feasible monomials in x variables. So

we obtain the representation $\overline{P}(x)$ in the form

$$\overline{P}(x) = \sum_{i=1}^k \left\{ \left[\sum_{i=1}^n c_{ij} M_{ij}(x) \right]^2 \right\},$$

where $\{c_{ij}\}_{i=1}^k$ are vectors of coefficients of linear functions $l_i^{u^*}$, $i = 1, \dots, k$, $M_{ij}(x)$ are corresponding monomials. So $\overline{P}(x)$ can be represented as a sum of squares of real polynomials.

The Main Theorem is proved.

A great mathematician D.Hilbert [21] considered the problem of representation of nonnegative polynomial forms as a sum of squares polynomials more than hundred years ago. He proved that if dimension $n = 1$ or 2 , arbitrary nonnegative forms (homogenous polynomials) can be represented as a sum of squares. If $n = 3$ and the number of variables ≤ 3 , also the corresponding nonnegative forms can be represented as a sum of squares. But if $n = 3$ and degree $2m \geq 4$ there exist nonnegative polynomial forms which cannot be represented as a sum of squares. Our approach gives a possibility not only to determine whether a given nonnegative polynomial can be decomposed in a sum of squares, but to find such decomposition if it exists using the described above algorithm of finding Langrangian quadratic bounds with use of full set of superfluous constraints.

4. Quadratic-type models and upper bounds for the problems of finding the maximum weighted independent set in graphs

Let an undirected graph $G(V, E)$ be given: $V = \{1, \dots, n\}$ is the set of vertices, E is the set of edges; $(i, j) \in E$ is the edge with end points i and j belonging to V , (i, j) and (j, i) are equivalent symbols. The subset $I \subseteq V$ is called independent (stable) if there is no pair i and j such that $i, j \in I$ and $(i, j) \in E$, $i \neq j$.

A subset $K \subseteq V$ is called a clique if for all pairs $(i, j) \in K$, $(i, j) \in E$, $i \neq j$.

Denote by $A_G = \{a_{ij}\}_{i,j=1}^n$ the adjacency matrix of $G(V, E)$:

$$a_{ij} = 1, \text{ if } (i, j) \in E; \quad a_{ij} = 0, \quad \text{if } (i, j) \notin E.$$

The **complement** graph to $G(V, E)$ is $\overline{G}(V, \overline{E})$ with the same vertex set and

$$\overline{E} = \{(i, j), \quad i \neq j / (i, j) \in \overline{E}, \quad \text{if } (i, j) \notin E\}.$$

The graph $G(V, E)$ may be vertex weighted if for every $i \in V$ the weight $w_i \geq 0$ is given.

Let $w = \{w_i\}_{i \in V}$. For a subset $S \subseteq V$ we define $W(S)$ (weight of S) as

$$W(S) = \sum_{i \in S} w_i.$$

We call

$$G(S) = G(S, E \cap S \times S)$$

the subgraph of $G(V, E)$ induced by S . The maximum weight clique problem is to find a clique of maximum weight.

The maximum (weight) independent (stable) set problem is to find an independent set of maximum cardinality (of maximum weight). The size of a maximum independent set is the stability number of G , denoted by $\alpha(G)$. The maximum weight independent set is denoted by $\alpha_w(G)$. It is easy to see that S is the clique of G if and only if S is an independent set of complement graph \overline{G} . So, any result obtained for one of the mentioned problems can be reformulated for another problem. Both of these problems are *NP*-complete for the class of arbitrary graphs. But for some specific classes of graphs the polynomial-time algorithms were constructed.

The weighted maximum stable set problem in graphs can be formulated as the following 0–1 problem: to find

$$\alpha_w(G) = \max(w, x), \quad x = \{x_1, \dots, x_n\}, \quad (22)$$

$$x_i + x_j \leq 1 \quad \text{for all } (i, j) \in E, \quad (23)$$

$$x_k \in \{0, 1\} \quad \text{for all } k \in V. \quad (24)$$

We introduce the stable set polytope

$$\text{STAB}(G) := \text{conv}\{x^S \in R^V \mid S \subseteq V \text{ is a stable set}\}$$

defined as the convex hull of the incidence vectors of all stable sets of vertices of G . $\alpha_w(G)$ is equal to the maximum value of linear function (w, x) on convex polytope $\text{STAB}(G)$. Of course, it is very useful to represent the $\text{STAB}(G)$ by a system of linear inequalities. Unfortunately, in general case, it is a very hard problem. Therefore we consider some particular cases.

The linear relaxation of the problem (22)–(24) is to find:

$$\alpha_1(G, V) = \max(w, x), \quad (25)$$

$$x_i + x_j \leq 1, \forall (i, j) \in E, \quad (26)$$

$$0 \leq x_k \leq 1, \quad k = 1, \dots, n. \quad (27)$$

Theorem [20]. *The inequalities (26), (27) give full description of $\text{STAB}(G)$ if and only if G is bipartite. Hence, for bipartite graphs the problem (22)–(24) can be solved in polynomial time, as LP problem, if weights are rational.*

The minimal graphs for which inequalities (26), (27) are not sufficient to describe $\text{STAB}(G)$ are the odd circuits. In fact, if $G(V, E)$ is an odd circuit then $|V| = 2m + 1$, and the point $x_i = \frac{1}{2}, i \in V$ satisfies the inequalities (26), (27) but does not belong to $\text{STAB}(G)$. So we can propose a new class of linear inequalities valid for $\text{STAB}(G)$, so-called **odd circuit constraints**,

$$\sum_{i \in V(C)} x_i \leq \frac{|V(C)| - 1}{2} \quad \text{for each odd circuit } C, \quad (28)$$

where $V(C)$ is the set of vertices that lie in the circuit C .

Let us call the graph t -perfect if (26), (27) and (28) are enough to describe $\text{STAB}(G)$.

In general case we do not know whether the problem of checking the t -perfectness is in NP or in P. Despite this fact a maximum weight stable set in a t -perfect graph can be found in polynomial time by using a slight modification of the ellipsoid method (see [20], p.276). This problem is reduced to the LP-problem (25)–(27), (28) with possibly exponentially many odd circuits constraints (28). But for obtaining cutting plane in the ellipsoid method it is enough to have one constraint of type (28) that is not satisfied for a given x . The search of such constraint is equivalent to finding the shortest weight odd circuit. For the last problem there exist polynomial time algorithms ([20], p.236).

A partition of V into stable sets (cliques, respectively) is called a **coloring (clique covering, respectively)** of G . The coloring number (respectively, clique covering number) is the smallest number of stable sets in a coloring (respectively, cliques in a clique covering) of G , and is denoted by $\chi(G)$ (respectively, $\bar{\chi}(G)$). It is clear that

$$\chi(G) = \bar{\chi}(\bar{G}).$$

Each stable set in G has no more than one representative in each clique, hence, we have the inequality:

$$\alpha(G) \leq \bar{\chi}(G).$$

Similarily we obtain

$$\omega(G) \leq \chi(G),$$

where $\omega(G)$ is a clique number of graph G . Berge called a graph G **perfect** if the equality

$$\omega(G') = \chi(G')$$

holds for every induced subgraph G' of G . The first Berge conjecture about perfect graphs was the following [5], [6]: *The complement of a perfect graph is also perfect.* This was proved by well known Hungarian mathematician Lovász in 1972 [25]. In 1962 Berge also proposed the **Strong Perfect Graph Conjecture**. *The graph is perfect if and only if it or its complement, does not contain an odd circuit of length at least five as an induced subgraph.* This conjecture is still unsolved.

In [24] L. Lovász proposed and proved some interesting upper bounds for $\alpha(G)$, which are exact for the class of perfect graphs. Similar results one may find in [26], [31]. Further these results were generalized for a weighted problem of finding $\alpha_w(G)$ (see [20], chapter 9.3).

Below we shall formulate two general extremal matrix problems studied while finding upper bounds $\nu_w(G)$ for the weighted independent set problem (see [20], §9.3).

I. Let $G = \{V, E\}$ be a graph, $|V| = n$, $w = \{w_i\}_{i=1}^n$ be a vector of vertex weights. Consider F , the class of symmetric matrices $n \times n$ $S(y) = \{s_{ij}(y)\}_{i,j=1}^n$ dependent of parameter vector $y = \{y_1, \dots, y_n\}$, which have following properties for arbitrary $y \in R^n$:

- (a) $s_{ij}(y) = 0$, if $(i, j) \in E$;
- (b) the matrices $S(y) \in F$ are positive semidefinite;
- (c) $\sum_{i=1}^n s_{ii}(y) = 1$.

The problem of finding Lagrangian bound $\nu_w(G)$ was reduced to the following extremal problem: to find

$$\max_{\{y: S(y) \in F\}} \sum_{i,j=1}^n \sqrt{w_i w_j} s_i(y) s_j(y).$$

II. The other upper bounds equal to the first one has the form:

$$\vartheta_w(G) = \min_{A \in \Sigma_n} \lambda_{\max}[A + W], \quad (29)$$

where Σ_n is the class of symmetric $n \times n$ matrices, $A = \{a_{ij}\}$, $a_{ii} = 0$ for all $i \in V$, $a_{ij} = 0$ for all i, j nonadjacent in G , $W = \{\sqrt{w_i \cdot w_j}, i, j \in V\}$, $\lambda_{\max}(\cdot)$ denotes the maximal eigenvalue.

Since $\lambda_{\max}(X)$ is a nondifferentiable convex function of entries X , the problem (29) is a typical problem of nondifferentiable optimization and may be solved particularly by r -algorithm. In the case when all weights are 1, the problem (29) is reduced to minimization of maximal eigenvalue on some set of matrices with variable entries. The upper bounds (I) and (II) for $\alpha_w(G)$ were obtained by specific technique of coding theory, namely, by orthonormal representation of graphs (see [24], [20]).

These investigations were stimulated by works of C. Shannon, founder of classical information theory (see [30]). In this paper C. Shannon give the notion of information capacity of graph, which is connected with the problem of speed of transferring long messages, in which some symbols can be mixed. He proposed a linear upper bound for information capacity of graph:

$$\mu(G) = \max \sum_{i=1}^n x_i, \quad (30)$$

$$\sum_{i \in S} x_i \leq 1, \quad \text{for an arbitrary clique } S \in G, \quad (31)$$

$$0 \leq x_k \leq 1, \quad k = 1, \dots, n. \quad (32)$$

(Here $G(V, E)$ is a graph, V is a set of symbols and E is a set of edges, corresponding to mixed symbols.) For perfect graph such estimate is exact and coincides with Lóvasz $\nu_w(G)$ bounds.

We can obtain the same bounds using quadratic-type formulation of the problem of finding maximal weighted stable set in graph G and calculating corresponding Lagrangian bounds.

Namely, let the graph $G = (V, E)$, $|V| = n$, $w = \{w_i\}_{i=1}^n$ (the vector of weights) be given. Consider the quadratic-type problem:

to find

$$\alpha_w(G) = \max \sum_{i=1}^n w_i x_i \quad (33)$$

subject to constraints:

$$x_k^2 - x_k = 0, \forall k = 1, \dots, n \text{ (boolean property)} \quad (34)$$

$$x_i x_j = 0, \forall (i, j) \in E. \quad (35)$$

Let $u = (\{u_k\}_{k=1}^n, \{u_{ij}\}_{i,j} \in E)$ be a vector of Lagrange multipliers. Consider the optimal Lagrangian bound

$$\rho_w(G) = \inf_u \sup_x L(x, u).$$

In our case the function $\psi(u) = \sup L(x, u)$ has nonempty domain. As proved in [33], $\rho_w(G)$ has the same value, as Lóvasz’s bounds.

These bounds are exact for arbitrary weights if G is a perfect graph. For nonperfect graphs one can find such weights, that corresponding bounds are not exact.

Thus, using a natural quadratic-type formulation of the problem of finding maximal weight stable number in a graph and applying to this problem standard technique of finding Lagrangian bounds we get the same results as obtained by using “specific” approach of orthonormal representation of graphs.

Consider dual quadratic Lagrangian bounds $\nu_w(G)$ for $\alpha_w(G)$. We want to improve such bounds. For this aim one may use superfluous quadratic inequalities in the problem formulation, for example

$$x_i x_j \geq 0 \text{ for all (some) nonadjacent pairs } (i, j) \tag{36}$$

or

$$x_k(x_i + x_j) \leq x_k, \text{ for any } k \text{ and } (i, j) \in E, (i, j, k) \in V. \tag{37}$$

The corresponding modification of Lagrange functions may lead sometimes to more precise dual bound for $\alpha_w(G)$ without essential complication of calculations. So, if we add to the constraints (34), (35) the constraints of the form (36), (37) we then take into attention not only clique constraints (31) but also the odd circuit constraints (28).

Even if we use only one trivial family of superfluous inequalities (36) we can considerably improve the dual upper bounds for $\alpha_w(G)$ in some cases. The corresponding Lagrange function $L_1(x, \lambda)$ can be constructed as follows:

$$L_1(x, \lambda) = \sum_{i=1}^n w_i x_i + \sum_{(i,j) \in E} \lambda_{ij} x_i x_j + \sum_{k=1}^n \lambda_k (x_k^2 - x_k) - \sum_{(i,j) \in \bar{E}} \lambda_{ij}^{(1)} x_i x_j,$$

where $\lambda = \{ \{ \lambda_k \}_{k=1}^n, \{ \lambda_{ij} \}_{i,j \in E}, \{ \lambda_{ij}^{(1)} \}_{i,j \in \bar{E}} \}$.

Let

$$\varphi_1(\lambda) = \sup_x L_1(x, \lambda), \quad \varphi_1^*(G) = \inf_{\lambda \in \Omega^+} \varphi_1(\lambda)$$

(here Ω^+ is the domain of function φ_1 in λ).

There are some graphs G for which the upper bound φ_1^* for $\alpha(G)$ is much better than $\nu(G)$. For example, let vertices of graph $G_6 = (V, E)$ correspond to the integer numbers from 0 to 63, written in binary codes of length 6, and two vertices v_1, v_2 are joined by edge if Hamming distance $d(v_1, v_2)$ is no more than 3. It is easy to show that $\alpha(G_6) = 4$ (see

[26]) (for instance the maximum stable set is : $\{v_1 = (000000), v_2 = (111100), v_3 = (110011), v_4 = (001111)\}$. It was calculated, that $\nu(G_6) = \frac{16}{3}$, but $\varphi_1^*(G_6) = 4$. It is a bright example of the fact that adding of superfluous constraints to nonconvex problem may considerably improve the dual (Lagrange) bound.

The number of constraints (36) is large, but we may use the fact, that in optimal solution the different dual variables, corresponding to pairs (i, j) , which have the same Hamming distance, are equal. Due to this fact in our example the number of dual variables for the constraints (36) can be reduced to 6 variables. The results of numerical experiments one may find in [33], p.259, [36].

5. Using of the Fan Ky theorem for obtaining dual bounds in some problems of graph theory

The Lovasz's estimate $\nu_w(G)$ may be obtained also in the following way. Let $G = (V, E)$ be a simple undirected graph without loops, $V = \{1, \dots, n\}$. Denote by $x(S) = \{x_i(S)\}_{i=1}^n$ the indicator vector of the subset $S \subseteq V$. Let $W = \{\sqrt{w_i w_j}\}_{i, j \in V}$ be a $n \times n$ - matrix, $u = \{u_{ij}\}_{i, j \in E}$;

$$A_G(u) = \{u_{ij}, (i, j) \in E; 0, \text{ otherwise}\}$$

be a subclass of symmetric $n \times n$ matrices. For arbitrary $S \subseteq V$ consider an n -dimensional vector

$$w(S) = \left\{ \frac{\sqrt{w_i}}{\sqrt{\sum_{j \in S} w_j}}, i \in S; 0, \text{ otherwise} \right\}.$$

Note that $\|w(S)\| = 1$, and for arbitrary u and a stable set S

$$(W + A_G(u)w(S), w(S)) = (Ww(S), w(S)) = \frac{\sum_{i, j \in S} w_i w_j}{\sum_{k \in S} w_k} = \sum_{i \in S} w_i. \quad (38)$$

Let $\nu_G(u)$ be the maximal eigenvalue of $M(u) = W + A_G(u)$. From (38) one can obtain:

$$\nu_w(u) = \lambda_1[W + A_G(u)] = \max_{\{z: \|z\|=1\}} (M(u)z, z) \geq (M(u)w(S), w(S)) = \sum_{i \in S} w_i$$

for arbitrary u and a stable set S . So, $\nu_w(u)$ is an upper bound for the maximum weight of stable sets for every u .

Thus, $\nu_w(G) = \min_u \lambda_1(W + A_G(u)) \geq \alpha_w(G)$ is a well known estimate of Lovasz.

In a similar way one can obtain the upper bound for $\alpha^{(k)}(G)$, where $\alpha^{(k)}(G)$ is the size of the largest induced k -partite subgraph of $G = (V, E)$, i.e. the maximum number of nodes, that can be covered by k subsets of V so that no edge has both ends in any subset.

Let S_1, S_2, \dots, S_k be k pairwise nonintersecting stable subsets of V . Introduce the family of n -dimensional vectors $\{Y_r\}_{r=1}^k, Y_r = (y_r^{(1)}, \dots, y_r^{(n)})$,

$$y_r^{(i)} = \begin{cases} \frac{1}{\sqrt{|S_i|}} & , i \in S_r, r = 1, \dots, k; \\ 0 & , \text{otherwise} \end{cases}$$

Vectors $Y_r, r = 1, \dots, k$, form an orthonormal system. Denote by $Y(k)$ the $n \times k$ matrix with columns $Y_r, r = 1, \dots, k$.

Let $A_G(u), u = \{u_{ij}\}_{(i,j) \in E}$, be the parametric family of symmetric $n \times n$ matrices with entries $a_{ij}(u)$

$$a_{ij} = \begin{cases} u_{ij} & , \text{if } (i, j) \in E, \\ 1 & , \text{otherwise.} \end{cases}$$

It is easy to verify that for arbitrary $u, \text{tr}(A(u)Y(k)[Y(k)]^{(t)}) = \sum_{r=1}^k |S_r|$. Due to Fan Ky Theorem (Theorem 1), $\max_{Y \in M_{n,k}} \text{tr}[A(u)Y Y^{(t)}]$, where $M_{n,k}$ is the class of matrices $n \times k$ with orthonormal system of k columns, is equal to the sum of k largest eigenvalues of symmetric matrix $A(u)$.

So,

$$S_{n,k}(A(u)) = \sum_{r=1}^k \lambda_r[A(u)] \geq \sum_{r=1}^k |S_r|$$

for arbitrary u and any k -partite induced subgraph of G with stable sets S_1, \dots, S_k . Thus,

$$\nu^{(k)}(G) = \min_u S_{n,k}(A(u)) \geq \alpha^{(k)}(G). \tag{39}$$

Calculation of $\nu^{(k)}(G)$ is the problem of nonsmooth convex optimization and can be solved by subgradient-type methods, particularly, by r -algorithm. Procedure of finding of a subgradient is describe in section 1. of this article.

The upper bound (39) for $\alpha^{(k)}(G)$ was first derived by Narasimhan and Manber in [27]. When $k = 1$, one obtains the Lovasz's estimate $\nu(G)$ for the stable number $\alpha(G)$.

Calculation of $\nu^{(k)}(G)$ is reduced to the minimization of $|E|$ -dimensional convex nonsmooth function of matrix parameters u , which enter into matrix affinely.

If $\alpha^{(k)}(G) = |V| = n$, it means that graph $G = (V, E)$ has a true vertex colouring in k colours. So if $\nu^{(k)}(G) < n$ the graph G cannot be coloured by k colours. Thus, we can use the upper bound $\nu^{(k)}(G)$ for studying some of the colouring problems.

The problem of minimizing weighted sums of the k largest eigenvalues of a parametric family of symmetric matrices has many combinatorial applications. The best known of them is the graph partitioning problem. The problem is to divide n nodes of a given graph $G(V, E)$ into k disjoint subsets with given cardinalities $m_1 \geq \dots \geq m_k$ in a way to minimize the total number of edges connecting different subsets. The problem is NP-hard and therefore we have no hope to get a computationally "good" algorithm for its precise solution in general case. But Donath and Hoffman proposed in [13], [14] an effective lower bound for this problem that gives us possibility to obtain good enough approximate solutions or test the accuracy of the solution that can be received by other relatively simple heuristic methods.

Let $S_n^k(x) = \sum_{i=1}^k \lambda_i(A(x))$ be a sum of k largest eigenvalues of matrix $A(x)$. We know that $S_n^k(x)$ is a convex function in x . In the case of graph partitioning problem the variables x are contained only in diagonal elements of $A(x)$. The lower bound for the partitioning problem given in [13], [14] is:

$$\rho^*(A, m) = -\frac{1}{2}s^*, \quad (40)$$

where s^* is the solution of the following problem: minimize

$$S_n^k(x, m) = \sum_{i=1}^k m_i \lambda_i(A_0 + D(x)) \quad (41)$$

subject to the constraint $\text{tr}D(x) = 0$,

where the nondiagonal elements a_{ij}^0 of the symmetric matrix A_0 are equal to one, if the i -th and j -th nodes are connected, and zero-otherwise; the diagonal elements are defined by

$$a_{ii}^0 = - \sum_{j=1; j \neq i}^n (a_{ij}^0), \quad i = 1, \dots, n,$$

$D(x)$ is a diagonal matrix with elements

$$d_{ii} = x_i, \quad i = 1, \dots, n, \quad \sum_{i=1}^n x_i = 0.$$

Note that $m_1 \geq m_2 \geq \dots \geq m_k$, therefore due to Lemma 1 the problem (41) is the problem of convex programming.

The problem (41) is a special case of minimization of the weighted sum of largest eigenvalues of parametric family of matrices (see Section 1.).

One can give a more general formulation of the graph partitioning problem (GPP) by introducing nonnegative weights of edges. As earlier we have a set of integers $m_1 \geq m_2 \geq \dots \geq m_k$ with $\sum_{j=1}^k m_j = n$. Denote by m the k -vector made up of m_j -s.

Let $G = (V, E)$ be a complete graph with $|V| = n$ and each edge $(i, j), i < j$ has weight $w_{ij} \geq 0$. We want to part a set of vertices V into k subsets such that the j -th subset has a prescribed cardinality m_j ; and that sum of the weights of those edges whose endpoints are in different subsets is minimized. Let us denote this minimum value by $\pi_m(G)$. Let $W = \{a_{ij}\}$ be a matrix with $a_{ij} = w_{ij}, i < j$, and $a_{ii} = 0, i = 1, \dots, n$.

Donath and Hoffman [13] also proved the following inequality

$$\pi_m(G) \geq U_m^W(G) = -\frac{1}{2} \min_x \sum_{j=1}^k m_j \lambda_j(W + \text{diag}(x))$$

subject to constraint

$$\sum_{i=1}^n x_i = - \sum_{(i,j) \in E} w_{ij}.$$

We can use subgradient-type methods for calculating the lower bounds for $\pi_m(G)$.

The first practically efficient algorithm for obtaining the lower bounds for graph partitioning problems using the minimization of $S_n^k(x, m)$ (see (41)) was proposed in [9] by Cullum et al. for equal $m_i, i = 1, \dots, k$. They noticed that the problem (41) is a problem of nondifferentiable optimization and, moreover as a rule, the optimum point is the point of nondifferentiability of minimized function. Their method is based on idea of smoothing the function $S_n^k(x, m)$ by using information not only about the k largest eigenvalues and corresponding eigenvectors, but also the information about other eigenvalues and eigenvectors, if they are slightly different from k -th eigenvalue.

We used for this purpose one of the modifications of the subgradient-type method with space dilation in the direction of difference of two successive subgradients (r -algorithm) and obtained good results in test experiments (see [33]). The technique for calculation of subgradients is described in the section 1 of this article.

When n is even, $k = 2$, and $m_1 = m_2 = \frac{n}{2}$, the graph partitioning problem can be reduced to the so-called graph bisection problem that can be considered as max-cut problem with one additional constraint.

6. Lagrangian bounds for the maximum cut problem

One of the most bright examples of using quadratic superfluous constraints for improving dual bounds in quadratic-type problems is connected with the max-cut problem.

Let $G(V, E)$ be an ordinary graph with the vertex set $V = \{1, \dots, n\}$ and the edge set $E = \{(i, j) = (j, i)\}$, where (i, j) is the edge, linking vertices (i, j) . The weight function W is given by symmetric $n \times n$ matrix

$$W = \begin{cases} 0 & \text{for } (i, j) \notin E, \\ w_{ij} & \text{for } (i, j) \in E. \end{cases}$$

Let the vertex set be divided into two nonempty nonintersecting parts V_1 and V_2 : $V = V_1 \cup V_2$. We say that the edge (i, j) belongs to the cut $R(V_1, V_2)$ if this edge have its ends in different subsets of subdivision $V = V_1 \cup V_2$. We must find such partition $V = V_1 \cup V_2$ that the sum of weights of all edges, belonging to the corresponding cut, is maximal.

The max-cut problem is NP-complete, it is proved [37] that it preserves this property even for class of graphs with degrees of vertices not exceeding 3. But for the subclass of the so-called weakly bipartite graphs the max-cut problem with positive weights of edges can be solved by polynomial-time algorithm.

Let $G = (V, E)$ be a graph and $F \subseteq E$ be an edge subset. The vector $y^F \in R^E$ with $y_e^F = 1$ if $e \in F$ and $y_e^F = 0$ if $e \notin F$ is called the incidence vector of F . The polytope $P_B(G) := \text{conv}\{y^F \in R^E | (V, F) \text{ is a bipartite subgraph of } G\}$ is called the **bipartite subgraph polytope** of G . It is clear that for positive edge weights $W(e)$, $e \in E$, every optimum basic solution of the linear program

$$\max(W, y), \quad y \in P_B(G),$$

corresponds to a cut.

Consider the trivial inequalities:

$$0 \leq y_e \leq 1. \tag{42}$$

The inequalities (42) determine $P_B(G)$ completely if and only if G is bipartite. Consider any odd cycle C in a graph G . It is obvious that full edge set of C cannot belong to a cut. So, one can formulate for an incidence vector y of a cut the following odd cut inequalities:

$$y(C) := \sum_{e \in C} y_e \leq |C| - 1, \quad C \text{ is an odd cycle in } G. \quad (43)$$

Definition 1. A graph $G(V, E)$ that has the property: $P_B(G) = \{y \in R^E \mid y \text{ satisfies all inequalities (42) and (43)}\}$ is called **weakly bipartite**.

In [20] (see chapter 9.3) the polynomial time algorithm is given for checking the feasibility of the vector $\{y_l\}_{l \in E}$; $0 \leq y_l \leq 1$ for weakly bipartite graphs. This algorithm gives also the odd cycle for which constraint (43) is not fulfilled, if such cycle exists.

Thus, one may use the ellipsoid method for finding a maximal cut for weakly bipartite graph with positive weights (see [20]). But this algorithm is not good for practical calculations, because it is very complex and converges slowly. Therefore we consider below quadratic-type formulation of max-cut problem.

Let $x_k \in \{-1, 1\}$ be a binary variable corresponding to the vertex k ($k \in \{1, 2, \dots, n\}$),

$$x_k = \begin{cases} -1, & \text{if } x_k \in V_1, \\ +1, & \text{if } x_k \in V_2. \end{cases}$$

Without loss of generality one can assume that the graph $G(V, E)$ is full. In this case the value of cut may be represented in terms of binary variables as a quadratic function:

$$f(x, W) = \frac{1}{8} \sum_{(i,j), i \neq j} w_{ij} (x_i - x_j)^2 = \frac{1}{4} \left(\sum_{(i,j)} w_{ij} - \sum_{(i,j)} w_{ij} x_i x_j \right).$$

Since $\max f(x, W)$ subject to constraints $x_k^2 - 1 = 0$, $k = 1, \dots, n$, equal to $1/4 \sum_{(i,j)} w_{ij} - \min_x \sum_{(i,j)} w_{ij} x_i x_j$ subject to the same constraints, the max-cut problem can be reduced to the quadratic-type problem:

to find

$$S(W) = \min_x \sum_{i,j=1}^n w_{ij} x_i x_j \quad (44)$$

subject to constraints

$$x_k^2 - 1 = 0, \quad k = 1, \dots, n. \quad (45)$$

Denote by $R^*(W)$ the optimal value of the max-cut. Due to (44), (45)

$$R^*(W) = \frac{1}{4} \left(\sum_{(i,j)} w_{ij} - S(W) \right). \quad (46)$$

Let $u = \{u_1, \dots, u_n\}$ be the vector of Lagrange multipliers corresponding to equalities (45). Then Lagrange function $L(x, u)$ of reduced problem (44)-(45) has the form

$$L(x, u) = (W(u)x, x) - \sum_{k=1}^n u_k,$$

where

$$W(u) = W + \text{diag } u,$$

(diag u is a diagonal matrix with the components $d_{ii} = u_i, i = 1, \dots, n$). Let $\lambda_{\min}(A)$ denotes the minimal eigenvalue of matrix A .

The quadratic part of $L(x, u)$ is a homogeneous quadratic in x function, so

$$\inf_x L(x, u) = \begin{cases} -\infty, & \text{if } \lambda_{\min}(W(u)) < 0, \\ -\sum_{k=1}^n u_k, & \text{if } W(u) \succeq 0, \text{ i.e. } \lambda_{\min}(W(u)) \geq 0. \end{cases}$$

Consider

$$\varphi^* = \min \sum_{k=1}^n u_k, \quad W(u) \succeq 0.$$

By using an exact nonsmooth penalty function (see [7], Chapter 4), the problem of finding φ^* is reduced to an unconstrained optimization of non-differentiable function

$$f(u, s) = \sum_{k=1}^n u_k + s \left[\lambda_{\min}^-(W(u)) \right], \quad (47)$$

where

$$\lambda_{\min}^- = \begin{cases} 0, & \text{if } \lambda_{\min} \geq 0, \\ \lambda_{\min}, & \text{if } \lambda_{\min} \leq 0, \end{cases}$$

s is a penalty multiplier. One can prove, that if $s \geq n$, the problem of minimization of $f(u, s)$ is equivalent to the problem (47).

One can solve problem (47) by using r -algorithm. After obtaining an optimal value of (47) $S(W)$ we can obtain the upper bound for the optimal cut value, using (46).

The alternative approaches for finding upper bounds of the max-cut problems are represented in [14],[2], [10]. In articles [1],[22] the bounds similar to ours are obtained by using another technique. In [16] it is shown how to describe the max-cut problem in terms of semidefinite programming.

Consider a weighted graph $G = (V, E)$ with even number of vertices: $|V| = n = 2s$. The bisection problem for G can be formulated as follows:

$$\text{maximize } f(y) = \frac{1}{4} \left(\sum_{(i,j) \in E} w_{ij} - \sum_{i,j \in E} w_{ij} y_i y_j \right) \tag{48}$$

subject to constraints:

$$y_i^2 - 1 = 0, \forall i \in V = \{1, \dots, n\}, \tag{49}$$

$$\sum_{i=1}^n y_i = 0 \tag{50}$$

(for feasible solution \bar{y} the number of \bar{y}_i having value 1 must be equal to number \bar{y}_i having value -1). One may propose several ways for obtaining upper bounds for the problem (48)–(50):

- (i) use an estimate of the type (40), (41) with $m_1 = m_2 = s$;
- (ii) due to (50), $y_n = \sum_{i=1}^{n-1} y_i$. Set $\tilde{y} = \{y_1, \dots, y_{n-1}\}$.

Consider the problem in variables \tilde{y} which is equivalent to the problem (48)–(50):

$$\text{maximize } \bar{f}(\tilde{y}) = \frac{1}{4} \left(\sum_{(i,j) \in E} w_{ij} - \sum_{i,j \in E} [w_{ij} y_i y_j - \sum_{i=1}^{n-1} w_{in} (y_i y_n + y_n y_i)] \right) \tag{51}$$

subject to constraints:

$$y_i^2 - 1 = 0, \forall i \in \{1, \dots, n-1\}, \tag{52}$$

$$\left(\sum_{i=1}^{n-1} y_i \right)^2 - 1 = 0. \tag{53}$$

One can construct Lagrange function for the problem (51)–(53) in the form:

$$L(\tilde{y}, u) = -\bar{f}(\tilde{y}) + \sum_{i=1}^{n-1} u_i (y_i^2 - 1) + u_n [(\sum_{i=1}^{n-1} y_i)^2 - 1]$$

and consider a marginal function

$$\psi(u) = \inf_{\tilde{y}} L(\tilde{y}, u).$$

$\psi(u)$ is convex function, and $\text{dom } \psi \subseteq \bar{\Omega}^+$, where $\bar{\Omega} + (\Omega^+) = \{\bar{u} \in R^n : L(\tilde{y}, \bar{u}) \text{ is convex (strictly convex) function in } \tilde{y} \text{ for fixed } \bar{u}\}$.

Since $L(\tilde{y}, u)$ has no linear in \tilde{y} terms

$$\psi(u) = -\frac{1}{4}\bar{W} + \sum_{i=1}^n u_i, \text{ if } u \in \text{dom } \psi,$$

where $\bar{W} = \sum_{i,j=1}^n w_{ij}$.

Let $Q(u)$ be a matrix of the quadratic part of $L(x, u)$. Let $\psi^* = \sup_{u \in \text{dom } \psi} \psi(u)$

$$\psi^* = \sup_{u \in \Omega^+} \psi(u) = \{\sup \psi(u) : Q(u) \succeq 0\}.$$

Thus the upper dual bound φ^* for the maximum bisection problem (48)–(50) can be calculated by solving the following problem:

$$\text{find } \psi^* = \min \left[\sum_{i=1}^n u_i + \frac{1}{4}\bar{w} \right] \quad (54)$$

subject to constraint:

$$\lambda_n[Q(u)] \succeq 0. \quad (55)$$

The convex programming problem (54), (55) can be solved by application of r -algorithm for nonsmooth penalty function

$$\varphi_p(u) = \frac{1}{4}\bar{w} + \sum_{i=1}^n u_i + s\lambda_n^+[Q(u)],$$

where

$$\lambda_n^+(u) = \begin{cases} \lambda_n(u), & \text{if } \lambda_n(u) < 0; \\ 0, & \text{if } \lambda_n(u) \geq 0; \end{cases}$$

$s > 0$ is a penalty multiplier;

- (iii) the bisection problem (48)–(50) can be approximately reduced by use of quadratic-type penalty $s(\sum u_i)^2$ for constraint (50) to usual max-cut problem with other weights:

$$w_{ij}^c = w_{ij} + s,$$

where s is a penalty multiplier.

The max-cut problem may be formulated also as the problem of linear programming in edge Boolean variables $y = \{y_{ij}\}_{i,j=1, i \neq j}^n$, $y_{ij} \in \{0, 1\}$ ($y_{ij} = 1$, if (i, j) belongs to the cut $R(V_1, V_2)$, and 0 in the opposite case) (see [3], [4], [19]).

Let $M(G)$ be the convex hull of all feasible solutions of the max-cut problem in edge variables $\{y_{ij}\}$. So, one may reformulate the max-cut problem as a linear programming problem: to find

$$l^*(G, W) = \max_{y \in M(G)} \sum_{(i,j)} w_{ij} y_{ij}.$$

Edge variables y_{ij} and vertex binary variables are linked by simple formula:

$$y_{ij} = \frac{1 - x_i x_j}{2} \text{ for all } y_{ij}, i \neq j. \quad (56)$$

In a large graph the set $M(G)$ has tremendous number of faces. But for special graphs all its faces can be described in compact form. For example, if graph is planar, for each triangle $\{i, j, k\}$, one may write down the inequality

$$y_{ij} + y_{jk} + y_{ik} \leq 2. \quad (57)$$

Barahona and co-workers [3, 4] generalized the algorithm for planar graphs on the family of graphs not contractible to graph K_5 (5- clique).

Denote by $y(N)$ the sum $\sum_{(i,j) \in N} y_{ij}$, where N is a set of edges. Using this notation we can record more general family of inequalities of the following form:

$$\begin{aligned} y(F) - y(C \setminus F) &\leq |F| - 1 \text{ for all circuits} \\ C \text{ of } G \text{ and subsets } F &\subseteq C \text{ with } |F| \text{ odd.} \end{aligned} \quad (58)$$

These inequalities are valid for $M(G)$.

Barahona and co-workers [3], [4] have shown that the solution set of (58) is equal to $M(G)$ if and only if G has no subgraphs contractible to K_5

(K_5 is 5-clique). There is a simple algorithm (see [20], §8.3, p.235-236) for verifying validation of system (58) for given \bar{y} ($0 \leq \bar{y}_{ij} \leq 1$ for all (i, j)).

We construct a new graph $H = (V' \cup V'', E' \cup E'' \cup E''')$, consisting of disjoint copies $G'(V', E')$ and $G''(V'', E'')$ of G and an additional edge set E''' that contains the edges $u'v''$, $u''v'$ for each $uv \in E$. The edges $u'v' \in E'$ and $u''v'' \in E''$ get the weight y_{uv} while the edges $u'v''$ and $u''v' \in E'''$ get the weight $1 - y_{uv}$. For each node $u \in V$ we calculate a shortest path (with respect to weights just defined). Such a path contains an odd number of edges of E''' and corresponds to a closed walk in G , containing u . Clearly, if the shortest of these $[u', u'']$ -paths has the length equal at least one, then y satisfies (58), otherwise there exists a cycle C and a set $F \subseteq C$, $|F|$ odd, such that \bar{y} violates the corresponding inequality. Such linear in \bar{y} inequalities can be converted by substitution (56) in quadratic inequalities in binary variables x which may be added to the set of obtained earlier inequalities in x as superfluous constraints. So we described one of the possible ways to generate the superfluous inequalities for model (44). Further adding of superfluous constraints as a rule essentially improves dual quadratic upper bounds for the max-cut problem. In combination with heuristic methods of constructing feasible cuts such approach allows to obtain an exact solution for graphs with integer weight matrix which subgraphs are not contractible to K_5 ([3], [20]). A more detailed description of corresponding algorithm one can find in [33], §8.1.

Thus, in the case of weakly bipartite graphs the max-cut problem with positive weights of edges reduces to LP problem but the number of inequalities of the type (42) may increase as exponential function of the size of a graph. To overcome this difficulty for obtaining a polynomial time algorithm of solving the max-cut problem in the mentioned above case one can use the ellipsoid method. Certain steps of this method need to solve the problem of finding the odd cycle of minimum weight for checking the feasibility of the inequalities (42). The search of such a cycle can be reduced to at most n applications of the algorithm of finding the shortest even path in a weighted graph (see [20], chapter 8.4).

The quadratic dual bound for the max-cut problem in the case of weakly bipartite graphs is not always exact, but if all weights are nonnegative, the following estimate is proven.

Theorem 4. *Let G be a weakly bipartite graph with non-negative weights. Then*

$$\varphi(G) \leq \frac{5\sqrt{5}(1 + \sqrt{5})}{32} mc(G).$$

In [18, 15] there was proposed a way of generating "good" feasible solutions for the max-cut problem using the eigenvectors associated with $\lambda_n(\overline{W} + \text{diag}(\overline{u}))$ where \overline{u} is an approximation of u^* . Let $s(\overline{u}) = \{s_i(\overline{u})\}_{i=1}^n$ be an eigenvector associated with $\lambda_n(\overline{W} + \text{diag}(\overline{u}))$. Write its entries in a nonincreasing order:

$$s_{i_1}(\overline{u}) \geq s_{i_2}(\overline{u}) \geq \dots \geq s_{i_n}(\overline{u}).$$

For different k , $1 \leq k \leq n$, construct the partition of V :

$$V = S(k) \cup (V \setminus S(k)),$$

where $S(k) = \{i_1, \dots, i_k\}$, and find $\max_k c(S(k)) = c(S(\overline{k}))$, where $c(S(k))$ is the value of cut corresponding to two subdivisions: $S(k)$ and $V \setminus S(k)$. Choose a feasible vector $y(s(\overline{u}))$ corresponding to this partition:

$$y(s(\overline{u})) = \begin{cases} 1, & \text{for } i \in S(\overline{k}), \\ -1, & \text{for } i \in V \setminus S(\overline{k}). \end{cases}$$

In many cases $y(s(\overline{u}))$ is a good approximation of the optimal solution. So, in the process of minimization of $S_N(u)$ by r -algorithm one may use $u(k)$ at each step k not only for obtaining an upper bound for $f(y^*) = mc(G)$, but also for obtaining a feasible integral solution $y(u(k))$ in the described above way. One can use the record value of the objective function for generated feasible solutions as lower bound for $mc(G)$. The results of numerical experiments are presented in [34].

7. The computational results

(1) We describe the results of test calculations for finding global minimum value of some multiextremal polynomial functions of one variable. For beginning, consider the family of 6-th degree polynomials of the form:

$$P_6(x) = x^6 - 2tx^4 + (t^2 - \varepsilon)x^2 = x^2((x^2 - t)^2 - \varepsilon), \quad (59)$$

where t and ε are parameters of this family.

It is easy to see that if $t > 0$, $\varepsilon > 0$, $t^2 - \varepsilon > 0$, polynomials (59) have two global minima:

$$x_1^* = \sqrt{\frac{2t}{3} + \frac{\sqrt{t^2 + 3\varepsilon}}{3}}; \quad x_2^* = -\sqrt{\frac{2t}{3} + \frac{\sqrt{t^2 + 3\varepsilon}}{3}}.$$

t	ε	ψ_2^*	P^*	u_1^*	u_2^*	u_3^*
1.00	1.0000	-1.185185	-1.185185	1.7778	.0000	2.6667
1.00	.1000	-.102387	-.102387	1.0956	.0000	2.0935
1.00	.0100	-.010025	-.010025	1.0100	.0000	2.0099
1.00	.0010	-.001000	-.001000	1.0010	.0000	2.0010
1.00	.0001	-.000100	-.000100	1.0001	.0000	2.0001
1.00	.0000	.000000	.000000	1.0000	.0000	2.0000
.50	1.0000	-.758076	-.758076	.8728	.0000	1.8685
.50	.1000	-.054288	-.054288	.3370	.0000	1.1611
.50	.0100	-.005049	-.005049	.2598	.0000	1.0194
.50	.0010	-.000500	-.000500	.2510	.0000	1.0020
.50	.0001	-.000050	-.000050	.2501	.0000	1.0002
.50	.0000	.000000	.000000	.2500	.0000	1.0000
.25	1.0000	-.562500	-.562500	.5625	.0000	1.5000
.25	.1000	-.031676	-.031676	.1350	.0000	.7347
.25	.0100	-.002593	-.002593	.0718	.0000	.5361
.25	.0010	-.000251	-.000251	.0635	.0000	.5040
.25	.0001	-.000025	-.000025	.0626	.0000	.5004
.25	.0000	.000000	.000000	.0625	.0000	.5000

Table 1. Results of experiments for $P_6(x)$

If $\varepsilon = 0$, polynomials (59) has three global minima: x_1^* , x_2^* and 0.

The results are given in Table 1. Here x_2^* is an approximation of optimal dual bounds obtained for the quadratic-type problems (8) – (11), which correspond to the global minimum of polynomial $P_6(x)$; P^* is an exact value of global minimum of polynomial $P_6(x)$; u_1^* , u_2^* , u_3^* are obtained values of Lagrange multipliers for constraints (9) – (11).

Table 1 shows that in all test experiments ψ_2^* is equal to global minimum of polynomial $P_6(x)$ with good precision. Optimal values of dual variables u_1^* , u_2^* , u_3^* lie at the border of positive definite region of matrix

$$K(u) = K(u_1, u_2, u_3) = \begin{pmatrix} u_1 & u_2/2 & -u_3/2 \\ u_2/2 & u_3 - 2t & 0 \\ -u_3/2 & 0 & 1 \end{pmatrix}.$$

Taking into account that $u_2^* = 0$,

$$K(u_1^*, 0, u_3^*) = \begin{pmatrix} u_1^* & 0 & -u_3^*/2 \\ 0 & u_3^* - 2t & 0 \\ -u_3^*/2 & 0 & 1 \end{pmatrix}$$

the eigenvalues of $K(u_1^*, 0, u_3^*)$ will be following

$$\lambda_1 = u_3^* - 2t; \quad \lambda_{2,3} = \frac{(u_1^* + 1) \pm \sqrt{(u_1^* - 1)^2 + (u_3^*)^2}}{2}.$$

For $\varepsilon > 0$ the polynomials $P_6(x)$ have two global minima and eigenvalues of matrices $K(u_1^*, 0, u_3^*)$ satisfy the condition $\lambda_1 > 0$, $\lambda_2 > 0$ and $\lambda_3 = 0$. If $\varepsilon = 0$, $P_6(x)$ has three global minima, and $\lambda_2 > 0$, $\lambda_1 = \lambda_3 = 0$.

Discribed results show that the method is stable also in the case, when solution is at the border of the set of positive definiteness of the matrix $K(u)$.

Some examples of finding Lagrangian bounds were solved for polynomials of degree greater than 6. For polynomial of eight degree

$$P_8(x) = x^8 - 76/3x^6 + 222x^4 - 756x^2,$$

which has two global minimum points: $x_1^* = \sqrt{3}$ and $x_2^* = -\sqrt{3}$, $P_8(x_1^*) = P_8(x_2^*) = -873$, the following results were obtained: the precision 10^{-6} was reached in 66 iterations of r -algorithms; the precision 10^{-10} was reached in 98 iterations. For the polynomial

$$P_8(x) = x^8 - 8x^7 + 112x^5 - 158x^4 - 392x^3 + 840x^2$$

having one global minimum point: $x_1^* = 5$, $P_8(x_1^*) = -11125$, the following results were obtained: for precision 10^{-11} the 105 iterations were needed. Some experiments for Legendre and Chebyshev polynomials were made. The Legendre polynomials were generated by recurrent formula:

$$P_{n+1}(x) = \frac{2n+1}{n+1}xP_n(x) - \frac{n}{n+1}P_{n-1}(x), \quad P_0(x) = 1, \quad P_1(x) = x$$

the Chebyshev polynomials were generated by formula:

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad T_0(x) = 1, \quad T_1(x) = x.$$

The results are given for even degrees $2m = 4, 6, 8, 10, 12$ and 14 in the table 2. Here *iter* – the number of iterations and *nf* – the number of calculations of function and subgradients in r -algorithm procedure, necessary for obtaining the defined precision $\varepsilon_x = 1.E - 6$. In columns 4 and 7 the deviations of Lagrangian bounds from global minimums are given.

(2) For the max-cut problem some numerical experiments were accomplished. For planar graph in the form of icosahedron (12 vertices and 30 edges) we use the edge weights from Table 3. At first we obtain dual

$2m$	Chebyshev polynomials			Legendre polynomials		
	$iter$	nf	$\psi_r^* - P^*$	$iter$	nf	$\psi_r^* - T^*$
4	21	31	.47684E-06	21	31	.44740E-06
6	35	71	.17072E-04	47	95	.72306E-08
8	70	129	.52532E-04	95	189	.11723E-09
10	124	268	.83521E-04	144	257	.15465E-10
12	183	377	.83487E-03	432	793	.43077E-13
14	234	458	.88431E-03	505	921	.34885E-12

Table 2. Results for polynomials Chebyshev and Legendre polynomials

No. edge	i	j	w_{ij}	No. edge	i	j	w_{ij}	No. edge	i	j	w_{ij}
1	1	2	20.0	11	3	9	18.0	21	7	8	18.0
2	1	3	30.0	12	3	10	15.0	22	7	9	27.0
3	1	4	40.0	13	4	5	32.0	23	7	10	36.0
4	1	5	50.0	14	4	10	24.0	24	7	11	45.0
5	1	6	60.0	15	4	11	20.0	25	7	12	54.0
6	2	3	16.0	16	5	6	40.0	26	8	9	14.0
7	2	6	48.0	17	5	11	30.0	27	8	12	42.0
8	2	8	12.0	18	5	12	25.0	28	9	10	21.0
9	2	9	10.0	19	6	8	30.0	29	10	11	28.0
10	3	4	24.0	20	6	12	36.0	30	11	12	35.0

Table 3. Edge weights for "icosahedron"

quadratic upper bound 668. Further by heuristic algorithm [18] which use optimal dual variables we obtain a feasible solution with cut value 642 ($V_1 = \{1, 2, 9, 10, 11, 12\}$; ($V_2 = \{3, 4, 5, 6, 7, 8\}$). For obtaining optimal cut we generate successively the triangle constraints (57). In the table 4 the vertices of triangle cycles are represented. After adding ten superfluous constraints in the form of triangles constraints, which were converted by substitution (56) in quadratic inequalities, we found the upper bound φ^* such that $\varphi^* - 642 < 1$. Thus, the found feasible solution gives the optimal cut. Note, that our result corresponds to theorem 4:

$$(668 - 642)/642 = 26/642 = 23/321 \approx 0.0405 \approx 4\%.$$

(3) Optimal bisection of graphs (BiSection). BiSection problem for graph $G(V, E)$ is formulated for even $|V|$ and corresponds to MaxCut problem under condition: the number of vertices in subsets V_1 and V_2 are equal.

No. cycle	i, j, k	φ^*	No. cycle	i, j, k	φ^*
1	6,8,12	665.53	6	5,6,12	647.06
2	1,5,6	661.94	7	2,6,8	646.72
3	1,4,5	657.13	8	5,11,12	646.41
4	7,11,12	652.82	9	7,8,12	645.14
5	7,9,10	650.66	10	7,10,11	642.90

Table 4. Adding of triangle cycles for "icosahedron"

We obtain the quadratic-type problem (44)–(45), with a new constraint

$$\sum_{i=1}^n x_i = 0, \quad (60)$$

or its quadratic analogue:

$$\left(\sum_{i=1}^n x_i \right)^2 = 0. \quad (61)$$

One may find the upper Lagrangian for BiSection in different ways: (a) using (60) for excluding x_n ; (b) using Lagrange multiplier u_{n+1} for account (61); (c) using penalty nonsmooth function with parameter s , where $s > \sum_{(i,j) \in E} |w_{ij}|$. Algorithm for solving "BiSection" by scheme (c) is similar to scheme of finding dual bounds for max-cut (see Section 6.). The difference is that instead of matrix $W(u)$ we use the matrix $\tilde{W}(u) = sJ_n + W(u)$, where J_n is the $n \times n$ matrix with all entries equal to 1.

The results of experiments are given in Table 6. For test problems we use the Petersen graph (10 vertices and 15 edges) with enumeration of vertices as in [34]. The values of edge weights for nine examples P_1, \dots, P_9 are given in Table 5.

We use the following parameters of r -algorithm: $\alpha = 2$, $h_0 = 1.0$, $q_1 = 0.9$, $n_h = 3$, $q_2 = 1.1$. We use $u_0 = \{1, 1, 1, 1, 1, 0, 0, 0, 0, 0\}$ as a starting point. r -algorithm stops if at the k -th iteration $\|u_k - u_{k-1}\| \leq \varepsilon_u = 10^{-3}$ or $\varphi_{up} - \varphi_{lo} < 1$. Here φ_{up} is an upper bound obtained at the current step, φ_{lo} is the value of cut obtained by heuristic procedure of finding "good" feasible cut (see Section 6.). This cut is shown in Table 6 by representation $V_1 \cup V_2$.

From Table 6 we see, that for 5 examples we obtain exact values of the problems after rounding of bounds, and the number of iterations was relatively small.

(i, j)	w_{ij}								
	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9
(1,2)	1	5	2	5	1	5	2	5	1
(1,5)	1	1	3	6	8	2	7	2	3
(1,10)	1	3	4	3	1	2	2	2	3
(2,3)	1	2	7	2	3	1	3	6	8
(2,9)	1	2	6	2	5	5	1	5	9
(3,4)	1	3	1	6	7	3	4	3	1
(3,8)	1	4	3	4	2	1	2	1	2
(4,5)	1	1	2	1	2	2	6	2	5
(4,7)	1	1	5	1	3	4	3	4	2
(5,6)	1	5	1	5	9	3	1	6	7
(6,8)	1	5	5	5	4	1	5	1	3
(6,9)	1	3	1	3	1	5	5	5	4
(7,9)	1	2	2	2	3	3	1	3	1
(7,10)	1	1	2	1	8	1	2	1	8
(8,10)	1	1	1	1	1	1	1	1	1

Table 5. Edge weights for Petersen graph

$Test$	$iter$	φ_{up}	φ_{lo}	V_1	V_2
P_1	41	12.50045	11.0	1,2,5,7,8	3,4,6,9,10
P_2	7	34.91024	34.0	2,3,6,7,10	1,4,5,8,9
P_3	2	40.84215	40.0	2,5,7,8,10	1,3,4,6,9
P_4	44	41.63351	40.0	1,3,6,7,10	2,4,5,8,9
P_5	35	52.52589	51.0	1,2,4,6,7	3,5,8,9,10
P_6	32	34.14082	33.0	1,4,5,8,9	2,3,6,7,10
P_7	17	39.99753	39.0	1,2,4,8,9	3,5,6,7,10
P_8	11	40.99564	40.0	2,4,6,8,10	1,3,5,7,9
P_9	7	52.92904	52.0	1,2,4,6,7	3,5,8,9,10

Table 6. BiSection for Petersen graph

Conclusion

As we have shown in this article the finding of Lagrangian bounds in polynomial multiextremal and Boolean extremal models can be reduced in many cases to the problems of nondifferentiable optimization with specific constraints in the form of semidefiniteness of some parametric families of symmetric matrices. These problems can be formulated also in the form of semidefinite programming (SDP) (see [29]) and solved by corresponding algorithms (for example, by interior point method [12]).

Since 1970, we solved by r -algorithms many hundreds of sophisticated tests and applied problems using r -algorithms. The results of Nemirovsky-Yudin on information complexity of convex programming algorithms [28] show that in general case one must make $O(n \log \frac{1}{\epsilon})$ measurements of function and subgradients in current points to guarantee the relative accuracy ϵ on minimized function value and discrepancy in constraints.

The results of testing of r -algorithms show that if the errors of rounding are not essential, the objective function values as a rule may be majored by a geometrical progression of the form $Cq^{\frac{k}{n}}$, where k is the number of current step and $q = \frac{1}{2}$. So, as a rule the convergence of r -algorithm approximately $2n$ times faster than of well known ellipsoid methods.

Our numerical experiments showed that application of nondifferentiable optimization models for obtaining Lagrangian bounds for multiextremal and combinatorial problems have some advantages in comparison with *SDP* methods: (a) the possibility to take in account specific structure of the problem and the use of decomposition schemes, if the problem has quasi-block structure; (b) simple ways for exchanging models in the case of adding new quadratic constraints; (c) possibility to use exact nonsmooth penalty functions; (d) rather fast rate of convergence and simplicity of r -algorithms.

In [8] the long history of finding good algorithms for solving some extremal mechanical problems posed by Lagrange in 1773 is described. Now these problems were reduced to nonsmooth optimization problems of maximization of the least eigenvalue of a self-adjoint fourth-order differential operator.

Only after more than 200 years good algorithms for solving such problems have been found and implemented. New algorithms were built on the base of last developments of convex analysis and study the differential properties of nonsmooth functions, particularly, of extremal eigenvalues of differential operators. Note, that similar problems arise when we want to find Lagrangian bounds for the combinatorial problem. As we have shown, the r -algorithms solve such problems successfully.

We want to emphasize that methods of nondifferentiable optimization must become necessary part of the courses of applied mathematics for mathematical and technical education.

Bibliography

1. ALIZADEH, F. (1995), Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM Journal on Optimization* 5, 13–51.
2. ANJOS, M.F. AND WOLKOWICZ, H. A strengthened SDP relaxation via second lifting for the max-cut problem, University of Waterloo, Department of Combinatorics and Optimization, *Research Report CORR 99-55*, 28 pp.
3. BARAHONA, F., GRÖTSHEL, M. AND MAHJOUR, A. (1985), Facets of the bipartite subgraph polytope. *Mathematics of Operations Research*, 10, 340–358.
4. BARAHONA, F. AND MAHJOUR, A. (1986), On the cut polytope. *Mathematics Programming*, 36, 157–173.
5. BERGE, C. (1961), Farbung von Graphen, deren samtliche bzw. deren ungerade Kreise starr sind (Zusammenfassung), *Wissenschaftliche Zeitschrift*, Martin Luther Universitat Halle-Wittenberg, Mathematisch-Naturwissenschaftliche Reihe, pp. 114–115,.
6. BERGE, C. (1962), Sur une conjecture relative au probleme des codes optimaux, *Communication*, 13eme assemblee generake de l'URSI, Tokyo.
7. BERTSEKAS, D.P. (1995), Nonlinear Programming, *Athena Scientific, Belmont, Massachusetts*, P. 646.
8. COX, S.J. (1992), The shape of the ideal column, *Mathematical Intelligencer*, 14, 16–31.
9. CULLUM, J., DONATH, W.E. AND WOLFE, P. (1975), The minimization of certain nondifferentiable sums of eigenvalues of symmetric matrices, *Mathematical Programming Study* 3, 35–55.
10. DELORME, C. AND POLJAK, S. (1993), Laplacian eigenvalues and the maximum cut problem, *Math. Programming* 62(3), 557–574.

11. DEMYANOV, V.F. AND RUBINOV, A.M. (1980), On quasidifferentiable functionals, *Soviet Mathematics Doklady* 21, 14–17.
12. DIKIN, I.I. AND ZORKALTSEV, V.I. (1980), Iterative solvig of mathematical programming problem (interior point method: algoritms), Nauka, Moscow, (in Russian).
13. DONATH, W.E. AND HOFFMAN, A.J. (1972), Algorithm for partitioning graphs and computer logic based on eigenvectors of connection matrices, *Math. Prog. Study* 3, 35–55.
14. DONATH, W.E. AND HOFFMAN, A.J. (1973), Lower bounds for the partitioning of graphs, *IBM J.Res.Dev.* 17.
15. FRIEZE, A. AND JERUM, M. (1995), Improved approximation algorithms for max k-cut and max bisection, *Reprint, Carnegie Mellon University*, pp. 1–17.
16. GOEMANS, M.X. (1997), Semidefinite programming in combinatorial optimization, *Math. Programming* 79, 143–162.
17. FAN K. (1949), On a theorem of Weyl concerning the eigenvalues of linear transformations, *Proceedings of the National Academy of the Sciences of U.S.A.* 35, 652–655.
18. GOEMANS, M.X. AND WILLIAMSON, D.P. (1994), 0.878-Approximation algorithms for MAX-CUT and MAX 2SAT, *Proceedings of the 26th Annual ACM Symposium on Theory of Computing*, 422–431.
19. GRÖTSCHEL, M. AND PULLEYBLANK, W. (1981), Weakly bipartite graphs and the max-cut problem. *Operations Research letters* 1, 23–27.
20. GRÖTSCHEL, M., LÓVASZ, L. AND SCHRIJVER A. (1988), *Geometric Algorithms and Combinatorial Optimization*, Vol. 2 *Combinatorics*. Springer, Berlin.
21. HILBERT D. (1888), Über die Darstellung definiten Formen als Summen von Formen quadraten *Math. Ann.*, Leipzig, Bd. 22, 342–350.
22. LAWRENT, M., POLJAK, S. AND RENDL, F. (1997), Connections between the semidefinite relaxations of the max-cut and stable set problems, *Math. Programming* 77, 225–246.
23. LEWIS, A.S. AND OVERTON, M.L. (1996), Eigenvalue Optimization, *Acta Numerica*, 5, 149–190.

24. LÓVASZ, L. (1979), On the Shannon-capacity of graph. *IEEE Transactions on information theory*. IT-25, 1, 1–7.
25. LÓVASZ, L. (1972), Normal hypergraphs and the perfect graph conjecture, *Discrete Mathematics* 2, 253–267.
26. McELICE, R., RODEMICH, E. AND ROMSEY, JR.H. (1978), The Lovash bound and some generalization. *Journal of combinatorics, Information and System Science*. 3(3), 134–152.
27. NARASIMHAN, G. AND MANBER, R. (1990), A generalization of Lovasz's sandwich theorem, *Polyhedral combinatorics: Proc. of a DIMACS Workshop AMS*.
28. NEMIROVSKY, A. AND YUDIN, D.B. (1983), Problem Complexity and Methods Efficiency in Optimization, *J. Wiley, New York*.
29. NESTEROV, YU. AND NEMIROVSKII, A. (1994), Interior Point Polynomial Algorithm in Convex Programming, *SIAM*, p. 406.
30. SHANNON, C. (1956), The zero-error capacity of a noisy channel. *IRe Transactions Information Theory*. IT, 2, 8–19.
31. SHRIJVER, A. (1979), A comparison of the Delsarte and Lovash bounds. *IEEE Transactions on information theory*. IT-25, 4.
32. SHOR, N.Z. (1970), Methods of non-differentiable optimization and its applications. *Ph.D.Thesis*, Institute of Cybernetics of NASU, Kiev, Ukraine.
33. SHOR, N.Z. (1998), *Nondifferentiable Optimization and Polynomial Problems*, Kluwer, Dordrecht.
34. SHOR, N.Z. AND BEREZOVSKI, O.A. (1995), New algorithms for solving weight max-cut problem, *Kibernetika i sistemny analiz* 2, 100–106.
35. SHOR, N.Z. AND STETSENKO, S.I. (1989), *Quadratic extremal problems and nondifferentiable optimization*, Naukova dumka, Kiev, (in Russian).
36. SHOR, N.Z. AND STETSYUK, P.I.(1997), Using of r -algorithm modifications for obtaining of global minimum of polynomial functions, *Kibernetika i sistemny analiz* 4, 28–48.
37. YANNAKAKIS, M. (1978), Node and edge deletion NP-complete problems. *Proc. 10th Ann. ACM Symp. on Theory of Computing*. New York, 253–264.

Список трудов Н. З. Шора

1961

1. МИХАЛЕВИЧ В.С., ШОР Н.З. *Про оптимальне регулювання марковської послідовності з двома фазовими станами* // Збірник праць обчислювальної математики і техніки. – К.: Ін-т кібернетики АН УРСР. – Т. 1.

2. МИХАЛЕВИЧ В.С., ШОР Н.З. *Метод последовательного анализа вариантов при решении вариационных задач управления, планирования и проектирования* // IV Всесоюзный математический съезд. Тезисы докладов. – Ленинград. – С. 91.

1962

3. БАКАЕВ О.О., БРАНОВИЦЬКА С.В., МИХАЛЕВИЧ В.С., ШОР Н.З. *Визначення характеристик транспортної сітки за методом послідовного аналізу варіантів* (Представив академік АН УРСР В.М.Глушков) // ДАН УРСР. – № 4. – С. 469–472.

4. БАКАЕВ А.А., БРАНОВИЦЬКА С.В., МИХАЛЕВИЧ В.С., ШОР Н.З. *Определение наиболее выгодных маршрутов по заданной сети дорог* // Заочный семинар: Кибернетика на транспорте. – К.: Дом научно-технической пропаганды.

5. БАКАЕВ А.А., БРАНОВИЦЬКА С.В., МИХАЛЕВИЧ В.С., ШОР Н.З. *Определение рациональных связей для решения транспортных задач большого объема* // Материалы Госэкономсовета СССР. СОПС. – Бюл. № 2.

6. МИХАЛЕВИЧ В.С., ШОР Н.З., ЛЯХОВСКИЙ В.Н. *Определение на ЭЦВМ наиболее выгодного положения красной линии продольного профиля на вольном ходу* // Транспортное строительство. – № 4.

7. МИХАЛЕВИЧ В.С., ШОР Н.З. *Численное решение многовариантных задач по методу последовательного анализа вариантов* // Научно-методические материалы семинара: О численных методах решения многовариантных плановых и технико-экономических задачах. – М.: Академия Наук СССР: Лаборатория по применению математических методов в экономических исследованиях и планировании, ВЦ АН УССР. – Вып. 1. – С. 15–48. – (Ротапринт).

8. МИХАЛЕВИЧ В.С., ШОР Н.З. *Математические вопросы оптимального размещения предприятий* // Материалы Госэкономсовета СССР. Лаборатория математических методов. – Ротапринт. – Бюл. № 2. – С. 15–42.

9. ШОР Н.З. *Закрытая*. – КВИРТУ.

10. ШОР Н.З. *Закрытая*. – КВИРТУ.

11. ШОР Н.З., ШКУРБА В.В. *Вероятностный расчет среднего времени выполнения арифметической операции на ЭЦВМ* // VI Всесоюзный совет по теории вероятности.

12. ШОР Н.З., БЕРЕСТОВЕНКО К.М. *Автоматизация тяговых расчетов с помощью ЭЦВМ* // Заочный семинар: Кибернетика на транспорте. – КДНТП.

13. ШОР Н.З. *Применение метода градиентного спуска для решения сетевой транспортной задачи*. В сб. Материалы научного семинара по теоретическим и прикладным вопросам кибернетики и исследования операций. – Киев, НС по кибернетике АН УССР. – Вып. 1. – С. 9–17.

14. ШОР Н.З., БЕРЕСТОВЕНКО К.М. *Автоматизация тяговых расчетов с помощью ЭЦВМ* // Заочный семинар: Кибернетика на транспорте. – К.: Дом научно-технической пропаганды.

1963

15. БАКАЕВ А.А., РОСИНА Н.И., ШОР Н.З. *Алгоритм определения грузонапряженности на звеньях сети путей сообщения* // Научно-методические материалы семинара по теоретическим и прикладным вопросам кибернетики и исследования операций. Экономическая кибернетика и исследование операций. НС по кибернетике АН УССР. – К.: Дом научно-технической пропаганды. – Вып. 4. – № 6. – С. 3–13.

16. ГАЛУСТОВА Л.А., ШОР Н.З. *Методика выбора оптимальных сечений сети 6-10 кв. методом последовательного анализа вариантов* // Научно-методические материалы семинара по теоретическим и прикладным вопросам кибернетики и исследования операций. Экономическая кибернетика и исследование операций. НС по кибернетике АН УССР. – К.: Дом научно-технической пропаганды. – Вып. 2. – С. 3–5.

17. БАКАЕВ А.А., МИХАЛЕВИЧ В.С., БРАНОВИЦКАЯ С.В., ШОР Н.З. *Методика и опыт решения сетевых транспортных задач большого объёма на ЭЦВМ* // Математические методы и проблемы производства. – М.: Экономиздат. – С. 247–257.

1964

18. МИХАЛЕВИЧ В.С., ШОР Н.З. *Математические основы решения задачи выбора оптимального очертания продольного профиля* // *Материалы ЦНИИС*. – М.: Транспорт. – С. 22–28.

19. ГАЛУСТОВА Л.А., ШОР Н.З. *Определение наимыгоднейшего варианта сети 35-10 кв. с проверкой на минимальный режим* // *Кибернетика и техника вычислений*. – К.: Наук. думка. – С. 144–147.

20. МИХАЛЕВИЧ В.С., ШОР Н.З. *Метод последовательного анализа вариантов для численного решения задач оптимизации* // *Материалы по вопросам применения ЭВМ в народном хозяйстве*. – Горький.

21. МИХАЛЕВИЧ В.С., ШОР Н.З. *Математические методы решения некоторых задач размещения* // *Применение математики при размещении произв. сил*. – К.

22. ШОР Н.З. *Оптимальный расчет электрических сетей с тремя ступенями напряжения с помощью ЭВМ* // *Материалы математической конференции по применению вычислительной техники в электроэнергетике*. – М. – № 2.

23. БЕШЕНИ М.И., МИХАЛЕВИЧ В.С., ШОР Н.З. *Выбор оптимальных вариантов сложных отводов от магистрального газопровода. Удостоверение о регистрации*. – № 48944. – (Изобретение)

24. ШОР Н.З. *О структуре алгоритмов численного решения задач оптимального планирования и проектирования*. Автореферат диссертации на соискание степени кандидата физ.-мат наук. – Объединенный Ученый Совет Института Математики, Кибернетики и Главной Астрономической Обсерватории АН УССР. – К. – 9 с.

25. ШОР Н.З. *О структуре алгоритмов численного решения задач оптимального планирования и проектирования*. Диссертация на соискание степени кандидата физ.-мат наук. – К.: Ин-т кибернетики АН УССР. – 135 с.

1965

26. МИХНОВСКИЙ С.Д., ШОР Н.З. *Оценка минимального числа пересылок при динамическом распределении страничной памяти* // *Кибернетика*. – № 5. – С. 18–20.

[English transl.: MIKHNOVSKII S.D. AND SHOR N.Z.]

An estimate of the minimum number of transfers for dynamic allocation of a page memory // *Cybernetics*. – V. 1. – № 5. – P. 16–18.

1966

27. ШОР Н.З. *Логическая структура* // Тезисы доклада на Международном Конгрессе математиков. Секция 13. – М.

1967

28. ШОР Н.З. *Об одном подходе к задачам многоцелевого планирования* // I Всесоюзная конференция по математическим вопросам сетевого планирования и управления. Тезисы и доклады. – К.: Ин-т кибернетики АН УССР. – С. 94–101.

29. ГАЛУСТОВА Л.А., ШОР Н.З. *Нахождение оптимальной конфигурации разомкнутых электрических сетей* // Экономическая кибернетика и исследование операций. – К.: Ин-т кибернетики АН УССР. – Вып. 2.

30. МИХАЛЕВИЧ В.С., ШОР Н.З., БИДУЛИНА Л.М. *Математические методы выбора оптимального варианта сложного магистрального газопровода при стационарном режиме течения газа* // Экономическая кибернетика и исследование операций. – К.: Ин-т кибернетики АН УССР. – Вып. 4. – С. 57–59.

31. ШОР Н.З., ДОНЕЦ Г.А. *Алгебраический подход к исследованию задачи о четырех красках* // Теория оптимальных решений. – К.: Ин-т кибернетики АН УССР. – Вып. 3. – С. 57–72.

32. ЦАЛЬ Р.Я., ШОР Н.З. *Численные методы расчета потоко-распределения в системах промышленной вентиляции* // Экономическая кибернетика и исследование операций. – К.: Ин-т кибернетики АН УССР. – С. 3–9.

33. ЕРМОЛЬЕВ Ю.М., ШОР Н.З. *О минимизации недифференцируемых функций* // Кибернетика. – № 1. – С. 101–102.

[English transl.: ERMOL'EV YU.M. AND SHOR N.Z. *On the minimization of non-differentiable functions* // Cybernetics. – V. 3. – № 1. – P. 72.

34. МИХАЛЕВИЧ В.С., ЕРМОЛЬЕВ Ю.М., ШКУРБА В.В., ШОР Н.З. *Сложные системы и решение экстремальных задач* // Кибернетика. – № 5. – С. 29–39.

[English transl.: MIKHALEVICH V.S., ERMOL'EV YU.M., SHKURBA V.V. AND SHOR N.Z. *Complex systems and the solution of extremal problems* // Cybernetics. – V. 3. – № 5. – P. 25–34.

35. ШОР Н.З. *Применение обобщенного градиентного спуска в блочном программировании* // Кибернетика. – № 3. – С. 53–55.

[English transl.: SHOR N.Z. *Application of generalized gradient descent in block programming* // Cybernetics. – V. 3. – № 3. – P. 43–45.

36. КОЗАЧКОВ Л.С., ХУРСИН Л.А., ШОР Н.З. *Количественный анализ некоторых тенденций в развитии исследования операций и математической экономики* // Кибернетика. – № 5. – С. 102–108.

[English transl.: KOZACHKOV L.S., HURSHIN L.A. AND SHOR N.Z. *Quantitative analysis of some trends in development of operations research and mathematical economics* // Cybernetics. – V. 3. – № 5. – P. 90–96.

37. ШОР Н.З., БАКАЕВ А.А. *Задачи планирования грузопотоков на транспортной сети* // Материалы МИИТ. – М.: Транспорт. – Вып. 237.

38. ШОР Н.З., РОСИНА Н.И. *Схема разложения задач линейного и выпуклого программирования и ее применение для решения задач планирования перевозок* // Доклады I Всесоюзной конференции по оптимизации и моделированию транспортных сетей. – К. – С. 225–237.

39. ШОР Н.З. *Многоэтапное стохастическое выпуклое программирование* // Теория оптимальных решений. – К.: Ин-т кибернетики АН УССР. – № 1. – С. 48–58.

40. ШОР Н.З., ЩЕПАКИН М.Б., АТУТОВА С.Я. *Алгоритмы решения задачи многоэтапного линейного стохастического программирования* // Теория оптимальных решений. – К.: Ин-т кибернетики АН УССР. – Вып. 1. – С. 73–86.

41. ШОР Н.З., ЩЕПАКИН М.Б. *Об одной задаче стохастического программирования* // Теория оптимальных решений. – К.: Ин-т кибернетики АН УССР. – № 2. – С. 38–44.

1968

42. ШОР Н.З., ГОРБАЧ Г.И. *Решение задач распределительного типа методом обобщенного градиентного спуска* // Теория оптимальных решений. – К.: Ин-т кибернетики АН УССР. – № 1. – С. 59–71.

43. ШОР Н.З. *О скорости сходимости обобщенного градиентного спуска* // Кибернетика. – № 3. – С. 98–99.

[English transl.: SHOR N.Z. *The rate of convergence of the generalized gradient descent method* // Cybernetics. – V. 4. – №3. – P. 79–80.

44. ШОР Н.З., РОСИНА Н.И. *Выбор оптимальной совокупности типовых элементов* // Теория автоматов и методы формализованного синтеза вычислительных машин и систем. – К. – Вып. 5.

45. ЕРМОЛЬЕВ Ю.М., ШОР Н.З. *Метод случайного поиска для задач двухэтапного стохастического программирования и его обобщение* // Кибернетика. – № 1. – С. 90–92.

[English transl.: ERMOL'EV YU. M. AND SHOR N. Z. *Method of random walk for the two-stage problem of stochastic programming and its generalization* // Cybernetics. – V. 1. – № 1. – P. 59–60.

46. ШОР Н.З., ЩЕПАКИН М.Б. *Алгоритм решения двухэтапной задачи стохастического программирования* // Кибернетика. – № 3. – С. 56–58.

[English transl.: SHOR N. Z. AND SCHEPAKIN M. B. *Algorithms for the solution of the two-stage problem in stochastic programming* // Cybernetics. – V. 4. – № 3. – P. 48–50.

1969

47. ШОР Н.З., БИЛЕЦКИЙ В.И. *Метод растяжения пространства для ускорения сходимости в задачах овражного типа* // Теория оптимальных решений. – К.: Ин-т кибернетики АН УССР. – Вып. 2. – С. 3–18.

48. ШОР Н.З., МАРАХОВСКИЙ Л.Ф. *Выбор оптимальной совокупности типовых элементов* // Теория автоматов и методы формализованного синтеза вычислительных машин и систем. – АН УРСР. НС по кибернетике. – К. – № 5. – С. 103–109.

49. ШОР Н.З., РОСИНА Н.И., СПЕКТОР Л.И. *Выбор оптимального диаметра нефтепровода с учетом поэтапного наращивания мощности* // Строительство трубопроводов. – М. – № 11.

50. ШОР Н.З., ЩЕПАКИН М.Б. *Задачи многоэтапного стохастического программирования в параметрической форме* // Теория оптимальных решений. – К.: Ин-т кибернетики АН УССР. – № 3. – С. 63–73.

51. ШОР Н.З., ИВАНОВА Л.В. *Об одном итеративном методе решения задач линейного программирования и матричных игр* // Теория оптимальных решений. – К.: Ин-т кибернетики АН УССР. – № 3. – С. 22–30.

52. ШОР Н.З. *Обобщенный градиентный спуск* // Труды первой зимней школы по математическому программированию (г. Дрогобыч). – М. – Вып. 3. – С. 578–585.

1970

53. ШОР Н.З., ЩЕПАКИН М.Б. *О некоторых подходах к решению задач многоэтапного стохастического программирования* // Материалы второй зимней школы по математическому программированию (г. Дрогобыч). – М. – Вып. 3. – С. 664–682.

54. ШОР Н.З., ЗЕМЛЯНУХИНА Л.И. *Некоторые комбинаторные задачи теории графов, связанные с максимальным внутренне устойчивым множеством* // Математические методы исследования и оптимизации систем. – К. – № 5. – С. 13–24.

55. ШОР Н.З. *Использование операции растяжения пространства в задачах минимизации выпуклых функций* // Кибернетика. – № 1. – С. 6–12.

[English transl.: SHOR N.Z. *Utilization of the operation of space dilation in the minimization of convex functions* // Cybernetics. – V. 6. – № 1. – P. 7–12.

56. ШОР Н.З. *Методы минимизации недифференцируемых функций и их приложения*. Автореферат диссертации на соискание степени доктора физ.-мат наук. – Ин-т кибернетики АН УССР. – К. – 43 с.

57. ШОР Н.З. *Методы минимизации недифференцируемых функций и их приложения*. Диссертация на соискание степени доктора физ.-мат наук. – К.: Ин-т кибернетики АН УССР. – 203 с.

58. ШОР Н.З. *О скорости сходимости метода обобщенного градиентного спуска с растяжением пространства* // Кибернетика. – № 2. – С. 80–85.

[English transl.: SHOR N.Z. *Convergence rate of the gradient descent method with dilation of the space* // Cybernetics. – V. 6. – № 2. – P. 80–85.

59. ШОР Н.З., КАРМИНСКИЙ Б.Б., ПОЛЯК И.М. *Математическая модель и машинный алгоритм оптимального планирования загрузки станков производством проката* // Управление в черной металлургии. – М.: Металлургия. – Вып. XII. – С. 15–42.

1971

60. ШОР Н.З., ЖУРБЕНКО Н.Г. *Метод минимизации, использующий операцию растяжения пространства в направлении разности двух последовательных градиентов* // Кибернетика. – 1971. – № 3. – С. 51–59.

[English transl.: SHOR N.Z. AND ZHURBENKO N.G. *A minimization method using the operation of extension of the space in the direction of*

the difference of two successive gradients // Cybernetics. – V. 7. – № 3. – P. 450–459.

61. ШОР Н.З., ГАМБУРД П.Р. *Некоторые вопросы сходимости обобщенного градиентного спуска* // Кибернетика. – № 6. – С. 82–84.

[English transl.: SHOR N.Z. AND GAMBURD P.R. *Certain Questions of Convergence of Generalized Gradient Descent* // Cybernetics. – V. 7. – № 6. – P. 1033–1036.

1972

62. МИХАЛЕВИЧ В.С., ГАЛУСТОВА Л.А., ШОР Н.З., МОМОТ А.И., СТРУТИНСКАЯ С.П. *Программа и инструкция по выбору оптимальной конфигурации электрических сетей* – К.: Ин-т кибернетики АН УССР, НС по кибернетике. – 79 с.

63. ШОР Н.З., ШАБАШОВА Л.П. *О решении минимаксных задач методом обобщенного градиентного спуска с растяжением пространства* // Кибернетика. – № 1. – С. 82–88.

[English transl.: SHOR N.Z. AND SHABASHOVA L.P. *Solution of minimax problems by the method of generalized gradient descent with dilation of the space* // Cybernetics. – V. 4. – № 1. – P. 88–94.

64. ШОР Н.З., ШАБАШОВА Л.П. *Алгоритм и численные эксперименты решения задач минимизации с использованием операции растяжения пространства* // Математика и механика. – Днепропетровск, ДГУ.

65. ШОР Н.З. *О классе почти-дифференцируемых функций и одном методе минимизации функций этого класса* // Кибернетика. – № 4. – С. 65–70.

[English transl.: SHOR N.Z. *A class of almost differentiable functions and a minimization method for functions of this class* // Cybernetics. – V. 8. – № 4. – P. 599–606.

66. КУКСА А.И., ШОР Н.З. *О методе оценки количества условно-оптимальных траекторий дискретного сепарабельного динамического программирования* // Кибернетика. – № 6. – С. 37–44.

[English transl.: KUKSA A.I. AND SHOR N.Z. *The Method of evaluating the Number of Conditionally Optimal Trajectories of Discrete Separable Dynamic Programming* // Cybernetics. – V. 8. – № 6. – P. 910–919.

67. ШОР Н.З., КАВЕЛИН С.С., ЮН Г.Н., ВДОВИЧЕНКО И.А., ВАСИЛЬЕВ В.Г., ГОРДЕЕВ Г.М., СИБИРКО А.Н. // *Закрытая*. – К.

68. ШОР Н.З., ЮН Г.Н., ВДОВИЧЕНКО И., ВАСИЛЬЕВ В.Г., ДУДАНКИН В.И. // *Закрытая*. – К.

69. ШОР Н.З., СТРУТИНСКАЯ С.П., МОМОТ А.И. *Оптимальное потоко-распределение в Единой газоснабжающей системе* // Применение методов в экономических исследованиях и планировании. – К.: Ин-т кибернетики АН УССР. – С. 3–21.

1973

70. ШОР Н.З., ЩЕПАКИН М.Б. *Об одном способе регулировки шага в алгоритмах, использующих случайный поиск* // Теория оптимальных решений. – К.: Ин-т кибернетики АН УССР. – С. 68–74.

1974

71. ШОР Н.З., ЩЕПАКИН М.Б. *Метод случайного поиска с растяжением пространства* // Теория оптимальных решений. – К.: Ин-т кибернетики АН УССР. – С. 25–27.

72. ШОР Н.З., ЩЕПАКИН М.Б. *Об оценке скорости сходимости метода случайного поиска* // Кибернетика. – № 4. – С. 55–58.

SHOR N.Z. AND SHCHERAKIN M.B. *Estimating the Rate of Convergence of the Random Search Method* // Cybernetics. – V. 10. – № 4. – P. 615–618.

73. ВДОВИЧЕНКО И.А., ТРУБИН В.А., ШОР Н.З., ЮН Г.Н. *Задачи типизации технических объектов* // Кибернетика. – № 6. – С. 106–108.

[English transl.: VDOVICHENKO I.A., TRUBIN V.A., SHOR N.Z. AND YUN G.N. *Problems of Typification of Technological objects* // Cybernetics. – V. 10. – № 6. – P. 1032–1034.

74. ШОР Н.З., ЮН Г.Н., ТИТКОВ Б.М., ПИНАЕВ Е.Г. *Математические модели, методы и алгоритмы решения некоторых задач программного планирования развития гражданской авиационной техники* // Материалы Всесоюзной НТК: Эффективность и оптимизация систем и процессов гражданской авиации. – М.

75. ШОР Н.З. *Беллмана принцип оптимальности* // Энциклопедия кибернетики. – К. – Т. 1.

76. SHOR N.Z. *Беллмана уравнение* // Энциклопедия кибернетики. – К. – Т. 1.

77. ШОР Н.З., ТРУБИН В.А. *Двойственности теория в программировании линейном* // Энциклопедия кибернетики. – К. – Т. 1.

78. ШОР Н.З. *Дробно-линейного программирования задача* // Энциклопедия кибернетики. – К. – Т. 1.

79. ШОР Н.З. *Задача о кратчайшем пути* // Энциклопедия кибернетики. – К. – Т. 1.
80. МИХАЛЕВИЧ В.С., ШОР Н.З., ИВАНОВ В.В. *Минимизации функций методы* // Энциклопедия кибернетики. – К. – Т. II.
81. МИХАЛЕВИЧ В.С., ШОР Н.З. *Последовательный анализ вариантов* // Энциклопедия кибернетики. – К. – Т. II.
82. ШОР Н.З. *Обобщенных градиентов метод* // Энциклопедия кибернетики. – К. – Т. II.

1975

83. ШОР Н.З., БИЛЕЦКИЙ В.И. *Применение метода обобщенного градиентного спуска с растяжением пространства в блочном программировании* // Теория оптимальных решений. – К.: Ин-т кибернетики АН УССР. – № 1. – С. 3–9.
84. ШОР Н.З., ЮН Г.Н., ТАРАСЕНКО В.П., ГОРДЕЕВ Г.М. // *Закрытая.* – К.
85. ШОР Н.З., ЮН Г.Н., ТОКАРЕВА О.Н., РЖЕВСКИЙ С.В. // *Закрытая.* – К.
86. ШОР Н.З. *Исследование сходимости метода градиентного типа с растяжением пространства в направлении разности двух последовательных субградиентов* // Кибернетика. – № 4. – С. 48–53.
 [English transl.: SHOR N.Z. *Convergence of a gradient method with space dilation in the direction of the difference between two successive gradients* // Cybernetics. – V. 11. – № 4. – P. 564–570.
87. ЛЯШЕНКО И.Н., КАРАГОДОВА Е.А., ЧЕРНИКОВА Н.Е., ШОР Н.З. *Линейное и нелинейное программирование.* – К.: Вища школа. – 372 с.
88. ШОР Н.З., ЮН Г.Н., ПИНАЕВ Е.Г. *О некоторых общесистемных аспектах оптимального проектирования гражданских самолетов* // Материалы Всесоюзной НТК. Автоматизация несущей способности и длительной ЛА. Тезисы докладов. – Харьков.
89. ШОР Н.З., ВЛАСЕНКО Ю.В., ПРОСКУРИН Е.А., ТРАЙНИН Э.З. *К вопросу о выборе метода решения одного класса задач линейного программирования* // Разработка математических и технических средств автоматизированных систем управления. – К.: Ин-т кибернетики АН УССР. – С. 61–68.

90. ШОР Н.З. *Обобщенные градиентные методы минимизации негладких функций и их применение к задачам математического программирования* // Экономика и математические методы. – Вып. 2. – № 12. – С. 337–356.

91. МИХАЛЕВИЧ В.С., ШОР Н.З., ТОКАРЕВА О.Н., ЮН Г.Н., РЖЕВСКИЙ С.В. *К вопросу оптимального проектирования источников питания технических объектов* – К. – 17 с. – (Препринт // АН УССР. Ин-т кибернетики АН УССР. – С. 76–56).

92. ШОР Н.З., ЩЕПАКИН М.Б. *Некоторые задачи приближения функциями с ограничениями на первую и вторую производные* // Теория оптимальных решений. – К.: Ин-т кибернетики им. В.М. Глушкова АН УССР. – С. 11–20.

93. ШОР Н.З., ПАК С.В. *Комплекс задач оптимизации и размещения самолетного парка для ПАНХ* // Задачи и процедуры формирования перспективного самолет-вертолет. парка ПАНХ. – К. – С. 19–22.

94. БЕЛЯЕВА Л.В., ЖУРБЕНКО Н.Г., ШОР Н.З. *Об одной нелинейной транспортной задаче* // Теория оптимальных решений. – К.: Ин-т кибернетики АН УССР. – С. 83–89.

95. ОВРУЦКИЙ И.Г., ШОР Н.З. *Применение методов минимизации негладких функций для решения задач интерпретации гравиметрических наблюдений* // Кибернетика. – № 2. – С. 57–64.

[English transl.: OVRUTSKII I.G. AND SHOR N.Z. *Application of methods of minimization of nonsmooth functions to the interpretation of gravimetric observations* // Cybernetics. – V. 12. – № 2. – P. 232–239.

96. ЖУРБЕНКО Н.Г., ПИНАЕВ Е.Г., ШОР Н.З., ЮН Г.Н. *Выбор состава и распределение гражданских самолетов по авиалиниям* // Кибернетика. – № 4. – С. 138–141.

[English transl.: ZHURBENKO N.G., PINAEV E.G., SHOR N.Z. AND YUN G.N. *Choice of fleet composition and allocation of aircraft to civil airline routes* // Cybernetics. – V. 12. – № 4. – P. 636–640.

97. ШОР Н.З., ТОКАРЕВА О.Н., ЮН Г.Н. и др. *К вопросу оптимизации углов установки панелей неориентируемой солнечной батареи с учетом температуры* // Материалы V Гагаринских чтений. – М.: ВИАИ.

98. ШОР Н.З., ТОКАРЕВА О.Н., ЮН Г.Н. и др. *О нахождении оптимальной ориентации панелей солнечной батареи несимметричной формы* – К.: Ин-т кибернетики АН УССР. (Препринт 76-56: Вопросы оптимального проектирования технических объектов.)

99. ШОР Н.З., ЮН Г.Н., АНДРУСЕНКО С.К., КАСЬЯНЕНКО В. *Динамическая модель формирования перспективного планирования парка гражданских самолетов* // Кибернетика. – № 5. – С. 149–150.

1977

100. ШОР Н.З. *Метод отсечения с растяжением пространства для решения задач выпуклого программирования* // Кибернетика. – № 1. – С. 94–95.

[English transl.: SHOR N.Z. *Cut-off method with space extension in convex programming problems* // Cybernetics. – V. 13. – № 1. – P. 94–96.

101. ШОР Н.З., РЖЕВСКИЙ С.В., ТОКАРЕВА О.Н., ЮН Г.Н. *Об одной задаче оптимизации углов установки панелей солнечной батареи выпуклой формы* // Ученые записки ЦАГИ. – М. – Т. 8. – № 3.

102. МИХАЛЕВИЧ В.С., ШОР Н.З., ГАЛУСТОВА Л.А., ЖУРБЕНКО Н.Г., МОМОТ А.И., СИБИРКО А.Т., ТРУБИН В.А., ЮН Г.Н. *Вычислительные методы выбора оптимальных проектных решений*. – К.: Наук. думка. – 178 с.

103. ШОР Н.З. *Новые направления в развитии методов негладкой оптимизации* // Кибернетика. – № 6. – С. 87–91.

[English transl.: SHOR N.Z. *New development trends in nondifferentiable optimization* // Cybernetics. – V. 13. – № 6. – P. 881–886.

1978

104. БЕЛЯЕВА Л.В., ЖУРБЕНКО Н.Г., ШОР Н.З. *О методе решения одного класса динамических распределительных задач* // Экономика и матем. методы. – Т. 14. – Вып. 1. – С. 137–146.

105. ШОР Н.З., ГАЛУСТОВА Л.А., МОМОТ А.И. *Применение математических методов при оптимальном проектировании Единой газоснабжающей системы с учетом динамики ее развития* // Кибернетика. – № 1. – С. 69–74.

[English transl.: SHOR N.Z., GALUSTOVA L.A. AND MOMOT A.I. *Application of mathematical methods for optimal design of a single gas supply system considering the dynamics of its development* // Cybernetics. – V. 14. – № 1. – P. 68–73.

106. ШОР Н.З., ГАЛУСТОВА Л.А., РОЗЕН Н.Е. *Алгоритм оптимизации для решения задач развития Единой системы газоснабжения* // Газовая промышленность. – № 11.

107. ШОР Н.З., АСЕЛЬДЕРОВА И.М. *Исследование одной элементарной задачи преследования на сети* // Кибернетика. – № 4. – С. 66–72.

[English transl.: SHOR N.Z. AND ASELDEROVA I.M. *Investigation of one elementary problem of pursuit on a grid* // Cybernetics. – V. 14. – № 4. – P. 550–555.

108. МИХАЛЕВИЧ В.С., САКАЧ Р.В., ШОР Н.З., ПИНАЕВ Е.Г. *Совершенствование системы комплексного планирования развития гражданской авиации на основе экономико-математических моделей* – К. – 20 с. – (Препринт // АН УССР. Ин-т кибернетики: 78–22).

1979

109. ШОР Н.З. *Методы минимизации недифференцируемых функций и их приложения* – К.: Наук. думка. – 199 с.

[English transl.: SHOR N.Z. *Minimization Methods for Non-Differentiable Functions* – Berlin: Springer-Verlag. – 199 p.

110. МИХАЛЕВИЧ В.С., СЕРГИЕНКО И.В., БЫКОВ В.Н., КУКСА А.И., ЛЕБЕДЕВА Т.Т., РОЦИН В.А., СТУКАЛО А.С., ТРУБИН В.А., ШОР Н.З. *О математическом обеспечении пакета программ ДИС-ПРО для решения задач дискретного программирования* – К. – 57 с. – (Препринт: 79-3 // Ин-т кибернетики АН УССР).

111. ШОР Н.З., ЛЕЙЧКИС И.М., ЩЕПАКИН М.Б. *Влияние концентрации примесей разделяющей суспензии на величину оптимальных параметров при фильтровании с применением вспомогательных веществ* // Химическая промышленность. – № 1.

112. ШОР Н.З., ГЕРШОВИЧ В.И. *Об одном семействе алгоритмов для решения задач выпуклого программирования* // Кибернетика. – № 4. – С. 62–67.

[English transl.: SHOR N.Z. AND GERSHOVICH V.I. *Family of algorithms for solving convex programming problems* // Cybernetics. – V. 15. – № 4. – P. 502–508.

113. ШОР Н.З., ИВАНОВ В.В., ЛЮДВИЧЕНКО В., МИХАЛЕВИЧ В.С. *Вопросы повышения эффективности алгоритмов минимизации функций и математического программирования* – К. – (Препринт // АН УССР. Ин-т кибернетики: 79-59).

1980

114. Михалевич В.С., Шор Н.З., Зайцев Р.В., Билецкий В.И. и др. *О комплексе задач оптимизации проектных решений по профилю сложных участков дорог (на примере БАМ)* – К. – 46 с. – (Препринт // АН УССР. Ин-т кибернетики: 80-29).

115. Гладышевская Г.Н., Косиченко Е.Ф., Шор Н.З., Щепакин М.Б., Юн Г.Н. *Об одном способе прогнозирования экономических показателей развития отрасли (на примере гражданской авиации)* // Теория оптимальных решений. – К.: Ин-т кибернетики АН УССР. – С. 11–20.

116. Шор Н.З. *О возможности создания практически эффективных процедур решения задач линейного программирования на базе методов недифференцируемой оптимизации* // VI Всесоюзный симпозиум программного обеспечения решения задач оптимального планирования. Тезисы докладов. – М.

117. Михалевич В.С., Шор Н.З., Зайцев Р.В. и др. *О диалоговом проектировании профиля сложных участков железных дорог* // 2-я школа-семинар: Интерактивные системы. Тезисы докладов. – Тбилиси: МецМИЭРЭба.

1981

118. Шор Н.З., Галустова Л.А., Константинова И.М., Атутова С.Я. *Выбор структуры развивающихся газопроводов с учетом ограничения ресурсов* // В кн.: Вычислительные аспекты в пакетах программ и опыт решения оптимизационных задач. – К.: Ин-т кибернетики АН УССР. – С. 3–7.

119. Шор Н.З., Галустова Л.А. *Оптимизация этапов развития ГТС*. – М.: ГВЦ Мингазпрома. – Отраслевой фонд. – № 00321.

120. Михалевич В.С., Сергиенко И.В., Лебедева Т.Т., Рощин В.А., Стукало А.С., Трубин В.А., Шор Н.З. *Пакет прикладных программ ДИСПРО, предназначенных для решения задач дискретного программирования* // Кибернетика. – № 3. – С. 117–137.

[English transl.: MIKHALEVICH V.S., SERGIENKO I.V., LEBEDEVA T.T., ROSCHIN V.A., STUKALO A.S., TRUBIN V.A. AND SHOR N.Z. *DISPRO: A discrete programming package* // Cybernetics. – V. 17. – № 3. – P. 416–437.

121. МИХАЛЕВИЧ В.С., СЕРГИЕНКО И.В., ШОР Н.З. *Исследование методов решения оптимизационных задач и их приложения* // Кибернетика. – № 4. – С. 89–113.

[English transl.: MIKHALEVICH V.S., SERGIENKO I.V. AND SHOR N.Z. *Investigation of optimization methods and their applications* // Cybernetics. – V. 17. – № 4. – P. 522–548.

122. ШОР Н.З., ГЕРШОВИЧ В.И. *Об одной модификации алгоритмов градиентного типа с растяжением пространства для решения задач большой размерности* // Кибернетика. – № 5. – С. 67–70.

[English transl.: SHOR N.Z. AND GERSHOVICH V.I. *Solution of problems of high dimensionality of modified gradient minimization with space expansion* // Cybernetics. – V. 17. – № 5. – P. 642–645.

123. ШОР Н.З., ЮН Г.Н., АНДРУСЕНКО С. *Динамическая модель формирования самолетно-вертолетного парка гражданской авиации* // Тезисы. – М.: ММИГА.

124. ШОР Н.З., РЖЕВСКИЙ С.В. *ϵ -субградиентный алгоритм безусловной минимизации выпуклой необязательно дифференцируемой функции* // ОФАП МОМ. – № 1025. – Инв. 307П государственной регистрации.

1982

125. ШОР Н.З., ГЕРШОВИЧ В.И. *Метод эллипсоидов, его обобщения и приложения* // Кибернетика. – № 5. – С. 61–69.

126. ШОР Н.З., ЮН Г.Н., АНДРУСЕНКО С., ЕМЕЛЬЯНОВ В. *Динамическая модель формирования перспективного парка пассажирских самолетов с учетом ресурсных ограничений.* – К. – Препринт // АН УССР. Ин-т кибернетики им. В.М.Глушкова: 82-3.

127. ШОР Н.З., ДОНЕЦ Г.А. *Алгебраический подход в проблеме раскраски плоских графов.* – К.: Наук. думка. – 143 с.

128. SHOR N.Z. *Generalized gradient methods of nondifferentiable optimization using space expansion operations* // Proceedings of XI International symposium on Mathematical programming. – Bonn.

129. БАКАЕВ А.А., ШОР Н.З., ЯНОВЧИК В.А. *Стохастические модели оптимального планирования производства с учетом возможности внешних поставок и методы их реализации* // Математическое моделирование и применение ЭВМ в экономических исследованиях. – К.: Ин-т кибернетики им. В.М. Глушкова АН УССР. – С. 3–10.

1983

130. ЛЕБЕДЕВА Т.Т., МИХАЛЕВИЧ В.С., ШОР Н.З. *Структура, состав и назначение пакета программ ДИСПРО для решения задач дискретной оптимизации* // Техническая кибернетика. – М.: Известия АН СССР. – № 1.

131. КАНТОРОВИЧ Л.В., МИХАЛЕВИЧ В.С., ШОР Н.З. и др. *XI Международный симпозиум по математическому программированию* // Техническая кибернетика. – Обзор. – М.: Изв. АН СССР. – № 1. – С. 197–201.

132. МИХАЛЕВИЧ В.С., СЕРГИЕНКО И.В., ТРУБИН В.А., ШОР Н.З. и др. *Пакет прикладных программ для решения задач производственно-транспортного планирования большой размерности (ПЛАНЕР)* // Кибернетика. – № 3. – С. 57–71.

133. БЕЛЯЕВА Л.В., БИЛЕЦКИЙ В.И., ШОР Н.З. *О декомпозиционном алгоритме выбора оптимального профиля железной дороги* // Кибернетика. – № 3. – С. 76–79.

134. ШОР Н.З., ДОНЕЦ Г.А. *О задаче четырех красок* // Математика сегодня. – К.: Вища школа.

1984

135. МИХАЛЕВИЧ В.С., СЕРГИЕНКО И.В., ШОР Н.З., ЖУРБЕНКО Н.Г. и др. *Пакет прикладных программ для решения задач производственно-транспортного планирования большой размерности* // ППП. Методы оптимизации. – М.: Наука. – С. 66–84.

136. ШОР Н.З., БУРЧАК О.Т., ГЕРШОВИЧ В.И. *Алгоритмы негладкой оптимизации и параллельные вычисления* // Развитие теории многопроцессорных систем НС по кибернетике. – К.: Ин-т кибернетики им. В.М.Глушкова АН УССР. – С. 49–55.

137. СТЕЦЕНКО С.И., ШОР Н.З. *Связь оценок Ловаса с двойственными оценками в квадратичных булевых задачах* // Методы решения задач нелинейного и дискретного программирования. – К.: Ин-т кибернетики им. В.М.Глушкова АН УССР. – С. 20–25.

138. ШОР Н.З., СТЕЦЕНКО С.И. *Получение оценок хроматического класса графов с использованием моделей булева и квадратичного программирования* // Тезисы Республиканского семинара по дискретной оптимизации (г.Ужгород). – К.: Ин-т кибернетики им. В.М.Глушкова АН УССР. – С. 135–136.

139. КИСЕЛЁВА Е.М., ШОР Н.З. *Алгоритм решения многопродуктовой задачи оптимального разбиения с ограничениями* // Кибернетика. – № 1. – С. 76–81.

[English transl.: KISELEVA E.M. AND SHOR N.Z. *An algorithm of solution of multiproduct problem of optimal partitioning with constraints* // Cybernetics. – V. 21. – № 1. – P. 90–99.

140. SHOR N.Z. *Minimization Methods for Non-Differentiable Functions* – Berlin: Springer-Verlag. – 178 p.

141. ШОР Н.З., МИХАЛЕВИЧ В.С., СЕРГИЕНКО И.В., ТРУБИН В.А. и др. *Пакет прикладных программ для решения задач дискретной оптимизации ДИСПРО-3*. – К.: Ин-т кибернетики им. В.М. Глушкова АН УССР (Укр. РФАП). – Рег. № ПОО21. – 538 с.

142. МИХАЛЕВИЧ В.С., СЕРГИЕНКО И.В., ШОР Н.З., ТРУБИН В.А. и др. *Пакет прикладных программ ДИСПРО-3: назначение, классы решаемых задач, системное и алгоритмическое обеспечение* // Кибернетика. – № 1. – С. 56–71.

[English transl.: MIKHALEVICH V.S., SERGIENKO I.V., SHOR N.Z., TRUBIN V.A. ET AL. *The software package ДИСПРО-3: objectives, classes of problems, systemic and algorithmic support* // Cybernetics. – V. 21. – № 1. – P. 68–86.

143. SHOR N.Z. AND STECENKO S.I. *Quadratic boolean problems and Lovasz bounds* // Abstracts of 12th JFIP conf. on system modelling and optimization. – Budapest. – P. 302–304.

144. ШОР Н.З., АЛИЕВ Н.Т. *Об одной монотонной модификации r-алгоритма* // Методы решения сложных задач математического программирования – К.: Ин-т кибернетики им. В.М.Глушкова НАН Украины. – С. 45–50.

145. ШОР Н.З., ДАВЫДОВ А.С. *О методе получения оценок в квадратичных экстремальных задачах с булевыми переменными* // Кибернетика. – № 2. – С. 48–50.

[English transl.: SHOR N.Z. AND DAVYDOV A.S. *Method of obtaining estimates in quadratic extremal problems with Boolean variables* // Cybernetics and systems analysis. – V. 21. – № 2. – P. 207–211.

1986

146. МИХАЛЕВИЧ В.С., ТРУБИН В.А., ШОР Н.З. *Оптимизационные задачи производственно-транспортного планирования. Модели, методы, алгоритмы.* – М.: Наука. – 264 с.

147. ШОР Н.З., СТЕЦЕНКО С.И. *Экстремальные спектральные задачи на классах симметричных матриц и комбинаторные проблемы* // Кибернетика и вычислительная техника. – К.: Ин-т кибернетики им. В.М.Глушкова АН УССР. – Вып. 69. – С. 8–15.

148. ШОР Н.З., ЗЕЛЕНИН В.А. *Прогнозирование показателей разработки нефтяных месторождений на основе недифференцируемой оптимизации* // Математические модели планирования и управления сложными экономическими системами. – К.: Ин-т кибернетики им. В.М. Глушкова АН УССР. – С. 4–7.

149. ИВАНОВ Н.И., БРЕСЛАВЦЕВ А., ХИЖНЯК В.И., ШОР Н.З. и ДР. *Целевые комплексные программы развития производства (методические вопросы).* – К.: Наук. думка.

150. МИХАЛЕВИЧ В.С., ШОР Н.З., КАПИТОНОВА Ю., МОЛЧАНОВ И.Н. *Численные методы для многопроцессорного вычислительного комплекса ЕС.* – М.: ВВИА им. Н.Е.Жуковского.

151. МИХАЛЕВИЧ В.С., ТРУБИН В.А., ШОР Н.З. *Математические методы решения задач производственно-транспортного планирования. Модели, методы, алгоритмы.* – М.: Наука, – 240 с.

1987

152. ШОР Н.З., АЛИЕВ Н.Т. *Об одном алгоритме негладкой оптимизации* // Тезисы конференции: Методы математического программирования и программного обеспечения. – Свердловск.

153. ШОР Н.З. *Квадратичные оптимизационные задачи* // Техническая кибернетика. – К.: Ин-т кибернетики им. В.М. Глушкова АН УССР. – № 1. – С. 128–139.

154. ШОР Н.З. *Об одном подходе к получению глобальных экстремумов в полиномиальных задачах математического программирования* // Кибернетика. – № 5. – С. 102–106.

[English transl.: SHOR N.Z. *An approach to obtaining Global Extremum in Polynomial Mathematical Programming Problems* // Cybernetics and systems analysis. – V. 23. – № 5. – P. 695–701.

155. ШОР Н.З. *Об одном классе оценок глобального минимума полиномиальных функций* // Кибернетика. – № 6. – С. 9–11.

[English transl.: SHOR N.Z. *Class of Global Minimum bounds of polynomial Functions* // Cybernetics and systems analysis. – V. 23. – № 6. – P. 731–734.

1988

156. SHOR N.Z. AND STETSENKO S.I. *Optimization subject to positive semi-definite constraints in control problems* // 7th IFAC Workshop on Control nonlinear programming and optimization. – Tbilisi.

1989

157. ШОР Н.З., СТЕЦЕНКО С.И. *Квадратичные экстремальные задачи и недифференцируемая оптимизация*. – К.: Наук. думка. – 208 с.

158. ШОР Н.З., СОЛОМОН Д.И. *Декомпозиционные методы в дробно-линейном программировании*. – Кишинев: ШТИИЦа. – 204 с.

159. SHOR N.Z. *Nonsmooth optimization and dual bounds* // Nonsmooth Optimization and related topics. – New York, London: Plenum Press.

160. ШОР Н.З., СТЕЦЕНКО С.И. *Использование точных штрафов при построении описанных эллипсоидов минимального объёма* // Кибернетика. – № 2. – С. 117–119.

161. SHOR N.Z., STETSENKO S.I., BEREZOVSKI O.A. *Algorithms of constructing optimal inscribed and circumscribed ellipsoids* // Abstracts of 14th JFIP conference: System modelling and optimization. – Leipzig. – Heft 3.

162. ШОР Н.З. *Квадратичные оценки в полиномиальных задачах оптимизации* // Тезисы докладов Международной школы-семинара по методам оптимизации и их приложениям. – Иркутск.

163. ШОР Н.З., ЖУРБЕНКО Н.Г., КУНЦЕВИЧ А.В. *Об одном алгоритме планирования производственной программы объединения* // Кибернетика. – № 5. – С. 54–58.

[English transl.: SHOR N.Z., ZHURBENKO N.G. AND KUNTSEVICH A.V. *A production planning problem for an industrial association* // Cybernetics and systems analysis. – V. 25. – № 5. – P. 629–634.

164. ШОР Н.З., БЕРЕЗОВСКИЙ О.А. *Использование алгоритма субградиентного типа с растяжением пространства для построения эллипсоида максимального объёма, вписанного в многогранник* // Кибернетика. – № 6. – С. 119–120.

165. ШОР Н.З., ДАВЫДОВ А.С., СТЕЦЕНКО С.И. *Алгоритмы построения оптимальных описанных эллипсоидов на основе методов*

негладкой оптимизации // Методы решения экстремальных задач и смежные вопросы. – К.: Ин-т кибернетики им. В.М.Глушкова АН УССР.

1990

166. ШОР Н.З., СТЕЦЕНКО С.И. *Алгоритм последовательного сжатия пространства для построения описанного эллипсоида минимального объёма* // Исследование методов решения экстремальных задач. – АН УССР. Ин-т кибернетики им. В.М.Глушкова. НС по проблеме "Кибернетика". – К. – С. 25–29.

167. ШОР Н.З. *Перманенты и проблема раскраски плоских графов* // Перманенты: теория и приложения. – Красноярск.

168. ШОР Н.З., СТЕЦЕНКО С.И. *Исследование методов решения экстремальных задач* // Информационные технологии в научных исследованиях и испытаниях. – К.: Ин-т кибернетики им. В.М.Глушкова АН УССР. – С. 25-30.

169. SHOR N.Z. *Dual quadratic estimates in polynomial and boolean programming* // Annals of Operations Research. – 25. – P. 163–168.

170. SHOR N.Z. *Polynomial mathematical programming problems and nondifferentiable optimization* // Abstracts of France-Soviet workshop on mathematical programming. – Luminy.

1991

171. SHOR N.Z. *The Development of Numerical Methods for Nonsmooth Optimization in the USSR.* // History of Mathematical Programming. A Collection of Personal Reminiscences. (Edited by J.K.LENSTRAS, A.H.G. RIGNOOUY KAN, A.SCHRIJVER). – CWI. Amsterdam. North-Holland. – P. 135–139.

172. SHOR N.Z. *Nondifferentiable optimization and dual bounds* // Abstracts of the 14th International symposium on Math.Progr. – Amsterdam.

173. SHOR N.Z. *The short review of development of numerical methods for nonsmooth optimization in USSR* // The history of mathematical programming. – Amsterdam.

174. МИХАЛЕВИЧ В.С., СЕРГИЕНКО И.В., ШОР Н.З. и др. *Пакет прикладных программ для решения на ЕС ЭВМ в диалоговом режиме задач дискретной и нелинейной оптимизации* – К.: Ин-т кибернетики им. В.М. Глушкова АН УССР. – 449 с. – (Деп. в ГОСФАП СССР, 1991).

175. МИХАЛЕВИЧ В.С., СЕРГИЕНКО И.В., ШОР Н.З. и др. *Пакет прикладных программ для решения задач дискретной и нелинейной оптимизации (Пакет ДИСНЕЛ)* // Кибернетика. – № 3. – С. 36–45.

176. ШОР Н.З., БЕРЕЗОВСКИЙ О.А. *Нахождение глобального экстремума квадратичной функции на квадратичной поверхности* // Информационные технологии в научных исследованиях и испытаниях. – К.: Ин-т кибернетики им. В.М. Глушкова АН УССР. – С. 30–34.

177. ШОР Н.З., КИСЕЛЕВА Е.М. Решение непрерывной задачи разбиения в условиях неполной информации ее исходных данных. // ЖВМ и МФ. – Т. 31. – № 6. – С. 799–809.

178. ВАЙНЗОФ А.М., ЛИХОВИД А.П., ЧЕРНОБЕЛЬСКИЙ И.М., ШОР Н.З., ЯНОВЧИК В.А. *О концепции моделирования экономического поведения концерна в новых условиях хозяйствования.* – К. – 17 с. – (Препринт // АН УССР. Ин-т кибернетики им. В.М.Глушкова: 91-18).

1992

179. SHOR N.Z. AND BEREZOVSKI O.A. *New algorithms for constructing optimal circumscribed and inscribed ellipsoids* // Optimization Methods and Software. – 1. – P. 283–299.

180. /ШОР Н.З., БЕРЕЗОВСКИЙ О.А. *Построение эллипсоида максимального объема, вписанного в многогранник, с использованием последовательного растяжения пространства* // Кибернетика и вычислительная техника. – Вып. 93. – К.: Наук. думка. – С. 1–6.

181. SHOR N.Z. *Dual Estimates in Multiextremal Problems* // Journal of Global Optimization. – 2. – P 411–418.

182. SHOR N.Z. *Quadratic dual estimates for some multiextremal polynomial extremal polynomial* // Abstracts of workshop on computational optimization (March 22-24, 1992), Technion, Israel, Haifa.

1993

183. ШОР Н.З., ТУКАЛЕВСКИЙ С.Д. *Практические вопросы применения двойственного алгоритма решения многоэкстремальной распределительной задачи при оптимизации работы комплекса очистки загрязненных вод* // Математические методы в экономических исследованиях. – К.: Ин-т кибернетики им. В.М. Глушкова НАН Украины. – С. 4–13.

184. БЕРЕЗОВСКИЙ О.А., ШОР Н.З. *Алгоритм решения задачи минимизации квадратичной функции на пересечении двух квадратичных поверхностей* // Теория и вычислительные проблемы оптимизации. – К.: Ин-т кибернетики им. В.М. Глушкова НАН Украины. – С. 30–35.

1994

185. ШОР Н.З., СТЕЦЮК П.И. *Кусочно-вогнутая задача рюкзачного типа* // Методы исследования экстремальных задач. – К.: Ин-т кибернетики им. В.М. Глушкова НАН Украины. – С. 21–28.

186. КИСЕЛЕВА Е.М., ШОР Н.З. *Исследование алгоритмов решения одного класса непрерывных задач разбиения* // Кибернетика и системный анализ. – № 1. – С. 84–97.

[English transl.: KISELEVA E.M. AND SHOR N.Z. *Analysis of algorithms for a class of continuous partition problems* // Cybernetics and systems analysis. – V. 30. – № 1. – P. 64–74.

187. ШОР Н.З., БЕРЕЗОВСКИЙ О.А. *Применение аппарата двойственных квадратичных оценок при решении системы полиномиальных уравнений на множестве комплексных чисел* // Кибернетика и системный анализ. – № 5. – С. 67–75.

[English transl.: SHOR N.Z. AND BEREZOVSKII O.A. *Using the method of dual quadratic solution to solve systems of polynomial equations in the complex domain* // Cybernetics and systems analysis. – V. 30. – № 5. – P. 686–692.

188. SHOR N.Z. *Nonmonotonic-subgradient methods with space dilation* // Abstracts of 15th International Symposium on Mathematical Programming. – Michigan University.

1995

189. ШОР Н.З., БЕРЕЗОВСКИЙ О.А. *Новые алгоритмы решения взвешенной задачи о максимальном взвешенном разрезе графа* // Кибернетика и системный анализ. – 1995. – № 2. – С. 100–106.

[English transl.: SHOR N.Z. AND BEREZOVSKII O.A. *New algorithms for the weighted maximum cut problem on graph* // Cybernetics and systems analysis. – V. 31. – № 2. – P. 240–245.

190. SHOR N.Z. *Минимизация матричных функций и недифференцируемая оптимизация* // Обзорение прикладной и промышленной математики. – К. – Т. 2. – С. 113–138.

191. ШОР Н.З., БЕРЕЗОВСКИЙ О.А. *Алгоритмы построения инвариантного эллипсоида минимального объема для устойчивой динамической системы* // Кибернетика и системный анализ. – № 3. – С. 130–137.

[English transl.: SHOR N.Z. AND BEREZOVSKII O.A. *Algorithms to construct minimum-volume invariant ellipsoid for a stable dynamic system* // Cybernetics and systems analysis. – V. 31. – № 3. – P. 421–427.

1996

192. SHOR N.Z., VOITISHIN YU.V. *Using Dual Network Bounds in Algorithms for Solving Generalized Set Packing-Partitioning Problems* // Computational Optimization and Applications. – № 6. – P. 293–303.

1997

193. SHOR N. *Using of Nonsmooth Optimization Methods for Solving Quasiblock Structured Problems of Semidefinite Programming*. Program and Abstracts. 16-th International Symposium on Mathematical Programming, August 24–29, Lausanne, Switzerland. – P. 250.

194. SHOR N., BARDADYM T., КНОПОВ P., ЛИХОВИД А., СТЕТСЬУК P. *On the Use of Nonsmooth Optimization Methods in Statistics and Stochastic Programming*. Program and Abstracts. 16-th International Symposium on Mathematical Programming, August 24–29, Lausanne, Switzerland. – P. 41–42.

195. ШОР Н.З., СТЕЦЬОК П.И. *Использование модификации r -алгоритма для нахождения глобального минимума полиномиальных функций* // Кибернетика и системный анализ. – № 4. – С. 28–49.

[English transl.: SHOR N.Z. AND STETSYUK P.I. *Modified r -algorithm to find the global minimum of polynomial functions* // Cybernetics and systems analysis. – V. 33. – № 4. – P. 482–498.

196. ШОР Н.З., СТЕЦЬОК П.И. *Использование r -алгоритма в задачах полуопределенного программирования* // Праці міжнародної конференції "Питання оптимізації обчислень", Київ, 6–8 жовтня 1997. – С. 330–335.

1998

197. SHOR N.Z. *Nondifferentiable optimization and polynomial problems*. – Boston; Dordrecht; London: Kluwer Academic Publishers. – 394 p.

198. ШОР Н.З. *Роль избыточных ограничений в улучшении двойственных оценок для полиномиальных оптимизационных задач* // Кибернетика и системный анализ. – № 4. – С. 106–121.

[English transl.: SHOR N.Z. *Role of redundant constraints for improvement dual bounds in polynomial optimization problems* // Cybernetics and systems analysis. – V. 34. – № 4. – P. 564–576.

1999

199. ШОР Н.З. *Двойственные оценки в квадратичных оптимизационных задачах* // Вестник Международного Соломонова университета. – № 1. – С. 92–108.

200. ШОР Н.З., БАРДАДЫМ Т.А., ЖУРБЕНКО Н.Г., ЛИХОВИД А.П., СТЕЦЮК П.И. *Использование методов негладкой оптимизации в задачах стохастического программирования* // Кибернетика и системный анализ. – № 5. – С. 33–47.

[English transl.: SHOR N.Z., BARDADYM T.A., ZHURBENKO N.G., LIKHOVID A.P. AND STETSYUK P.I. *Nonsmooth-optimization methods in problems of stochastic programming* // Cybernetics and systems analysis. – V. 35. – № 5. – P. 708–720.

2000

201. СЕРГИЕНКО И.В., ШОР Н.З. *Академик В.С.Михалевич – ученый и организатор науки* // Кибернетика и системный анализ. – № 1. – С. 77–100.

[English transl.: SERGIENKO I.V. AND SHOR N.Z. *Academician Mikhailovich V.S. as a scientist and science organizer* // Cybernetics and systems analysis. – V. 36. – № 1. – P. 57–75.

202. ШОР Н.З., СТЕЦЮК П.И., КРЫЛОВ С.В. *Нахождение глобальных минимумов полиномиальных функций с использованием двойственных квадратичных оценок* // Вестник Международного Соломонова университета. – № 4. – С. 217–233.

2001

203. SHOR N.Z. AND STETSYUK P.I. *Dual Solution of Quadratic-Type Problems by r -algorithm (subroutine DSQTPr)* // Abstracts of Second International Workshop: Recent Advances in Non-Differentiable Optimization. – Kyiv, Ukraine. – October 1–4. – P. 36.

204. SHOR N.Z., LIKHOVID A.P., ZHURBENKO N.G. *Solving two stage stochastic problems using nonsmooth optimization methods* // Proc. of 9th International Conference on Stochastic Programming, Berlin.

205. SHOR N.Z. *Lagrangian Quadratic Bounds in Polynomial Nonconvex and Boolean Models with Superfluous Constraints* // Advances in Convex Analysis and Global Optimization. (Honoring the Memory of C. Caratheodory (1873-1950).) Nonconvex Optimization and its Application. – Dordrecht, Boston, London: Kluwer Academic Publishers. – 54. – P. 181–204.

2002

206. SHOR N.Z. AND STETSYUK P.I. *Constructing Utility Functions by Methods of Nondifferentiable Optimization* // Constructing and Applying Objective Functions (A.S. TANGIAN AND J. GRUBER – EDS.). – Lecture Notes in Economics and Mathematical Systems. – Berlin: Springer-Verlag. – 510. – P. 215–232

207. SHOR N.Z. AND STETSYUK P.I. *Lagrangian Bounds in Multi-extremal Polynomial and Discrete Optimization Problems* // Journal of Global Optimization. – 23. – P. 1–41.

208. ШОР Н.З. *Монотонные модификации r -алгоритмов и их приложения* // Кибернетика и системный анализ. – № 6. – С. 74–96.

[English transl.: SHOR N.Z. *Monotone modifications of r -algorithms and their applications* // Cybernetics and systems analysis. – V. 38. – № 6. – P. 855–873.

209. ШОР Н.З., СТЕЦЮК П.И., ЛЕЙБОВИЧ М.М. *К выбору метода оценки параметров линейной модели* // Матеріали ІХ-ої Міжнародної наукової конференції ім. академіка М.Кравчука (16-19 травня 2002 р., Київ, КПІ). – 399 с.

2003

210. ШОР Н.З., ЖУРБЕНКО Н.Г., ЛИХОВИД А.П., СТЕЦЮК П.И. *Развитие алгоритмов недифференцируемой оптимизации и их приложения* // Кибернетика и системный анализ. – № 4. – С. 80–94.

[English transl.: SHOR N.Z., ZHURBENKO N.G., LIKHOVID A.P. AND STETSYUK P.I. *Algorithms of nondifferentiable optimization: development and applications* // Cybernetics and systems analysis. – V. 39. – № 4. – P. 39.

2004

211. ШОР Н.З., СТЕЦЮК П.И., БЕРЕЗОВСКИЙ О.А. *Двойственные оценки для оптимизационной задачи квадратичного типа на многообразии Штифеля* // Теория оптимальных решений. – К.: Ин-т кибернетики им. В.М. Глушкова НАН Украины. – С. 3–10.

212. ШОР Н.З., СТЕЦЮК П.И., БЕРЕЗОВСКИЙ О.А. *Нахождение оценок глобального экстремума в задачах на многообразии Штифеля* // Автоматика-2004. Матеріали 11-ї міжнародної конференції по автоматичному управлінню, 27-30 вересня 2004 року. – К.: Національний університет харчових технологій. – Т. 1. – С. 47. – (Секція 1.)

213. ГЛАЗУНОВ Н.М., ПОСТНИКОВА Л.П., ШОР Н.З. *Арифметическое моделирование случайных процессов и эргодическая теория* // Кибернетика и системный анализ. – № 4. – С. 73–86.

[English transl.: GLAZUNOV N.M., POSTNIKOVA L.P. AND SHOR N.Z. *Ergodic theory and Arithmetical simulation of random processes* // Cybernetics and systems analysis. – V. 40. – № 4. – P. 527–536

2005

214. ШОР Н.З., СЕРГИЄНКО І.В., ШИЛО В.П., СТЕЦЮК П.І. ТА ІНШІ *Задачі оптимального проектування надійних мереж*. – К.: Наук. думка. – 230 с.

215. SHOR N., STETSYUK P. *Quadratic-type models and upper bounds for the problems of finding the maximum weighted independent set in graphs* // Conference on Systems Analysis, Data Mining and Optimization in Biomedicine. – University of Florida, Gainesville, FL.

2006

216. ШОР Н.З., ШАРИФОВ Ф.А. *Общая задача синтеза надежных сетей* // Проблемы управления и информатики. – № 1–2. – С. 184–202.

217. КИСЕЛЕВА Е.М., ШОР Н.З. *Непрерывные задачи оптимального разбиения множеств: теория, алгоритмы, приложения* – К.: Наук. думка. – 562 с.

СОДЕРЖАНИЕ

Академик Н. З. Шор 3

Ч А С Т Ь I

Развитие алгоритмов недифференцируемой оптимизации	16
и их приложения	
Применение обобщенного градиентного спуска в блочном	39
программировании	
Метод случайного поиска для задач двухэтапного стохастического	45
программирования и его обобщение	
Алгоритм решения двухэтапной задачи стохастического	51
программирования	
Метод минимизации, использующий операцию растяжения	
пространства в направлении разности двух последовательных	59
градиентов	
Метод отсечения с растяжением пространства для решения задач	76
выпуклого программирования	
О методе решения одного класса динамических распределительных	81
задач	
Использование методов негладкой оптимизации в задачах	95
стохастического программирования	

Ч А С Т Ь II

Минимизация матричных функций и недифференцируемая	116
оптимизация	
Использование модификации r -алгоритма для нахождения	147
глобального минимума полиномиальных функций	
Роль избыточных ограничений в улучшении двойственных оценок	176
для полиномиальных оптимизационных задач	
Lagrangian Bounds in Multiextremal Polynomial and Discrete	198
Optimization Problems	
Список трудов Н. З. Шора	244

НАЦИОНАЛЬНАЯ АКАДЕМИЯ НАУК УКРАИНЫ
ИНСТИТУТ КИБЕРНЕТИКИ ИМ. В. М. ГЛУШКОВА

МИНИСТЕРСТВО ПРОСВЕЩЕНИЯ И МОЛОДЕЖИ
РЕСПУБЛИКИ МОЛДОВА

АКАДЕМИЯ ТРАНСПОРТА,
ИНФОРМАТИКИ И КОММУНИКАЦИЙ

Н. З. ШОР
МЕТОДЫ НЕДИФФЕРЕНЦИРУЕМОЙ
ОПТИМИЗАЦИИ И СЛОЖНЫЕ
ЭКСТРЕМАЛЬНЫЕ ЗАДАЧИ

Сборник избранных трудов

Издательство „Эврика“, тел. 63-90-29
мун. Кишинэу, Республика Молдова
Мунчештское шоссе, 121а

Подписано в печать 26.02.2008

Формат 60x90/16

Усл. печ. листов 17

Тираж 150

Заказ № 7

Типография Академии наук Молдовы
Г.Кишинев, ул. П. Мовилэ, 8



АКАДЕМИК Н. З. ШОР

Шор Наум Зуселевич – основоположник направления недифференцируемой оптимизации в математическом программировании, профессор, академик НАН Украины. Родился 1 января 1937 г. в Киеве. После окончания в 1958 г. Киевского национального университета имени Тараса Шевченко работал в Институте кибернетики НАН Украины. Автор 10 монографий и более 200 статей по математическому программированию, вычислительной математике и теории графов. Лауреат Государственных премий УССР (1973 г.), СССР (1981 г.), Украины (1993 г., 2000 г.), премии им. В. М. Глушкова НАН Украины (1987 г.), премии им. В. С. Михалевича НАН Украины (1997 г.).