

П.И. Стецюк

**ДВОЙСТВЕННЫЕ ОЦЕНКИ
В КВАДРАТИЧНЫХ
ЭКСТРЕМАЛЬНЫХ ЗАДАЧАХ**

Кишинэу • Эврика • 2018

**НАЦИОНАЛЬНАЯ АКАДЕМИЯ НАУК УКРАИНЫ
ИНСТИТУТ КИБЕРНЕТИКИ ИМ. В.М. ГЛУШКОВА
МИНИСТЕРСТВО ПРОСВЕЩЕНИЯ, КУЛЬТУРЫ И НАУКИ
РЕСПУБЛИКИ МОЛДОВА
АКАДЕМИЯ ТРАНСПОРТА, ИНФОРМАТИКИ
И КОММУНИКАЦИЙ**

П. И. СТЕЦЮК

**ДВОЙСТВЕННЫЕ ОЦЕНКИ
В КВАДРАТИЧНЫХ
ЭКСТРЕМАЛЬНЫХ ЗАДАЧАХ**

КИШИНЭУ • ЭВРИКА • 2018

Серия научных публикаций
«Недифференцируемая оптимизация и ее приложения»,
посвященная академику Н. З. ШОРУ

Редакционная коллегия серии:
акад. НАН Украины И. В. СЕРГИЕНКО,
д-р техн. наук Д. И. СОЛОМОН,
д-р физ.-мат. наук П. И. СТЕЦЮК

CZU 519.8

C 799

Стецюк, П. И.

Двойственные оценки в квадратичных экстремальных задачах / П. И. Стецюк ; Нац. акад. наук Украины, Ин-т кибернетики им. В. М. Глушкова, М-во просвещения, культуры и науки Респ. Молдова [и др.]. – Кишинэу : Эврика, 2018 (Tipografia Academiei de Științe a Moldovei). – 503 p. : fig., tab. – (Серия научных публикаций «Недифференцируемая оптимизация и ее приложения», посвященная академику Н. З. Шору / редкол.: И. В. Сергиенко [и др.], ISBN).

Texte : lb. engl., ucr., rusă. – 300 ex.

ISBN 978-9975-62-420-6 (Tipografia AȘM).

519.8

C 799

Книга посвящена развитию одной из центральных идей Н.З. Шора, которая связана с использованием алгоритмов негладкой оптимизации для нахождения и уточнения двойственных оценок целевой функции в квадратичных экстремальных задачах. Значительное внимание уделяется новым семействам функционально избыточных ограничений для улучшения точности оценок в булевых квадратичных задачах.

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

Рекомендовано к печати Ученым советом Института кибернетики им. В. М. Глушкова НАН Украины (протокол № 15 от 30.10.18)

Рецензенты : **П. С. Кнопов**, член-корреспондент НАН Украины,
Г. А. Донец, доктор физико-математических наук,
С. Л. Кривой, доктор физико-математических наук.

Работа выполнена при поддержке НАН Украины (проекты № 0117U000327, № 0118U005227) и Норвежского центра SIU (grant СРЕА-LT-2016/10003)

ISBN 978-9975-62-420-6

© Изд-во Эврика, 2018

© Стецюк П.И., 2018



НАУМ ЗУСЕЛЕВИЧ ШОР (1937 – 2006) – один из основоположников направления недифференцируемой оптимизации в математическом программировании. Автор 10 монографий и более 200 статей.

Родился 1 января 1937 г. в г. Киеве. После окончания в 1958 г. Киевского государственного университета им. Т. Г. Шевченко работал в Институте кибернетики им. В. М. Глушкова НАН Украины (с 1983 по 1990 г. – заведующим лабораторией, с 1990 по 2006 г. – заведующим отделом). Профессор (1989), член-корреспондент (1990), академик (1997) Национальной академии наук Украины. Лауреат Государственных премий УССР (1973), СССР (1981), Украины (1993, 1999), премий НАН Украины имени В. М. Глушкова (1987), имени В. С. Михалевича (1997).

Работы Н. З. Шора положили начало численным методам оптимизации недифференцируемых функций и существенно повлияли на развитие линейного, нелинейного, дискретного и стохастического программирования. Разработанные им методы получили мировую известность и способствовали становлению Института кибернетики им. В. М. Глушкова НАН Украины как одного из мировых лидеров в области теории и практики методов оптимизации.

Серия научных публикаций
«Недифференцируемая оптимизация и ее приложения»,
посвященная Науму Зуселевичу Шору

Совместный проект Института кибернетики им. В.М. Глушкова НАН Украины и Академии транспорта, информатики и коммуникаций Республики Молдова связан с изданием серии книг под общим названием «Недифференцируемая оптимизация и ее приложения», где каждая отдельная книга посвящена современным теоретическим и практическим достижениям в тех или иных направлениях, связанных с научным наследием академика Н.З. Шора. Его результатом являются три книги:

1. СОЛОМОН Д.И. Дробное программирование и недифференцируемая оптимизация. – Кишинэу: Эврика, 2010. – 556 с.
2. СТЕЦЮК П.И. Методы эллипсоидов и r -алгоритмы. – Кишинэу, Эврика, 2014. – 488 с.
3. КИСЕЛЕВА Е.М., КОРЯШКИНА Л.С. Непрерывные задачи оптимального разбиения множеств и r -алгоритмы. – Киев, Наукова думка, 2015. – 400 с.

Первая книга посвящена методам решения задач дробного программирования с помощью субградиентных алгоритмов в сочетании со схемами декомпозиции. Значительное внимание в монографии уделено задачам большой размерности и задачам транспортного типа. Заметим, что именно в связи с решением транспортной задачи в сетевой форме и появился первый субградиентный метод (Шор, 1962).

Вторая книга посвящена развитию методов эллипсоидов, которые являются частным случаем алгоритмов с преобразованием пространства (Шор, 1969). Эти методы базируются на внешней аппроксимации множества экстремумов эллипсоидами с монотонным уменьшением их объема и имеют „красивую“ геометрическую интерпретацию. В книге приведены ускоренные модификации методов эллипсоидов и показана их тесная связь с r -алгоритмами Шора и другими методами.

Третья книга посвящена исследованию различных классов непрерывных задач оптимального разбиения множеств и родственных к ним по постановке непрерывных задач оптимального шарового покрытия. Показано, каким образом в каждом случае задача бесконечномерной

оптимизации может быть сведена к задаче оптимизации негладкой функции от конечного количества переменных. Большое внимание уделяется практическим приложениям теории оптимального разбиения множеств с использованием r -алгоритма Шора и его модификаций.

Книга П.И. Стецюка „Двойственные оценки в квадратичных экстремальных задачах“ продолжает эту серию. Она посвящена исследованиям по развитию идеи, связанной с техникой лагранжевой релаксации для квадратичных невыпуклых задач (Шор, 1985). Эта идея оказалась чрезвычайно плодотворной при решении комбинаторных проблем и позволяет выделить такие их подклассы, которые разрешимы за полиномиальное время. Основные результаты Н.З. Шора в этом направлении можно найти в его книгах „*Shor (1998) Nondifferentiable optimization and polynomial problems. Dordrecht: Kluwer*“ и „*Шор, Стеценко (1989) Квадратичные экстремальные задачи и недифференцируемая оптимизация. Киев: Наукова думка*“.

В настоящей книге представлены исследования автора по разработке алгоритмов негладкой оптимизации для нахождения двойственных лагранжевых оценок целевой функции в квадратичных невыпуклых задачах. Им предложено и обосновано новые схемы функционально избыточных ограничений для улучшения точности этих оценок в булевых квадратичных моделях, к которым сводится ряд NP -сложных экстремальных задач на графах. На их основе получен ряд новых результатов для задачи о максимальном разрезе графа и задачи о максимальном независимом множестве вершин графа. Обоснованы новые свойства оценок Шора для взвешенного числа внутренней устойчивости графа. Это позволило построить новые семейства графов, для которых взвешенное число устойчивости графа (в общем случае NP -сложная задача) может быть найдено за полиномиальное время.

Предложенные методы нашли практическое применение при решении задачи оптимальной компоновки 3D-объектов в круговом контейнере, для нахождения максимального взвешенного разреза графа с заданными количествами вершин в обоих подмножествах, для определения подмножеств вершин с заданными свойствами в различного рода сетях, для нахождения оптимальных параметров в продуктивной модели Леонтьева и других.

Надеемся, что данная книга окажется полезной для специалистов в области математического программирования, которые найдут в ней ряд оригинальных идей для построения эффективных алгоритмов решения многоэкстремальных и комбинаторных задач.

Редакционная коллегия

Предисловие

Многие задачи целочисленного линейного программирования могут быть переформулированы как экстремальные квадратичные задачи с булевыми переменными (булевы квадратичные задачи). Так, например, условие булевости переменной $x \in \{0, 1\}$ представляется квадратичным равенством $x^2 - x = 0$. Если две переменные x_i и x_j не могут одновременно принимать значение, равное единице, то это условие может быть записано в виде равенства $x_i x_j = 0$ (равносильно линейному неравенству $x_i + x_j \leq 1$). Для экстремальных квадратичных задач Н.З. Шор предложил использовать двойственный подход к получению и уточнению оценок целевой функции (1985 год). В квадратичных задачах на минимум эти оценки будут границами снизу для минимального значения целевой функции, а в квадратичных задачах на максимум – границами сверху для максимального значения целевой функции.

Идея Шора (1985) состоит в разработке алгоритмов нахождения двойственных оценок на основе методов недифференцируемой оптимизации и использовании функционально избыточных квадратичных ограничений (их добавление не изменяет множества допустимых решений исходной квадратичной задачи) для улучшения точности двойственных оценок. Впервые двойственный подход к получению оценок целевой функции был предложен в работе Н.З. Шора и А.С. Давыдова (1985) для задач булевого программирования (см. стр. 18–25 этой книги). В 1986–1987 годах он был дополнен процедурами уточнения оценок и применен к другим многоэкстремальным задачам, в том числе и к задаче нахождения глобального минимума полинома.

Опишем суть двойственного подхода для получения и уточнения оценок в квадратичных экстремальных задачах на примере экстремальной квадратичной задачи с ограничениями-равенствами: найти

$$Q^* = \max_{x \in R^n} Q_0(x) \quad (1)$$

при ограничениях

$$Q_i(x) = 0, \quad i = 1, \dots, m. \quad (2)$$

Здесь квадратичные функции имеют вид

$$Q_i(x) = (K_i x, x) + (b_i, x) + c_i,$$

где K_i – симметричные вещественные $n \times n$ -матрицы, b_i – n -мерные векторы из R^n , c_i – вещественные числа, $i = 0, \dots, m$. Некоторые из функций $Q_i(x)$, $i = 0, 1, \dots, m$ могут быть также линейными.

В общем случае задача (1), (2) многоэкстремальна и относится к классу NP -трудных задач. Оценки сверху для Q^* можно получить путем следующей лагранжевой релаксации. Пусть $u \in R^m$ – вектор множителей Лагранжа, соответствующий ограничениям (2). Функция Лагранжа для задачи (1), (2) имеет вид

$$L(x, u) = Q_0(x) + \sum_{i=1}^m u_i Q_i(x) \equiv (K(u)x, x) + (b(u), x) + c(u),$$

где

$$K(u) = K_0 + \sum_{i=1}^m u_i K_i, \quad b(u) = b_0 + \sum_{i=1}^m u_i b_i, \quad c(u) = c_0 + \sum_{i=1}^m u_i c_i.$$

Рассмотрим функцию

$$\Psi(u) = \max_{x \in R^n} L(x, u) \equiv \max_{x \in R^n} \left[(K(u)x, x) + (b(u), x) + c(u) \right].$$

Пусть $\lambda_{max}(K(u))$ – максимальное собственное число симметричной $n \times n$ -матрицы $K(u)$.

Функция $\Psi(u)$ является выпуклой функцией от переменных u (как результат операции нахождения максимума по переменным x для семейства линейных по переменным u функций). Область определения функции $\Psi(u)$ обозначим $\text{dom } \Psi$. Она состоит из подмножества

$$\Omega^- = \left\{ u \in R^m : \lambda_{max}(K(u)) < 0 \right\}$$

(подмножество $u \in R^m$, для которых матрица $K(u)$ отрицательно определена) и подмножества тех точек

$$u \in \Omega^0 = \left\{ u \in R^m : \lambda_{max}(K(u)) = 0 \right\},$$

для которых система линейных уравнений

$$2K(u)x + b(u) = 0 \tag{3}$$

имеет решение. Для всех других точек $\Psi(u) = +\infty$.

Если $\text{dom } \Psi \neq \emptyset$, то для любого $u \in \text{dom } \Psi$ значение функции $\Psi(u)$ является нетривиальной (т. е. не равной $+\infty$) оценкой сверху для Q^* – оптимального значения целевой функции в задаче (1), (2). Наилучшая

оценка сверху для Q^* в классе лагранжевых оценок вида $\Psi(u)$ связана с решением следующей задачи негладкой оптимизации:

$$\Psi^* = \min_{u \in \text{dom} \Psi} \Psi(u). \quad (4)$$

Точками негладкости функции $\Psi(u)$ есть точки границы множества Ω^- , где система линейных уравнений (3) имеет неединственное решение.

Оценку Ψ^* с любой заданной точностью можно найти за полиномиальное время методом эллипсоидов; для одной его итерации требуется $O(m^2) + O(n^3)$ арифметических операций. Из них $O(m^2)$ операций связано с растяжением пространства двойственных переменных (множителей Лагранжа), а $O(n^3)$ операций требуются для вычисления вектора, определяющего полупространство локализации оптимальных множителей Лагранжа (при фиксированных значениях множителей решается система линейных уравнений с n неизвестными, определяется максимальное собственное число симметричной $n \times n$ -матрицы и соответствующий этому числу собственный вектор).

Оценка Ψ^* обладает следующими свойствами. Если минимум в (4) достигается на $u^* \in \Omega^-$, то $\Psi^* = Q^*$, т. е. оценка точная. При этом находится и точка глобального минимума $x^* = x(u^*)$, где $x(u^*)$ – решение системы (3) при $u = u^*$. Если же минимум в (4) достигается на границе области Ω^- , то может существовать так называемый "разрыв двойственности" $\Delta^* = \Psi^* - Q^* > 0$. Один из способов уменьшения Δ^* связан с введением функционально избыточных ограничений (при этом может увеличиться и количество переменных в задаче).

Функционально избыточные ограничения – это ограничения, добавление которых не изменяет множества допустимых решений начальной квадратичной задачи. Однако при этом изменяется функция Лагранжа, что в некоторых случаях позволяет уменьшить Δ^* . Если к ограничениям исходной задачи (1), (2) прибавить r функционально избыточных квадратичных ограничений $Q_{m+1}(x) = 0, \dots, Q_{m+r}(x) = 0$, $r \geq 1$, то новая квадратичная задача принимает вид: найти

$$Q^* = \max_{x \in R^n} Q_0(x) \quad (5)$$

при ограничениях

$$Q_i(x) = 0, \quad i = 1, \dots, m + r. \quad (6)$$

Теорема 1. Если Ψ_1^* – оценка вида Ψ^* для задачи (5), (6), то $\Psi_1^* \leq \Psi^*$.

Доказательство. Задаче (5), (6) соответствует вектор множителей Лагранжа $U \in R^{m+r}$, и функция Лагранжа для нее имеет вид:

$$L_1(x, U) = Q_0(x) + \sum_{i=1}^{m+r} u_i Q_i(x) = L(x, u) + \sum_{i=m+1}^{m+r} u_i Q_i(x).$$

Поскольку $L_1(x, (\{u\}, 0, \dots, 0)) = L(x, u)$ и $\Psi_1(\{u\}, 0, \dots, 0) = \Psi(u)$, то

$$\Psi_1^* = \min_{U \in \text{dom } \Psi_1} \Psi_1(U) \leq \min_{u \in \text{dom } \Psi} \Psi(u) = \Psi^*.$$

Теорема 1 не только констатирует, что функционально избыточные ограничения могут улучшить точность двойственной оценки, но и объясняет, что это улучшение есть следствием увеличения количества множителей Лагранжа. Ограничения, которые являются линейными комбинациями уже существующих ограничений, не отражаются на точности двойственной оценки, т. е. $\Psi_1^* = \Psi^*$. Вклад таких ограничений в функцию Лагранжа эквивалентен лишь определенному изменению множителей Лагранжа при существующих ограничениях. Однако добавление функционально избыточных ограничений, которые являются нетривиальными следствиями из условий задачи, в ряде случаев приводит к тому, что двойственная оценка Ψ_1^* может стать точной для Q^* . Последнее означает, что оптимальное значение Q^* со сколь угодно большой точностью можно найти за полиномиальное время, которое зависит от числа переменных и числа ограничений в квадратичной задаче.

Функционально избыточными могут быть следующие ограничения:

- а) квадратичные следствия линейных ограничений: например, квадратичное ограничение в форме $(b_i^T x + c_i)(b_j^T x + c_j) \geq 0$ является следствием из двух линейных ограничений-неравенств: $b_i^T x + c_i \geq 0$ и $b_j^T x + c_j \geq 0$;
- б) квадратичные ограничения, которые характеризуют неоднозначность представления произведения трех либо большего числа переменных задачи. Как правило, они имеют место при сведении полиномиальной задачи к квадратичной. Например, имеются переменные x_1 , $x_2 = x_1^2$ и $x_3 = x_1^3$. Тогда квадратичное ограничение $x_2^2 - x_1 x_3 = 0$ есть следствием неоднозначного представления x_1^4 , а именно $x_1^4 = (x_2^2)^2 = (x_3^3)(x_1)$;

- в) квадратичные ограничения, которые являются следствиями булевости или бинарности переменных задачи. Например, для бинарных переменных $x_i^2 = 1$, $x_j^2 = 1$, $x_k^2 = 1$ всегда справедливо квадратичное неравенство $x_i x_j + x_i x_k + x_j x_k \geq -1$.

Более детальную информацию об этих и других семействах функционально избыточных ограничений, их использовании для нахождения глобального минимума полинома и в экстремальных задачах на графах (максимальное устойчивое множество вершин графа, максимальный разрез графа и др.) можно найти в этой книге.

Функционально избыточные ограничения сыграли огромную роль при получении эффективных оценок снизу для целевой функции в задачах нахождения глобального минимума полиномиальной функции $P(x)$ от одной или нескольких переменных. Эти задачи специальным образом сводятся к многоэкстремальным квадратичным задачам (на минимум) с определенными семействами функционально избыточных квадратичных ограничений. Н.З. Шор доказал, что двойственная оценка для таких квадратичных задач совпадает со значением p^* полинома $P(x)$ в точке глобального минимума тогда и только тогда, когда полином $\bar{P}(x) = P(x) - p^*$ может быть представлен как сумма квадратов других полиномов. Эти результаты имеют отношение к классическим работам Д. Гильберта по разложению неотрицательных полиномиальных форм в сумму квадратов. Разработанный метод дает возможность не только доказать существование такой декомпозиции (если она существует), но и найти одно из возможных представлений полинома $\bar{P}(x)$ в виде суммы квадратов других полиномов. Более того, этим методом можно определить значение глобального минимума полинома $P(x)$.

Чрезвычайно интересными оказались результаты Н.З. Шора для задачи нахождения взвешенного максимального независимого множества вершин графа. Эта задача играет большую роль в многочисленных приложениях: теории информации и кодировании, проектировании различных устройств при определенных условиях несовместности; она тесно связана с известными задачами выбора, разбиения множеств, раскраски графов и другими комбинаторными задачами, имеющими массу приложений. В то же время, она принадлежит к классу NP -полных задач (при целочисленных весах вершин). Двойственные оценки для этой задачи тесно связаны с известными числами Ловаса $\theta(G, w)$ и $\theta'(G, w)$, которые играют большую роль при обосновании результатов о полиномиальной разрешимости ряда задач в теории графов.

Идея Шора (1985) может быть использована при создании эффективных методов решения тех многоэкстремальных задач, которые мо-

жно описать с помощью квадратичных моделей (такие модели встречаются во многих приложениях). Эффективность таких методов обеспечивается двумя моментами. Во-первых, двойственные оценки для квадратичных моделей являются более точными, чем оценки в линеаризованных аналогах этих моделей, и целесообразно их использовать в сочетании с методом ветвей и границ. Во-вторых, использование функционально избыточных ограничений позволяет выделить среди многоэкстремальных задач такие их подклассы, которые разрешимы за полиномиальное время. При этом можно даже указать верхнюю границу сложности такого класса задач, которая зависит от количества квадратичных ограничений.

Особого внимания заслуживают и алгоритмы нахождения двойственных оценок в квадратичных моделях. Их можно считать альтернативой использованию методов внутренних точек для решения задач полуопределенного программирования (semidefinite programming), т. е. задач оптимизации, в которых в качестве ограничения фигурирует требование неотрицательной определенности некоторых матриц. Действительно, многие задачи полуопределенного программирования целесообразно рассматривать как задачи недифференцируемой оптимизации и применять для их решения эффективные методы минимизации негладких выпуклых функций. Условие неотрицательной определенности симметричной $(n \times n)$ -матрицы X (принято обозначать $X \succeq 0$) эквивалентно тому, что минимальное собственное число матрицы X не отрицательное: $\lambda_{\min}(X) \geq 0$. Но $\lambda_{\min}(X)$ – вогнутая недифференцируемая функция элементов матрицы, т. е. если элементы матрицы $X(u) = \{x_{ij}\}_{i,j=1}^n$ являются линейными функциями от вектора варьируемых параметров $u \in R^m$, то условие $X(u) \succeq 0$ эквивалентно выпуклому негладкому ограничению $\varphi(u) = -\lambda_{\min}(X(u)) \leq 0$.

Настоящая книга содержит результаты развития автором идеи Шора (1985) и ее применения для решения квадратичных экстремальных задач и задач комбинаторной оптимизации. Значительное внимание уделяется новым семействам функционально избыточных ограничений для улучшения точности двойственных оценок в булевых квадратичных задачах, к которым сводится ряд NP -сложных экстремальных задач на графах. На их основе получен ряд новых результатов для задачи о максимальном разрезе графа и задачи о максимальном независимом множестве вершин графа. Так, например, выявлены новые свойства оценок Шора для числа устойчивости графа. Они связаны с тем, что из ограничений квадратичной задачи для максимального независимого множества вершин графа вытекают линейные неравенства, справедливые

для ряда известных подструктур в графе, в том числе и для p -колеса (английский термин „ p -wheel“). p -Колесо состоит из клики с количеством вершин $p \geq 1$ и нечетного цикла с количеством вершин $(2k + 1)$, где $k \geq 2$, причем вершины клики и нечетного цикла не пересекаются, а каждая вершина клики связана со всеми вершинами нечетного цикла. Это дает возможность существенно расширить семейство графов, для которых задача нахождения числа устойчивости графа не является NP -трудной, а может быть решена за полиномиальное время. Как частный случай, сюда относится известное семейство W -совершенных графов, которому соответствует случай, когда $p = 1$, т. е. в p -колесо входит клика только с одной вершиной.

Материал книги размещен по тематической направленности в шести разделах. В первом разделе, который называется „Двойственные оценки в квадратичных задачах“, содержатся статьи:

- 1.1 Шор Н.З., Давыдов А.С. О методе получения оценок в квадратичных экстремальных задачах с булевыми переменными // Кибернетика. – 1985. – № 2. – С. 48–50;
- 1.2 Шор Н.З. Квадратичные оптимизационные задачи // Известия АН СССР. Техническая кибернетика. – 1987. – № 1. – С. 128–139;
- 1.3 Шор Н.З., Об одном подходе к получению глобальных экстремумов в полиномиальных задачах математического программирования // Кибернетика. – 1987. – № 5. – С. 102–106;
- 1.4 Шор Н.З. Об одном классе оценок глобального минимума полиномиальных функций // Там же. – 1987. – № 6. – С. 9–11;
- 1.5 Шор Н.З. Задачи минимизации матричных функций и недифференцируемая оптимизация // Обзорение прикладной и промышленной математики. – М.: научн. изд.-во "ТВП", 1995. – Т. 2. – С. 113–138.

Первые четыре статьи – это пионерские работы Шора по технике лагранжевых (двойственных) оценок для невыпуклых квадратичных задач. В статье 1.1 впервые предложен метод получения оценок в квадратичных экстремальных задачах с булевыми переменными, который впоследствии использовался при исследовании ряда NP -трудных комбинаторных задач и экстремальных задач на графах. Статья 1.2 посвящена численным методам решения оценочных задач в экстремальных задачах с квадратичными целевой функцией и ограничениями. Показано, что в многоэкстремальном случае оценки снизу минимального значения целевой функции можно получить методами двойственности, решая выпуклую оптимизационную задачу с ограничениями типа неотрицательной определенности матриц. В статье 1.3 методика оптимальных

лагранжевых оценок была применена к исследованию оценок снизу целевой функции в полиномиальных задачах математического программирования путем их сведения к квадратичным экстремальным задачам. Она была дополнена идеей использовать функционально избыточные ограничения для улучшения точности этих оценок. Для случая минимизации полиномов от одной переменной это позволило доказать, что соответствующие оценки являются точными. Эта идея нашла более полное изложение в статье 1.4 для задачи минимизации полиномов от многих переменных. Статья 1.5, опубликованная в "Обзрении прикладной и промышленной математики", содержит обзор по квадратичным экстремальным задачам и задачам матричной оптимизации, которые имеют широкий спектр приложений в теории устойчивости и оптимизации динамических систем. Здесь приведены примеры матричных задач, связанных с получением оптимальной функции Ляпунова, и методы их решения с помощью алгоритмов недифференцируемой оптимизации.

Второй раздел называется „Двойственные оценки Шора“, где представлены статьи 1997 – 2002 гг., в которых автору этой книги повезло быть соавтором Наума Зуселевича. В раздел включены следующие статьи¹:

- 2.1 ШОР, СТЕЦЮК (1997) Использование модификации g -алгоритма для нахождения глобального минимума полиномиальных функций [31];
- 2.2 ШОР, СТЕЦЮК (1997) Использование g -алгоритма в задачах полуопределенного программирования [32];
- 2.3 ШОР, СТЕЦЮК, КРЫЛОВ (2000) Нахождение глобальных минимумов полиномиальных функций с использованием двойственных квадратичных оценок [44];
- 2.4 SHOR, STETSYUK (2001) Dual Solution of Quadratic-Type Problems by g -algorithm (subroutine DSQTPPr) [50];
- 2.5 SHOR, STETSYUK (2002) Lagrangian bounds in multiextremal polynomial and discrete optimization problems [54].

Статьи 2.1 – 2.5 посвящены различным аспектам, связанным с применением методов недифференцируемой оптимизации в квадратичных задачах. Их объединяет использование подхода Шора для получения двойственных оценок для невыпуклых квадратичных моделей, которые встречаются во многих приложениях, например, в полиномиальных задачах оптимизации и NP -трудных экстремальных задачах на графах.

¹Здесь и далее приводятся сокращенные ссылки на включенные в книгу статьи, опуская даже инициалы авторов. Подробную ссылку можно найти по ее номеру в размещенном в конце книги списке работ автора.

Этот подход дает возможность среди NP -трудных невыпуклых квадратичных задач выделить такие их подклассы, для которых проблема нахождения значения глобального минимума целевой функции разрешима за полиномиальное время. Здесь рассмотрены свойства квадратичных двойственных оценок для полиномиальных функций, алгоритмы нахождения этих оценок, основанные на использовании метода эллипсоидов и модификаций r -алгоритмов. Приведены результаты тестовых экспериментов для экстремальных задач, которые описываются в виде моделей нелинейного программирования квадратичного типа. Показано, что использование функционально избыточных ограничений в ряде случаев существенно улучшает двойственные оценки.

Под названием „Новые семейства функционально избыточных ограничений в квадратичных булевых задачах“ в третьем разделе собраны следующие работы:

- 3.1 Стецюк (2005) О функционально избыточных ограничениях для булевых оптимизационных задач квадратичного типа [78];
- 3.2 Стецюк (2006) Новые модели квадратичного типа для задачи о максимальном взвешенном разрезе графа [80];
- 3.3 Стецюк, Пардалос (2006) Об уточнении лагранжевых двойственных оценок в бинарных и булевых квадратичных задачах [82];
- 3.4 Стецюк (2007) Об уточнении верхних оценок Н.З. Шора для числа независимости графа [86];
- 3.5 Стецюк, Золотых (2010) Бинарный квадратичный многогранник и его аппроксимации [112];
- 3.6 Золотых, Стецюк (2010) О внешней аппроксимации бинарного и булевого квадратичных многогранников [115].

В статьях 3.1–3.4 предложены и обоснованы новые семейства функционально избыточных ограничений для улучшения точности двойственных оценок в квадратичных экстремальных задачах с бинарными (± 1) и булевыми $(0, 1)$ переменными. Первые два семейства (3.1 и 3.3) основаны на расширении множества булевых и бинарных переменных и формировании полного набора функционально избыточных ограничений для связи новых и старых переменных. Третье семейство (3.2) использует свойство нечетного количества бинарных переменных и вытекающие из него квадратичные неравенства для четного и нечетного их количества. С их помощью уточнены верхние оценки Шора для числа независимости графа (3.4), сформулированы новые квадратичные модели для максимального взвешенного разреза графа (3.2). Из предложенных моделей вытекает известная линейная модель Барахоны-Маджуба и рекомендации для построения ее „усиленных“ аналогов.

В статьях 3.5 и 3.6 построено семейство многогранников, заданных с помощью семейств линейных неравенств, и показано, что они аппроксимируют сверху многогранник $ViQP_n$ (бинарный квадратичный многогранник, который введен по аналогии с булевым квадратичным многогранником Падберга $BoQP_n$). Для многогранников $ViQP_n$ и $BoQP_n$ при $n = 2, 3, 4$ построенные аппроксимации полностью совпали с неприводимыми системами линейных неравенств, полученными с помощью программы Skeleton (автор – Н.Ю. Золотых, Нижегородский государственный университет им. Н.И. Лобачевского).

Работы четвертого раздела объединены под названием „Оценки для числа устойчивости графа“. В раздел включены такие работы:

- 4.1 СТЕЦЮК, ПАРДАЛОС, КРОШКО (2006) О новых лагранжевых двойственных оценках для числа устойчивости графа [85];
- 4.2 СТЕЦЮК, БУТЕНКО, БЕРЕЗОВСКИЙ (2007) Об одной верхней оценке для взвешенного числа устойчивости графа [88];
- 4.3 СТЕЦЮК, ЛЯШКО, НУРМИНСКИЙ (2009) Точная ЛП-оценка для взвешенного числа устойчивости t -совершенных графов [111];
- 4.4 СТЕЦЮК (2008) О новых свойствах оценок Шора для взвешенного числа устойчивости графа [91];
- 4.5 СТЕЦЮК, БУТЕНКО, ЛИХОВИД (2008) ЛП-ориентированная верхняя оценка для числа устойчивости графа на основе p -колес [92];
- 4.6 СТЕЦЮК, ЛИХОВИД (2009) Об ЛП-ориентированных верхних оценках для взвешенного числа устойчивости графа [96];
- 4.7 СТЕЦЮК (2010) Оценки Н.З. Шора в квадратичных экстремальных задачах и их применение в комбинаторной оптимизации [113].

Статьи 4.1, 4.2 и 4.3 связаны с построением на основе функционально избыточных ограничений из раздела 3 неизвестных ранее оценок для числа устойчивости и взвешенного числа устойчивости графа и доказательством того, что они являются точными для специальных семейств графов. Статьи 4.4, 4.5 и 4.6 связаны с новыми свойствами оценок Шора, которые позволяют существенно расширить семейство графов, для которых задача нахождения числа устойчивости графа не является NP -трудной и может быть решена за полиномиальное время. На основе этих свойств создано программную реализацию алгоритма для нахождения ЛП-ориентированных верхних оценок взвешенного числа устойчивости неориентированных графов с тысячами вершин. Статья 4.7 была предметом пленарного доклада на конференции "Дискретная оптимизация и исследование операций" (Алтай, 27 июня – 3 июля 2010).

Пятый и шестой разделы посвящены специальным квадратичным задачам, для которых предложенные методы нашли практическое применение. Сюда включены 9 работ и две рукописи:

- 5.1 Шор, Стецюк, Березовский (2004) Двойственные оценки для специальной оптимизационной задачи квадратичного типа на многообразии Штифеля [71];
- 5.2 Стецюк, Березовский (2006) Лагранжевая оценка для максимального разреза графа с заданными количествами вершин в обоих подмножествах разбиения [81];
- 5.3 Стецюк (2006) Квадратичная оптимизационная задача для максимального k -плекса в неориентированном графе (рук.);
- 5.4 Стецюк, Бардадим, Ляшко (2017) Квадратична задача для максимального k -плекса в неорієнтованому графі [244];
- 5.5 Стецюк, Ляшко, Бардадим (2017) Властивості квадратичної задачі про максимальний k -плекс у неорієнтованому графі [247];
- 6.1 Стецюк, Кошлай (2010) Оптимальная нормированная структура спроса и добавленной стоимости в продуктивной модели Леонтьева [110];
- 6.2 Стецюк, Кошлай (2011) Об одной экстремальной задаче для связи прямой и двойственной моделей Леонтьева [127];
- 6.3 Стецюк (2011) О спектральных свойствах матриц Леонтьева [130];
- 6.4 Стецюк, Эмменеггер (2014) Максимальное сингулярное число матрицы и его экономическая интерпретация [171];
- 6.5 Стецюк, Романова, Шайтхауэр (2014) О глобальном минимуме целевой функции в задаче равновесной упаковки кругов [173];
- 6.6 Стецюк, Лиховид (2015) Задача равновесной упаковки кругов (рук.).

Они связаны с решением специальной оптимизационной задачи квадратичного типа на многообразии Штифеля (5.1), нахождением максимального разреза графа с заданными количествами вершин в обоих подмножествах (5.2), определением подмножеств вершин графа с заданными свойствами (5.3, 5.4, 5.5), нахождением оптимальных параметров в продуктивной модели Леонтьева (6.1, 6.2, 6.3), нахождением сбалансированной компоновки 3D-объектов в круговом контейнере (6.5 и 6.6).

Завершает книгу список научных трудов автора за 1982–2018 годы.

Автор признателен сотрудникам отдела методов негладкой оптимизации А.П. Лиховиду, Р.В. Машкиной и И.И. Парасюк, которые взяли на себя основной труд по подготовке этой книги в системе L^AT_EX.

П.И. Стецюк
Киев, октябрь 2018 г.

Р А З Д Е Л I

Двойственные оценки в квадратичных задачах (пионерские работы Шора)

В раздел включены статьи:

- 1.1. О методе получения оценок в квадратичных экстремальных
задачах с булевыми переменными (*Шор, Давыдов, 1985*) 18
- 1.2. Квадратичные оптимизационные задачи (*Шор, 1987*) 26
- 1.3. Об одном подходе к получению глобальных экстремумов
в полиномиальных задачах математического
программирования (*Шор, 1987*) 44
- 1.4. Об одном классе оценок глобального минимума
полиномиальных функций (*Шор, 1987*) 55
- 1.5. Задачи минимизации матричных функций и
недифференцируемая оптимизация (*Шор, 1995*) 62

О методе получения оценок в квадратичных экстремальных задачах с булевыми переменными

Н. З. Шор, А. С. Давыдов
Кибернетика. — 1985. — № 2. — С. 48–50.

Предложен новый метод получения оценок в квадратичных экстремальных задачах с булевыми переменными, имеющий теоретические и практические приложения в комбинаторике.

Рассматривается класс экстремальных задач следующего вида:

$$\min f_0(x), \quad x = \{x_1, \dots, x_n\} \quad (1)$$

при ограничениях

$$\begin{aligned} f_i(x) \leq 0; \quad i \in I^+, \quad f_i(x) = 0; \quad i \in I, \\ I^+ \cup I = \{1, \dots, m\}, \end{aligned} \quad (2)$$

$$x_j \in \{0, 1\}, \quad j = 1, \dots, n, \quad (3)$$

где $f_\nu(x)$ — квадратичные функции от переменных x_1, \dots, x_n , $\nu = 0, 1, \dots, m$.

Условие (3) можно заменить квадратичными равенствами:

$$x_j^2 - x_j = 0, \quad j = 1, \dots, n. \quad (3')$$

Заметим, что благодаря (3') без ограничения общности можно считать, что $f_\nu(x)$, $\nu = 0, \dots, m$, не содержат линейных членов.

Пусть f^* — оптимальное значение целевой функции в задаче (1)–(3). Условимся, что $f^* = +\infty$, если задача (1)–(3) не имеет допустимых решений.

Используем стандартный метод множителей Лагранжа [3] для получения оценок снизу для f^* . Получаем следующую функцию Лагранжа

$$L(x, u) = f_0(x) + \sum_{i=1}^m u_i f_i(x) + \sum_{j=1}^n \bar{u}_j (x_j^2 - x_j).$$

Здесь $u = \{u_1, \dots, u_m; \bar{u}_1, \dots, \bar{u}_n\}$ – $(m+n)$ -мерный вектор множителей Лагранжа.

Пусть U – подмножество $(n+m)$ -мерного линейного пространства множителей Лагранжа, вырезаемое ограничениями $u_i \geq 0$ для тех i , у которых в (2) стоит знак « \leq » ($i \in I^+$). Рассмотрим

$$\Psi(u) = \inf_x L(x, u). \quad (4)$$

При $u \in U$ для любого допустимого x $L(x, u) \leq f_0(x)$. Следовательно, $\Psi(u) \leq f^*$, $\forall u \in U$. Отсюда $\Psi^* = \sup_{u \in U} \Psi(u) \leq f^*$.

При любом $u \in U$ $L(x, u)$ – квадратичная функция, которую можно записать в виде

$$L(x, u) = (K(u)x, x) + l(x, u),$$

где $K(u)$ – симметричная матрица $n \times n$, $l(x, u)$ – линейная по x функция. Заметим, что элементы матрицы $K(u) = \{k_{ij}(u)\}_{i,j=1}^n$ линейно зависят от u .

Пусть U^+ – подмножество U значений u , при которых $K(u)$ является положительно определенной матрицей, $\bar{U}^+ = \{u \mid u \in U, K(u) \text{ – неотрицательно-определенной матрицей}\}$. При $u \in U^+$ задача (4) имеет единственный минимум $x(u)$, который получается путем решения соответствующей системы линейных уравнений. При $u \notin \bar{U}^+$ матрица $K(u)$ имеет отрицательные собственные числа и $\Psi(u) = -\infty$. Таким образом, поиск $\sup_{u \in U} \Psi(u)$ следует ограничить подмножеством \bar{U}^+ .

Лемма. U^+ является не пустым выпуклым множеством.

Доказательство. Зафиксируем произвольные допустимые значения u_1, \dots, u_m . Легко видеть, что за счет выбора достаточно больших $\bar{u}_1, \bar{u}_2, \dots, \bar{u}_n$ диагональные элементы матрицы $K(u)$ можно сделать сколь угодно большими, оставляя без изменения остальные элементы. Это доказывает существование значений u , при которых матрица $K(u)$ является положительно определенной, т. е. $U^+ \neq \emptyset$.

Докажем выпуклость U^+ . Пусть $u^{(1)}, u^{(2)} \in U^+$, $\alpha_1, \alpha_2 \geq 0$; $\alpha_1 + \alpha_2 = 1$. Рассмотрим $\bar{u} = \alpha_1 u^{(1)} + \alpha_2 u^{(2)}$. Так как элементы матрицы $K(u)$ являются линейными функциями от u , то $K(\bar{u}) = \alpha_1 K(u^{(1)}) + \alpha_2 K(u^{(2)})$. Для любого $x \neq 0$ $(K(u^{(1)})x, x) > 0$, $(K(u^{(2)})x, x) > 0$. Отсюда

$$(K(\bar{u})x, x) = \alpha_1 (K(u^{(1)})x, x) + \alpha_2 (K(u^{(2)})x, x) > 0,$$

что доказывает положительную определенность $K(\bar{u})$ и, следовательно, выпуклость множества U^+ . Лемма доказана.

Функция $\Psi(u)$, как инфимум по x линейных по u функций, является вогнутой на U^+ . Таким образом, задача определения

$$\Psi^* = \sup_{u \in \bar{U}^+} \Psi(u) = - \inf_{u \in \bar{U}^+} (-\Psi(u)) \quad (5)$$

сводится к задаче выпуклого программирования. Заметим, что множество U^+ является открытым, поэтому экстремум в (5), если он существует, может достигаться и на границе множества U^+ , т. е. во множестве \bar{U}^+ .

Пусть $-\Psi(u) = \varphi(u)$, $\varphi(u)$ – выпуклая функция. Будем решать задачу определения $\varphi^* = \min_{u \in \bar{U}^+} \varphi(u)$ в предположении, что минимум существует. Для решения этой задачи можно воспользоваться методами недифференцируемой оптимизации. Заметим, что условие $u \in \bar{U}^+$ содержит два типа ограничений:

- 1) $u_i \geq 0$ при $i \in I^+$, где I^+ – подмножество индексов ограничений (2), которые записаны в форме (\leq);
- 2) $K(u)$ – неотрицательно-определенная матрица, что эквивалентно бесконечной системе линейных по u неравенств

$$\sum_{i,j=1}^n y_i y_j k_{ij}(u) \geq 0, \quad \forall y \|y\| = 1, \quad y = \{y_1, \dots, y_n\},$$

которую можно переписать в компактном виде:

$$\lambda(u) = \min_{\|y\|=1} \sum_{i,j=1}^n k_{ij}(u) y_i y_j \geq 0, \quad (6)$$

где $\lambda(u)$ – минимальное собственное число матрицы $K(u)$, вогнутая функция от u .

Таким образом, получаем задачу

$$\min \varphi(u) \mid u_i \geq 0, \quad i \in I^+; \quad \lambda(u) \geq 0. \quad (7)$$

Так как функция $\varphi(u)$ и ее субградиент не определены за пределами области U^+ , нельзя воспользоваться методом негладких точных

внешних штрафных функций. Приходится применять для получения приближенного решения барьерные штрафные функции [6], например, в форме

$$S_{\varepsilon, \delta}(u) = \varphi(u) + \sum_{i \in I^+} \frac{\varepsilon}{u_i} + \frac{\delta}{\lambda(u)},$$

где ε и δ – достаточно малые положительные числа. Для минимизации $S_{\varepsilon, \delta}(u)$ можно использовать один из обобщенных градиентных методов недифференцируемой оптимизации (методы минимизации гладких функций в данном случае не пригодны, так как $S_{\varepsilon, \delta}(u)$ – недифференцируемая функция). Остановимся на способе вычисления обобщенных градиентов $g_\varphi(u)$, $g_\lambda(u)$ функций $\varphi(u)$ и $\lambda(u)$ при $u \in U^+$. Так как $\varphi(u) = -\psi(u)$, $g_\varphi(u) = -g_\psi(u)$. В соответствии с известными результатами выпуклого анализа [3], $g_\psi(u)$ получается как вектор невязок в ограничениях (2) при $x = x(u)$, т. е. $g_\psi(u) = \{f_i(x(u))\}_{i=1}^m$.

В соответствии с формулой (6) для $\lambda(u)$ обобщенный градиент функции $\lambda(u)$ в точке $u = \bar{u}$ вычисляется следующим образом:

а) находится минимальное собственное число $\lambda(\bar{u})$ матрицы $K(\bar{u})$ и соответствующий ему собственный вектор $y(\bar{u})$, $\|y(\bar{u})\| = 1$ (если $\lambda(\bar{u})$ – кратное собственное число, то в качестве $y(\bar{u})$ можно выбрать произвольный нормированный собственный вектор, соответствующий $\lambda(\bar{u})$; в этом случае в точке \bar{u} $\lambda(\bar{u})$ не является дифференцируемой);

б) в качестве $g_\lambda(\bar{u})$ берется градиент к линейной по u функции

$$e(u) = \sum_{i,j=1}^n y_i(\bar{u})y_j(\bar{u})k_{ij}(u).$$

Зная способы вычисления $g_\lambda(u)$ и $g_\varphi(u)$, легко построить алгоритм вычисления субградиента функции $S_{\varepsilon, \delta}(u)$ и применить для ее минимизации один из известных субградиентных методов с растяжением пространства (например, r -алгоритм [3]). Заметим, что при определении шагов по направлению в r -алгоритме нужно позаботиться о том, чтобы не переходить границу области U^+ .

Другой способ решения задачи (7) связан с возможностью применения метода эллипсоидов ([1, 4]) или его модификаций [2].

Суть метода эллипсоидов при решении задач выпуклого программирования состоит в следующем: перед началом очередной итерации оптимум локализован в некотором эллипсоиде; берется точка \bar{x} – центр эллипсоида. Если эта точка является допустимой, то с использованием

субградиента целевой функции строится отсекающая гиперплоскость и вокруг оставшейся части эллипсоида описывается эллипсоид по возможности меньшего объема (если \mathcal{T} не является допустимой, то для построения отсекающей гиперплоскости используется субградиент к ограничению, которое в данной точке не выполняется).

Таким образом, после завершения очередной итерации получаем в качестве области локализации оптимума опять эллипсоид, но меньшего объема. Модификации метода эллипсоидов предложены для ускорения сходимости этого метода и состоят в том, что погружение области локализации оптимума в эллипсоид происходит после двух или большего числа отсечений. Отметим, что метод эллипсоидов может быть реализован в форме обобщенного градиентного метода с растяжением пространства [1].

Как следует из сказанного, для реализации метода эллипсоидов и его модификаций при решении задачи (7) необходимо уметь находить $g_\varphi(u)$ для $u \in U^+$ или $g_\lambda(u)$ для $u \notin U^+$. На способах вычисления этих обобщенных градиентов останавливались выше.

Процедура вычисления $g_\lambda(u)$ основана на решении задачи нахождения собственных чисел и векторов симметричных матриц, которая является довольно трудоемкой, если требуется высокая точность результатов.

Воспользуемся способом, предложенным в [5]. Он позволяет строить отсекающие гиперплоскости, которые отделяют данную точку $\bar{u} \notin U^+$ от области U^+ без использования алгоритма нахождения минимального собственного числа и соответствующего ему собственного вектора.

Пусть $\bar{u} \notin U^+$. Это значит, что матрица $K(\bar{u})$ не является положительно определенной. Пусть $M_1(\bar{u}), M_2(\bar{u}), \dots, M_n(\bar{u})$ – последовательность главных миноров этой матрицы. По критерию Сильвестра найдется такой номер t , $1 \leq t \leq n$, что $d_i = \det M_i(\bar{u}) > 0$, $i = 1, \dots, t-1$, $d_t = \det M_t(\bar{u}) \leq 0$.

При $t = 1$ $k_{11}(\bar{u}) \leq 0$, и так как при $u \in U^+$ $k_{11}(u) > 0$ и $k_{11}(u)$ является линейной функцией от u , то гиперплоскость $k_{11}(u) = 0$ отсечет точку \bar{u} от множества U^+ .

Пусть $t > 1$. Рассмотрим подматрицы $M_t^i(\bar{u})$, которые получаются при вычеркивании t -го столбца и i -й строки из матрицы $M_t(\bar{u})$. Пусть

$$\begin{aligned} \xi_i &= (-1)^i \det M_t^i(\bar{u}), & i = 1, \dots, t, \\ \xi_t &= \det M_{t-1} \cdot (-1)^t \neq 0, & \xi = \{\xi_i\}_{i=1}^t, \end{aligned}$$

$$r_{\bar{u}}(u) = \sum_{i,j=1}^t k_{ij}(u) \xi_i \xi_j = \sum_{j=1}^t \xi_j \left(\sum_{i=1}^t k_{ij}(u) \xi_i \right). \quad (8)$$

При $u \in U^+$ $r_{\bar{u}}(u) > 0$, так как $M_t(u)$ является положительно определенной матрицей.

Пусть $u = \bar{u}$. $\sum_{i=1}^t k_{ij}(\bar{u}) \xi_i$ в силу определения вектора ξ с точностью до знака совпадает с определителем матрицы, получающейся из матрицы $M_t(\bar{u})$ заменой t -го столбца на i -й. Если $t \neq i$, то получающаяся матрица имеет два одинаковых столбца и ее определитель равен 0.

Таким образом,

$$r_{\bar{u}}(\bar{u}) = \sum_{i,j=1}^t k_{ij}(\bar{u}) \xi_i \xi_j = \xi_t \cdot (-1)^t d_t = d_{t-1} d_t \leq 0. \quad (9)$$

Заметим, что $r_{\bar{u}}(u)$ – линейная по u функция. Уравнение $r_{\bar{u}}(u) = 0$ задает гиперплоскость, которая отсекает согласно (9) точку \bar{u} от множества U^+ . Необходимый для определения $r_{\bar{u}}(u)$ вектор ξ вычисляется с использованием методов линейной алгебры.

Теперь можно описать алгоритм, основанный на методе эллипсоидов [1, 4] для решения задачи получения оценки (7): $\min \varphi(u)$ при ограничениях $u_i \geq 0$, $i \in I^+$; $\lambda(u) \geq 0$.

Начальная информация. Пусть нам априори известны точка $u^{(0)}$ и число $R_0 > 0$ такое, что оптимальное решение задачи (7) u^* удовлетворяет соотношению $\|u^* - u^{(0)}\| \leq R_0$.

Задана также матрица преобразования пространства $B_0 = I$ (единичная $(m+n) \times (m+n)$ матрица).

($k+1$)-й шаг алгоритма. $k = 0, 1, 2, \dots$ Перед началом этого шага имеем точку $u^{(k)}$, число R_k – радиус области локализации оптимума и матрицу B_k преобразования пространства. Проверяем неравенства $u_i \geq 0$, $i \in I^+$. Если для некоторого $\bar{i} \in I^+$ $u_{\bar{i}} < 0$, то берем в качестве $g(u^{(k)})$ вектор, у которого все координаты равны 0, кроме координаты \bar{i} , равной -1 . В противном случае вычисляем, как описано выше, величины d_1, d_2, \dots, d_n , где $d_i = \det M_t(u^{(k)})$. Если $d_1 \leq 0$, то берем $g(u^{(k)})$ равным антиградиенту функции $k_{11}(u)$. Если $d_1, \dots, d_{t-1} > 0$, а $d_t \leq 0$, $1 < t \leq n$, то формируем, как уже указано, функцию $r_{u^{(k)}}(u)$ и за $g(u^{(k)})$ принимаем антиградиент этой функции.

Если $d_i > 0$, $i = 1, \dots, n$, т. е. $u \in U^+$, то, решая относительно x

систему линейных уравнений

$$\frac{\partial L(x, u^{(k)})}{\partial x_j} = 0, \quad j = 1, \dots, n,$$

получаем $x(u^{(k)})$ и формируем компоненты вектора $g(u^{(k)})$ следующим образом:

$$\begin{aligned} g_i(u^{(k)}) &= -f_i(x(u^{(k)})), \quad i = 1, \dots, m, \\ g_{m+j}(u^{(k)}) &= -x_j^2(u^{(k)}) + x_j(u^{(k)}), \quad j = 1, \dots, n. \end{aligned}$$

Получив $g(u^{(k)})$, используем его стандартным способом в методе эллипсоидов [1]. Находим:

$$\begin{aligned} \xi^{(k)} &= \frac{B_k^T g(u^{(k)})}{\|B_k^T g(u^{(k)})\|}, \\ u^{(k+1)} &= u^{(k)} - h_k B_k \xi^{(k)}, \\ B_{k+1} &= B_k \cdot R_\beta(\xi^{(k)}), \quad \beta = \sqrt{\frac{m+n-1}{m+n+1}}, \\ h_k &= \frac{R_k}{m+n-1}, \quad R_{k+1} = R_k \cdot q, \quad q = \frac{m+n}{\sqrt{(m+n)^2 - 1}} \end{aligned}$$

и переходим к $(k+2)$ -му шагу.

Учитывая, что метод эллипсоидов сходится довольно медленно, на практике целесообразно использовать его ускоренные модификации, например, описанные в [2], при этом способ построения отсекающих гиперплоскостей остается прежним.

Квадратичные экстремальные задачи с булевыми переменными часто встречаются в приложениях (квадратичная задача о назначениях, задача о наилучшей аппроксимации в евклидовой метрике заданной точки линейной целочисленной комбинацией векторов из заданного семейства и др.).

Однако важно отметить, что многие задачи, которые, как правило, формулируются как задачи булева линейного программирования, могут быть переформулированы как квадратичные задачи с булевыми переменными, при этом предлагаемые в данной статье оценки функционалов могут оказаться более точными, чем соответствующие оценки линейного программирования. Особенно это характерно для моделей,

отражающих те или иные условия несовместности (задачи о максимальном внутренне устойчивом множестве, о раскраске графов и др.). Так, если булевы переменные x_i и x_j не могут одновременно принимать значения 1, то это условие можно записать в виде $x_i x_j = 0$, т. е. в форме (2) (сравните с линейным ограничением $x_i + x_j \leq 1$). В тех случаях, когда получение оценок линейного программирования требует больших затрат времени ЭВМ, более точные оценки, получаемые из квадратичных моделей, могут оказаться практически более эффективными при реализации метода ветвей и границ.

Отметим, что вопрос частного случая задачи (1)–(3) (ограничения (2) отсутствуют) о получении оценок рассмотрен в работе [7].

Литература

1. ШОР Н. З. Метод отсечения с растяжением пространства для решения задач выпуклого программирования // Кибернетика. – 1977. – № 1. – С. 94–95.
2. ШОР Н. З., ГЕРШОВИЧ В. И. Об одном семействе алгоритмов для решения задач выпуклого программирования // Кибернетика. – 1979. – № 4. – С. 62–67.
3. ШОР Н. З. Методы минимизации недифференцируемых функций и их приложения. – К.: Наук. думка, 1979. – 199 с.
4. ЮДИН Д. Б., НЕМИРОВСКИЙ А. С. Информационная сложность и эффективные методы решения выпуклых экстремальных задач // Экономика и мат. методы, 1976. – 12. – № 2. – С. 357–369.
5. GRÖTSHEL M., LOVÁSZ L., SHRIJVER A. The ellipsoid method and its consequences in combinatorial optimization // Combinatorica. – 1981. – 1. – № 2. – P. 169–197.
6. GROßMANN C., KAPLAN A. Strafmethoden und modifizierte Lagrangefunktionen in der nichtlinearen Optimierung. – Leipzig: Teubner Verlagsgesellschaft, 1979. – 200 S.
7. KÖRNER F., RICHTER CL. Zur effektiven lösung von booleschen, quadratischen Optimierungsproblemen // Numer. Math. – 1982. – 40. – P. 99–109.

Поступила в редакцию 28.11.83

Квадратичные оптимизационные задачи

Н. З. Шор

Известия АН СССР. Техническая кибернетика. — 1987. — № 1. — С. 128–139.

Рассматриваются численные методы, связанные с экстремальными задачами с квадратичными целевым функционалом и ограничениями. Показано, что в многоэкстремальном случае необходимые для применения методов типа ветвей и границ оценки снизу оптимального значения целевого функционала можно получить методами двойственности, решая выпуклую оптимизационную задачу с ограничениями типа неотрицательной определенности некоторых линейных по переменным матриц. Рассматриваются методы решения оценочных задач. Описано приложение намеченного подхода к некоторым дискретным оптимизационным задачам.

1. Постановка задачи

Под квадратичной оптимизационной задачей (к.з.) будем понимать задачу квадратичного программирования следующего вида:

определить

$$K^* = \inf K_0(x), \quad x \in M \subseteq E^n; \quad (1.1)$$

при ограничениях

$$K_i(x) \leq 0, \quad i \in I; \quad K_j(x) = 0; \quad j \in J; \quad I \cap J = \emptyset, \quad (1.2)$$

где $K_\nu(x) = (A_\nu x, x) + (l_\nu, x) + c_\nu$ — квадратичные функции, определенные на n -мерном евклидовом пространстве E^n , A_ν — симметричные матрицы размерности $n \times n$, l_ν — n -мерные векторы, c_ν — константы, $\nu \in \{0\} \cup I \cup J$, I, J — непересекающиеся конечные подмножества натуральных чисел. M либо совпадает с E^n , либо является выпуклым замкнутым многогранным подмножеством E^n .

Если задача (1.1), (1.2) является несовместной, будем считать, что $K^* = +\infty$. К.з. является далеко идущим обобщением выпуклой задачи квадратичного программирования, методам решения которой посвящено огромное количество литературы (см., например, [1, 2, 19]). Среди к.з. можно выделить достаточно интересные частные случаи:

- а) выпуклые к.з. (в.к.з.), когда $J = \emptyset$, A_ν являются неотрицательно-определенными для $\nu \in \{0\} \cup I$;
- б) к.з. с линейными ограничениями (к.з.л.о.), когда $A_\nu = 0$ для $\nu \in I \cup J$. Если A_0 – неотрицательно-определенная матрица, то получаем выпуклую задачу квадратичного программирования; в общем случае к.з.л.о. – многоэкстремальная задача и принадлежит к классу *NP*-трудных задач; в частности, к этому классу относятся задачи вогнутого квадратичного программирования;
- в) квадратичные оптимизационные задачи с булевыми переменными (б.к.з.), когда на компоненты вектора x накладываются условия $x_k \in \{0; 1\}$, $k = 1, \dots, n$. Эти условия эквивалентны квадратичным равенствам $x_k^2 - x_k = 0$, $k = 1, \dots, n$, иначе говоря б.к.з. можно рассматривать как частный случай квадратичных оптимизационных задач.

Отметим, что любую задачу нелинейного программирования с полиномиальными целевой функцией и ограничениями можно свести к к.з., вводя новые переменные и квадратичные подстановки следующего вида: $x_i x_j = z_1$ либо $x_i^2 = z_2$ и т. п., понижающие порядок полиномов до квадратичного.

К.з. имеют очень широкую область приложений. Так среди к.з.л.о. отметим задачи проектирования и размещения чипов, задачи линейной дополнительной [18]; среди б.к.з. – квадратичную задачу о назначениях, экстремальные задачи на графах (максимальное внутренне устойчивое множество, задачи раскраски и разрезания графов и др.), играющие большую роль при решении задач оптимального проектирования: размещения, компоновки, разбиения на блоки элементов сложных устройств; задачи выбора, упаковки и др.

По многоэкстремальным к.з. в достаточно общей постановке имеется небольшое число публикаций ([9], [16]). В общем случае, ввиду сложности проблемы, при решении к.з. не избежать перебора, поэтому естественно применять схему метода «ветвей и границ». Для реализации этой схемы следует разработать эффективный и достаточно общий метод получения оценок снизу для целевого функционала в задаче (1.1), (1.2). В данной статье будет исследован двойственный подход к получению нижних оценок, основанный на использовании функции Лагранжа. Такой подход впервые описан в [9]. Он приводит к необходимости решения специальных задач выпуклого программирования: задано некоторое параметрическое семейство симметричных матриц; среди множества значений параметров, при которых матрицы являются

неотрицательно-определенными, необходимо найти значение, на котором некоторая функция от параметров достигает максимума. Оказывается, что эти задачи можно достаточно эффективно решать, используя методы негладкой оптимизации. Отметим, что экстремальные задачи на классе неотрицательно-определенных матриц появляются в многочисленных приложениях, не связанных непосредственно с к.з.; разработка эффективных алгоритмов их решения наталкивается на значительные трудности (см. [14], [15]). Поэтому раздел, связанный с построением алгоритмов негладкой оптимизации для решения экстремальных задач на классе неотрицательно-определенных матриц, представляет и самостоятельный интерес.

Далее в работе более детально будут рассмотрены задачи с линейными ограничениями и булевы квадратичные задачи. Раздел, относящийся к б.к.з., основан на совместных работах С. И. Стеценко и автора [4], [8], [20].

2. Алгоритмы получения нижних оценок функционала в квадратичных оптимизационных задачах

Рассмотрим квадратичную оптимизационную задачу (1.1), (1.2). Пусть $|I| = m_1$; $|J| = m_2$; $N = m_1 + m_2$, $\lambda = \{\lambda_1, \dots, \lambda_{m_1}; \lambda_{m_1+1}, \dots, \lambda_N\}$ – вектор множителей Лагранжа, у которого первые m_1 компонент соответствуют ограничениям в форме неравенств, а остальные – ограничениям в форме равенств в (1.2), Λ^+ – подмножество N -мерного пространства E^N , состоящее из произвольных векторов, в которых первые m_1 компонент неотрицательны. Введем функцию Лагранжа к.з. (1.1), (1.2) $L(x, \lambda)$, определенную на прямом произведении $M \times \Lambda^+$, следующего вида

$$L(x, \lambda) = K_0(x) + \sum_{i=1}^N \lambda_i K_i(x).$$

Пусть T – множество допустимых решений задачи (1.1), (1.2). Если $\bar{x} \in T$, $\bar{\lambda} \in \Lambda^+$, то $\bar{\lambda}_i K_i(\bar{x}) \leq 0$ для $\{i : 1 \leq i \leq m_1\}$; $\bar{\lambda}_i K_i(\bar{x}) = 0$ для $i = m_1 + 1, \dots, N$. Таким образом, если $\bar{x} \in T$, $\bar{\lambda} \in \Lambda^+$, то $L(\bar{x}, \bar{\lambda}) \leq K_0(\bar{x})$, отсюда

$$\Psi(\bar{\lambda}) = \inf_{x \in M} L(x, \bar{\lambda}) \leq \inf_{x \in T} L(x, \bar{\lambda}) \leq \inf_{x \in T} K_0(x) = K^*$$

при любом $\bar{\lambda} \in \Lambda^+$. Значит,

$$\Psi(\lambda) = \inf_{x \in M} L(x, \lambda)$$

служит оценкой снизу для K^* при $\lambda \in \Lambda^+$. Для получения наиболее точной оценки такого типа нужно решить задачу нахождения

$$\Psi^* = \sup_{\lambda \in \Lambda^+} \Psi(\lambda).$$

Функция Лагранжа $L(x, \lambda)$ представима в следующем виде $L(x, \lambda) = (A(\lambda)x, x) + (l(\lambda), x) + c(\lambda)$, где

$$A(\lambda) = A_0 + \sum_{i=1}^N \lambda_i A_i, \quad l(\lambda) = l_0 + \sum_{i=1}^N \lambda_i l_i, \quad c(\lambda) = c_0 + \sum_{i=1}^N \lambda_i c_i.$$

Заметим, что элементы матрицы $A(\lambda) = \{a_{ij}(\lambda)\}_{i,j=1}^n$ линейно зависят от λ . Пусть $D(\bar{D})$ – подмножества E^N , состоящие из таких λ , что $A(\lambda)$ является положительно-определенной (соответственно – неотрицательно-определенной) матрицей. Легко показать, что D и \bar{D} – выпуклые множества, если они не пусты, при этом \bar{D} является замыканием D . На множестве D функция $\Psi(\lambda)$ принимает конечные значения и является вогнутой. В обобщенном смысле она является вогнутой и на \bar{D} .

Если $\lambda \in \bar{D}$, то задача нахождения

$$\Psi(\lambda) = \inf_{x \in M} L(x, \lambda)$$

является задачей выпуклого квадратичного программирования и может быть решена практически достаточно эффективно конечно-шаговым алгоритмом. Если $\lambda \notin \bar{D}$, то задача нахождения $\Psi(\lambda)$ является, вообще говоря, многоэкстремальной и трудно разрешимой. Поэтому вместо

$$\Psi^* = \sup_{\lambda \in \Lambda^+} \Psi(\lambda)$$

мы будем рассматривать более слабую оценку

$$\bar{\Psi}^* = \sup_{\lambda \in \bar{D} \cap \Lambda^+} \Psi(\lambda) \leq \Psi^*.$$

Если $M = E^n$, то при $\lambda \notin \bar{D}$, $\Psi(\lambda) = -\infty$. Поэтому в этом важном случае $\bar{\Psi}^* = \Psi^*$. Ниже для простоты будем рассматривать случай $M = E^n$. При $\lambda \in D$ матрица $A(\lambda)$ невырождена, и для нахождения

$\Psi(\lambda)$ нужно решить систему линейных уравнений $2A(\lambda)x(\lambda) + l(\lambda) = 0$, откуда $x(\lambda) = -\frac{1}{2}A^{-1}(\lambda)l(\lambda)$,

$$\Psi(\lambda) = -\frac{1}{4}(A^{-1}(\lambda)l(\lambda), l(\lambda)) + c(\lambda). \quad (2.1)$$

В области $D \cap \Lambda^+$ $\Psi(\lambda)$ является вогнутой непрерывно дифференцируемой, при этом ее градиент есть

$$g_{\Psi}(\lambda) = \{K_i(x(\lambda))\}_{i=1}^N. \quad (2.2)$$

Задача нахождения

$$\bar{\Psi}^* = \sup_{\lambda \in D \cap \Lambda^+} \Psi(\lambda) \quad (2.3)$$

является задачей выпуклого программирования. Рассмотрим возможные алгоритмы ее решения.

а) Метод эллипсоидов ([5, 11]). Суть метода эллипсоидов при решении задач выпуклого программирования $\min f_0(\lambda) \mid f_i(\lambda) \leq 0, i = 1, \dots, m$, состоит в следующем: перед началом очередной итерации оптимум локализован в некотором эллипсоиде; берется точка $\bar{\lambda}$ – центр эллипсоида. Если эта точка является допустимой, то с использованием субградиента целевой функции строится отсекающая гиперплоскость, и вокруг оставшейся части эллипсоида описывается новый эллипсоид по возможности меньшего объема. Если $\bar{\lambda}$ не является допустимой, то строится гиперплоскость, отделяющая $\bar{\lambda}$ от области допустимых значений λ , и вокруг части эллипсоида, содержащей допустимые точки, описывается новый эллипсоид. В качестве такой отделяющей гиперплоскости можно брать, в частности, гиперплоскость, проходящую через точку $\bar{\lambda}$ ортогонально к субградиенту ограничения, которое в данной точке не выполняется [11]. В обоих случаях ($\bar{\lambda}$ – допустимая или недопустимая точка) после завершения очередной итерации получаем в качестве области локализации оптимума опять эллипсоид, но меньшего объема. Модификации метода эллипсоидов, предложенные для ускорения его сходимости, состоят в том, что погружение области локализации оптимума в новый эллипсоид происходит после двух или большего числа отсечений. Отметим, что метод эллипсоидов может быть реализован в форме метода обобщенного градиентного спуска с растяжением пространства [5]. Как следует из сказанного выше, для реализации метода эллипсоидов и его модификаций при решении задачи (2.3) необходимо уметь находить $g_{\Psi}(\lambda)$ для $\lambda \in D \cap \Lambda^+$ (это можно делать по формуле (2.2)) и субградиент к функции-ограничению, если $\lambda \in D \cap \Lambda^+$. (Если

λ принадлежит границе области D , то $g_\Psi(\lambda)$ необязательно определен, поэтому в этом случае также будем использовать субградиент к ограничениям.)

Условие $\lambda \in D \cap \Lambda^+$ содержит два типа ограничений:

а) $\lambda_i \geq 0$ для $i \in I$;

б) $A(\lambda)$ – положительно-определенная матрица, что эквивалентно бесконечной системе линейных по λ неравенств

$$\sum_{i,j=0}^n a_{ij}(\lambda)x_i x_j > 0, \quad \forall x \in E^n; \quad x = \{x_1, \dots, x_n\}, \quad x \neq 0,$$

которую можно переписать в компактном виде

$$\mu(\lambda) = \min_{\{x \mid \|x\|=1\}} \sum_{i,j=1}^n a_{ij}(\lambda)x_i x_j > 0, \quad (2.4)$$

где $\mu(\lambda)$ – минимальное собственное число матрицы $A(\lambda)$.

Замечание. Из выражения (2.4) очевидно, что $\mu(\lambda)$ – вогнутая функция λ . В качестве суперградиента функции $\mu(\lambda)$ в точке $\bar{\lambda}$ $g_\mu(\bar{\lambda})$ можно взять градиент к линейной по λ функции

$$\Psi_{\bar{\lambda}}(\lambda) = \sum_{i,j=1}^n a_{ij}(\lambda)y_i(\bar{\lambda})y_j(\bar{\lambda}), \quad (2.5)$$

где $y_i(\bar{\lambda})$, $i = 1, \dots, n$, – компоненты нормированного собственного вектора матрицы $A(\bar{\lambda})$, соответствующего минимальному собственному числу $\mu(\bar{\lambda})$. $\mu(\lambda)$ – негладкая функция, ее субградиент терпит разрыв в тех точках λ , в которых матрица $A(\lambda)$ имеет кратное минимальное собственное число.

Для проверки принадлежности $\bar{\lambda}$ множеству $D \cap \Lambda^+$ рассматриваем сначала условие а). Если для некоторого $i = i^* \in I$; $\bar{\lambda}_{i^*} < 0$, тогда в качестве отсекающей гиперплоскости в методе эллипсоидов можно взять $\lambda_{i^*} = 0$.

Если условие а) выполняется, т. е. $\bar{\lambda} \in \Lambda^+$, то проверяем условие б). Для этого определяем $\mu(\bar{\lambda})$ и соответствующий собственный вектор матрицы $A(\bar{\lambda})$. Если $\mu(\bar{\lambda}) \leq 0$, то в качестве отсекающей гиперплоскости можно взять следующую $(g_\mu(\bar{\lambda}), \lambda - \bar{\lambda}) = 0$, где $g_\mu(\bar{\lambda})$ совпадает с градиентом функции $\Psi_{\bar{\lambda}}(\lambda)$ (см. (2.5)). Описанная выше процедура проверки условия б) и построения отсекающей гиперплоскости в методе

эллипсоидов, если это условие не выполняется, связана с нахождением минимального собственного числа матрицы $A(\lambda)$, что практически эквивалентно по сложности решению задачи на собственные значения для матрицы $A(\lambda)$. Решение этой задачи с высокой точностью требует при большом n значительного времени счета. В [13] в аналогичной ситуации предлагается другой способ построения отсекающей гиперплоскости в методе эллипсоидов. Рассмотрим его применительно к нашей задаче.

Пусть $\mu(\bar{\lambda}) \leq 0$. Это значит, что матрица $A(\bar{\lambda})$ не является положительно-определенной. Пусть $R_1(\bar{\lambda}), \dots, R_n(\bar{\lambda})$ – последовательность главных миноров этой матрицы. По критерию Сильвестра найдется такой номер t , $1 \leq t \leq n$, что $d_i = \det R_i(\bar{\lambda}) > 0$ для $i = 1, \dots, t-1$; $d_t = \det R_t(\bar{\lambda}) \leq 0$. Если $t = 1$, $d_t = a_{11}(\bar{\lambda}) \leq 0$, и так как при $\lambda \in D$ $a_{11}(\bar{\lambda}) > 0$ и $a_{11}(\lambda)$ является линейной функцией от λ , то гиперплоскость $a_{11}(\lambda) = 0$ отсекает точку $\bar{\lambda}$ от множества D . Пусть $t > 1$. Рассмотрим подматрицы $R_t^i(\bar{\lambda})$, которые получаются вычеркиванием t -го столбца и i -й строки ($i \leq t$) из матрицы $R_t(\bar{\lambda})$. Пусть

$$\xi_i = (-1)^i \det R_t^i(\bar{\lambda}), \quad i = 1, \dots, t, \quad \xi = \{\xi_i\}_{i=1}^t.$$

Введем функцию, линейную по λ

$$\rho_\xi(\lambda) = \sum_{i,j=1}^t a_{ij}(\lambda) \xi_i \xi_j = \sum_{j=1}^t \xi_j \left(\sum_{i=1}^t a_{ij}(\lambda) \xi_i \right).$$

Если $\lambda \in D$, $\rho_\xi(\lambda) > 0$, т. к. матрица $a_{ij}(\lambda)$ является положительно-определенной. При $\lambda = \bar{\lambda}$

$$S_j = \sum_{i=1}^t a_{ij}(\bar{\lambda}) \xi_i$$

в силу определения вектора ξ с точностью до знака совпадает с определением матрицы, получающейся из матрицы $R_t(\bar{\lambda})$ заменой t -го столбца на j -й. Если $t \neq j$, то эта матрица будет иметь два одинаковых столбца, т. е. $S_j = 0$. Таким образом,

$$\rho_\xi(\bar{\lambda}) = \xi_t \sum_{i=1}^t a_{it}(\bar{\lambda}) \xi_i = \xi_t (-1)^t d_t = d_{t-1} d_t \leq 0.$$

Значит, гиперплоскость $\rho_\xi(\lambda) = 0$ отделяет точку $\bar{\lambda}$ от области D . Описанный способ построения отделяющей гиперплоскости при рациональном $\bar{\lambda}$ и рациональных коэффициентах квадратичных форм в задаче

(1.1), (1.2) требует полиномиального числа операций. Недостатком этого способа есть слишком высокие требования к разрядности ЭВМ, на которой проводятся вычисления.

Для проверки положительной определенности матрицы $A(\bar{\lambda})$ не обязательно использовать правило Сильвестра, можно использовать процедуру выделения полных квадратов, соответствующую следующему разложению матрицы: $A(\bar{\lambda}) = BRB^T$, где $B(B^T)$ – треугольные матрицы, R – диагональная матрица. Процесс выделения полных квадратов ведем до тех пор, пока все диагональные элементы неразложенной части матрицы остаются положительными. Если процесс разложения проходит до конца, то $A(\bar{\lambda})$ – положительно-определенная матрица. Пусть процесс разложения заканчивается на k -ом шаге, $0 \leq k \leq n$. Тогда квадратичная форма с точностью до нумерации переменных представима в следующем виде:

$$(A(\bar{\lambda})x, x) = \sum_{i=1}^k r_{ii}(\bar{\lambda}) \left(x_i + \sum_{\{j:i < j \leq n\}} b_{ij}(\bar{\lambda})x_j \right)^2 + \Phi_{\bar{\lambda}}(x_{k+1}, \dots, x_n), \quad (2.6)$$

где $r_{ii}(\bar{\lambda}) > 0$, $i = 1, \dots, k$; $\Phi_{\bar{\lambda}}$ – квадратичная функция, у которой коэффициент при x_{k+1}^2 не превышает 0. Рассмотрим систему уравнений с треугольной матрицей

$$x_i + \sum_{\{j:i < j \leq n\}} b_{ij}(\bar{\lambda})x_j = 0; \quad i = 1, \dots, k; \quad (2.7)$$

$$x_{k+1} = 1; \quad x_{k+2} = x_{k+3} = \dots = x_n = 0, \quad (2.8)$$

и пусть $y(\bar{\lambda}) \neq 0$ – решение этой системы. В пространстве E^N векторов λ рассмотрим гиперплоскость

$$(A(\lambda)y(\bar{\lambda}), y(\bar{\lambda})) = 0. \quad (2.9)$$

(Заметим, что $A(\lambda)$ линейно зависит от λ , т. е. выражение (2.9) действительно задает гиперплоскость.) Из выражения (2.6) и определения $y(\bar{\lambda})$ (см. (2.7), (2.8)) следует, что $(A(\bar{\lambda})y(\bar{\lambda}), y(\bar{\lambda})) \leq 0$; с другой стороны, для $\lambda \in D$ $(A(\lambda)y(\bar{\lambda}), y(\bar{\lambda})) > 0$. Таким образом, гиперплоскость (2.9) отделяет точку λ от области D , т. е. ее можно использовать как отсекающую в методе эллипсоидов. На возможность такого способа построения отсекающей гиперплоскости указал автору А. С. Немировский. Хотя этот способ по форме отличается от предыдущего ([13]), но, по-видимому,

связан с близкими вычислительными проблемами неустойчивости алгоритма выделения полных квадратов, когда минимальное собственное число или несколько собственных чисел близки к 0.

б) Метод штрафных функций. Функция $\Psi(\lambda)$ за пределами области D в обычном смысле не определена: $\Psi(\lambda) = -\infty$, если $\lambda \notin \overline{D}$. Поэтому для решения задачи (2.3) нельзя применить метод внешних штрафных функций. Однако остается возможность применения барьерных штрафных функций, например, следующего вида:

$$S(\lambda, s) = \Psi(\lambda) - \frac{s_0}{\mu(\lambda)} - \sum_{i \in I} \frac{s_i}{\lambda_i},$$

где $S = \{s_0; s_i, i \in I\}$ – вектор с достаточно большими положительными компонентами. $S(\lambda, s)$ – вогнутая функция λ внутри области $D \cap \Lambda^+$. В любой внутренней точке λ области $D \cap \Lambda^+$ $S(\lambda, s) < \Psi(\lambda) \leq \overline{\Psi}$. Поэтому в качестве нижней оценки функционала задачи (1.1), (1.2) можно брать

$$\sup_{\lambda \in D \cap \Lambda^+} S(\lambda, s).$$

Так как $\mu(\lambda)$ – негладкая функция, то $S(\lambda, s)$ – также негладкая функция, стремящаяся к $-\infty$ при подходе к границе множества $D \cap \Lambda^+$ по внутренним точкам. Поэтому для нахождения

$$\sup_{\lambda \in D \cap \Lambda^+} S(\lambda, S)$$

нужно начинать из некоторой допустимой внутренней точки $\lambda^{(0)}$ и использовать один из монотонных методов негладкой оптимизации, обзор которых содержится в [6], [21].

Для решения задачи (2.3) можно применять также итеративный процесс типа метода «центров», решая на k -й итерации при заданной внутренней точке допустимой области $\lambda^{(k-1)}$ следующую задачу

$$\max_{\lambda} \left[\min \left[\Psi(\lambda) - \Psi(\lambda^{(k-1)}); \mu(\lambda); \lambda_i, i \in I \right] \right]. \quad (2.10)$$

В качестве $\lambda^{(k)}$ для следующей итерации нужно брать приближенное решение задачи (2.10), максимизацию в (2.10) можно осуществлять произвольным методом негладкой оптимизации. Специальный вариант метода «центров» описан также в [21].

в) Специальная форма r -алгоритма. Приведенные выше методы получения нижних оценок целевого функционала в задаче (1.1),

(1.2) недостаточно эффективны. Методы эллипсоидов при $N \geq 5$ сходятся медленно, методы барьерных функций и «центров» требуют вычисления минимальных собственных чисел матрицы $A(\lambda)$. Поэтому в качестве основного метода для получения оценок нами был выбран модифицированный метод обобщенных градиентов с растяжением пространства в направлении разности двух последовательных градиентов (r -алгоритм). Напомним, что в области $D \cap \Lambda^+$ нужно найти супремум вогнутой функции

$$\Psi(\lambda) = -\frac{1}{4}(A^{-1}(\lambda)l(\lambda), l(\lambda)) + c(\lambda),$$

причем λ входит в матрицу $A(\lambda)$, вектор $l(\lambda)$ в $c(\lambda)$ линейно. Ограничение $\lambda \in \Lambda^+$ легко учитывается путем введения негладкой функции штрафа вида

$$r(\lambda) = \sum_{i \in I} S(\lambda_i),$$

где

$$S(\lambda_i) = \begin{cases} 0, & \lambda_i \geq 0; \\ s\lambda_i, & \lambda_i < 0; \end{cases}$$

s – достаточно большое положительное число. Таким образом, будем рассматривать задачу нахождения $\sup_{\lambda \in D} \bar{\Psi}(\lambda)$, где $\bar{\Psi}(\lambda) = \Psi(\lambda) + r(\lambda) = -\frac{1}{4}(A^{-1}(\lambda)l(\lambda), l(\lambda)) + c(\lambda) + r(\lambda)$. Допустим, что $\lambda^{(0)}$ – начальная точка области D . $\sup_{\lambda \in D} \bar{\Psi}(\lambda)$ может достигаться либо в области $D(\lambda^* \in D)$, либо на границе этой области ($\lambda^* \in \bar{D} \setminus D$), причем второй случай практически встречается наиболее часто. Поэтому имеет смысл рассмотреть поведение функции $\Psi(\lambda)$ вблизи границы области D . Функция $c(\lambda)$ линейна, а $r(\lambda)$ – кусочно-линейна, поэтому в любой ограниченной области они ограничены. Особенности функции $\bar{\Psi}$ на множестве $\bar{D} \setminus D$ возникают в связи с поведением нелинейного выражения $(A^{-1}(\lambda)l(\lambda), l(\lambda))$. Пусть $\bar{\lambda} \in \bar{D} \setminus D$. Матрица $A(\bar{\lambda})$ вырождена. Обозначим размерность инвариантного подпространства $E_0(\bar{\lambda})$ матрицы $A(\bar{\lambda})$, соответствующего нулевому собственному значению, через $m(\bar{\lambda})$, $m(\bar{\lambda}) \geq 1$. Если проекция $l(\bar{\lambda})$ на $E_0(\bar{\lambda})$ отлична от 0, то

$$\lim_{\lambda \in D; \lambda \rightarrow \bar{\lambda}} (A^{-1}(\lambda)l(\lambda), l(\lambda)) = +\infty, \quad \text{т. е.} \quad \tilde{\Psi}(\lambda) = \lim_{\lambda \in D; \lambda \rightarrow \bar{\lambda}} \bar{\Psi}(\lambda) = -\infty.$$

Пусть $S_\varepsilon(\lambda)$ – ε -окрестность точки λ . Тогда, для того, чтобы

$$\Psi^+(\bar{\lambda}) = \lim_{\varepsilon \rightarrow 0} \sup_{D \cap S_\varepsilon(\bar{\lambda})} \Psi(\lambda) > -\infty,$$

необходимо, чтобы $l(\bar{\lambda})$ было ортогонально $E_0(\lambda)$. Это условие эквивалентно разрешимости уравнения

$$A(\bar{\lambda})x + l(\bar{\lambda}) = 0. \quad (2.11)$$

Пусть Q – подмножество симметричных матриц $A(\lambda)$, $\lambda \in \bar{D} \setminus D$, q – размерность этого множества, рассматриваемого как многообразие в пространстве симметричных матриц размерности $n(n+1)/2$. Условие разрешимости уравнения (2.11) может быть выражено в виде одного или нескольких алгебраических уравнений относительно $\bar{\lambda}$, и, вообще говоря, понижает размерность. Таким образом, размерность многообразия матриц $A(\bar{\lambda})$, принадлежащих Q , для которых $\Psi^+(\bar{\lambda}) > -\infty$ меньше q при $q \geq 1$. Это означает, что при движении из области D по некоторому направлению, ведущему в граничную точку $\bar{\lambda} \in \bar{D} \setminus D$, «вероятность» того, что функция $\Psi(\lambda)$ будет монотонно возрастать на открытом интервале, который заканчивается точкой $\bar{\lambda}$, равна 0.

Функция $\bar{\Psi}(\lambda)$ как бы «самоэкранируется» от выхода за пределы области своего определения при использовании алгоритмов монотонного подъема, т. е., как правило, отпадает необходимость в использовании барьерных штрафных функций. Появляется возможность использования монотонных вариантов r -алгоритма [6], [10]. Выход за пределы области D контролируется путем применения процедуры треугольного разложения матрицы $A(\lambda)$. Если на данном шаге произошел выход за пределы области D , то шаг дробится до тех пор, пока не будет получено допустимое значение. На основе указанной модификации r -алгоритма С. И. Стеценко разработал программу для решения некоторых булевых квадратичных задач и провел успешные вычислительные эксперименты. Более подробно об этом говорится в разд. 4.

3. Квадратичные экстремальные задачи с линейными ограничениями

Задачу минимизации квадратичной функции при линейных ограничениях (к.з.л.о.) запишем в следующем виде:

найти

$$\inf K_0(x), \quad K_0(x) = (A_0x, x) + (b, x) \quad (3.1)$$

при следующих ограничениях

$$l_j(x) = (l_j, x) + c_j \leq 0; \quad x = \{x_1, \dots, x_n\} \in E^n, \quad (3.2)$$

$l_j = \{l_j^{(1)}, \dots, l_j^{(n)}\} \in E^n$, c_j – константы, $j = 1, \dots, m$. Невыпуклые задачи квадратичного программирования возникают в ряде приложений. В частности, к ним сводятся проблемы линейной дополнителности, играющие большую роль в анализе макроэкономических моделей, задачи проектирования электронных модулей (чипов) на интегральных схемах и другие. В литературе описаны несколько подходов к решению такого рода задач, в частности, связанные с процедурой отсечения Бендерса (см. [18]), исследования в рамках общих схем решения многоэкстремальных задач или схем вогнутого программирования. В одной из недавних работ [18] в основе алгоритма лежит представление квадратичной формы в виде суммы выпуклой и вогнутой квадратичных функций и метод ветвей и границ. Для получения нижних оценок целевой функции вогнутая часть аппроксимируется линейной или кусочно-линейной функцией. В последнем случае задача оценки оказывается задачей смешанного типа: с непрерывными и булевыми переменными. Несмотря на использование мощной ЭВМ Cray 1S ($\sim 10^8$ операций в секунду), решение задач средней размерности (300 выпуклых переменных, 30 вогнутых, 50 ограничений) наталкивается на серьезные трудности. Так, достигнутая за счетное время 1500 с точностью всего порядка 3%. Поэтому актуальным остается вопрос о разработке новых практически эффективных алгоритмов.

Непосредственное применение предложенного нами выше двойственного подхода к получению нижних оценок целевой функции в общем случае к.з.л.о. ничего не дает, так как при изменении множителей Лагранжа меняется лишь линейная по x часть функции Лагранжа, квадратичная часть остается без изменений. Поэтому, если A_0 не является неотрицательно-определенной, то при любых множителях Лагранжа будем иметь тривиальную оценку: $\Psi(\lambda) = -\infty$.

Однако мы можем использовать линейные ограничения (3.2) для образования из них квадратичных следствий

$$K_{pj}(x) = -l_p(x)l_j(x) \leq 0; \quad p, j = 1, \dots, m. \quad (3.3)$$

Рассмотрим к.з. (3.1)–(3.3). Эта задача эквивалентна задаче (3.1), (3.2), однако формально содержит дополнительные ограничения, которые позволяют воздействовать на квадратичную часть функции Лагранжа. Обозначим множители Лагранжа, соответствующие ограниче-

ниям (3.2), через λ_{pj} ($p, j = 1, \dots, m$). Покажем, что при достаточно общих условиях множество $D \cap \Lambda^+$ задачи (3.1)–(3.3) не будет пустым.

Теорема. *Если выпуклая оболочка векторов l_j , $j = 1, \dots, m$, в ограничениях (3.2) содержит в качестве внутренней точки 0, то найдется неотрицательный вектор множителей Лагранжа $\{\lambda_{pj}\}_{p,j=1}^m$, при котором матрица $A(\lambda)$ становится положительно-определенной.*

Доказательство. Из условия теоремы следует, что произвольный вектор $y \in E^n$ может быть представлен в виде линейной комбинации с неотрицательными коэффициентами векторов l_j , $j = 1, \dots, m$. Пусть e_i – орт i -й координаты, $i = 1, \dots, n$. Тогда существует представление

$$e_i = \sum_{p=1}^m u_{ip}^{(+)} l_p; \quad -e_i = \sum_{q=1}^m u_{iq}^{(-)} l_q, \quad (3.4)$$

где $u_{ip}^{(+)}$, $u_{iq}^{(-)}$ – неотрицательные числа, $p, q = 1, \dots, m$; $i = 1, \dots, n$. Из (3.4) получаем соотношения

$$x_i = \sum_{p=1}^m u_{ip}^{(+)} (l_p, x); \quad -x_i = \sum_{q=1}^m u_{iq}^{(-)} (l_q, x), \quad i = 1, \dots, n.$$

Запишем тождества

$$x_i^2 = \sum_{p,q=1}^m u_{ip}^{(+)} u_{iq}^{(-)} \overline{K}_{pq}(x), \quad i = 1, \dots, n,$$

где $\overline{K}_{pq}(x) = -(l_p, x)(l_q, x)$ и отличается от $K_{pq}(x)$ лишь линейными членами ($p, q = 1, \dots, m$). Взяв вектор множителей Лагранжа

$$\{\lambda_{pq}\}_{p,q=1}^m = \left\{ \sum_{i=1}^n r_i u_{ip}^{(+)} u_{iq}^{(-)} \right\}_{p,q=1}^m,$$

где $r_i > 0$, $i = 1, \dots, n$, мы увеличим диагональный элемент матрицы $A_0 a_{ii}^0$ на r_i , оставляя недиагональные элементы без изменения. При достаточно больших r_i матрица функции Лагранжа станет положительно-определенной.

Следствие. *Если многогранное множество W , образованное ограничениями (3.2), непусто и ограничено, то найдутся неотрицательные множители Лагранжа $\{\lambda_{pq}\}_{p,q=1}^m$, при которых функция Лагранжа будет неотрицательно-определенной по x .*

Доказательство. Рассмотрим произвольный вектор $s \in E^n$, $\|s\| = 1$ и допустимую точку \bar{x}_0 . Тогда $\max_{1 \leq j \leq m} (s, l_j) > 0$, так как в противном случае весь луч $x(t) = \bar{x}_0 + st$, $t \geq 0$, состоял бы из допустимых точек, что противоречит ограниченности W . Отсюда следует, что

$$\min_{\{s, \|s\|=1\}} \max_{1 \leq j \leq m} (s, l_j) > 0,$$

но это условие выражает тот факт, что выпуклая оболочка $\{l_j\}_{j=1}^m$ содержит 0 в качестве своей внутренней точки, т. е. условие теоремы выполняется.

Таким образом, искусственный прием, связанный с введением ограничений (3.3), позволяет применить разработанную методику получения нижних оценок и к задачам с линейными ограничениями.

4. Булевы квадратичные задачи

Среди оптимизационных задач вида (1.1), (1.2) булевы квадратичные задачи (б.к.з.) выделяются наличием в числе ограничений (1.2) специальных булевых ограничений вида $x_j^2 - x_j = 0$; $j = 1, \dots, n$. Вводя эти ограничения с произвольными множителями Лагранжа в целевую функцию, можно произвольным образом менять диагональные элементы матрицы $A(\lambda)$, не изменяя недиагональные элементы. Легко подобрать множители Лагранжа так, чтобы $A(\lambda)$ стала положительно определенной; значит, множество D в задачах б. к. з. непусто.

Типичным примером б. к. з. являются задачи нахождения максимального взвешенного внутренне устойчивого множества графов, т. е. совокупности вершин графа, попарно не связанных ребрами, с максимальным суммарным весом.

Пусть задан граф $G(V, E)$, где V – множество вершин, E – множество ребер; $V = \{1, \dots, n\}$; ребра, соединяющие вершины $i, j \in V$, будем обозначать (i, j) . Сопоставим каждой вершине $i \in V$ булеву переменную x_i и вес $c_i > 0$. Тогда задача нахождения максимального взвешенного внутренне устойчивого множества (МВ ВУМ) состоит в нахождении

$$\alpha_c(G) = \max \sum_{i=1}^n c_i x_i \quad (4.1)$$

при ограничениях

$$x_i x_j = 0; \quad \forall (i, j) \in E; \quad (4.2)$$

$$x_i^2 - x_i = 0; \quad i = 1, \dots, n. \quad (4.3)$$

Задача МВ ВУМ играет большую роль в многочисленных приложениях: теория информации и кодирование, проектирование различных устройств при определенных условиях несовместности; она тесно связана с известными задачами выбора, разбиения множеств, раскраски графов и другими комбинаторными задачами, имеющими, в свою очередь, массу приложений. С другой стороны, она принадлежит к классу NP -полных задач (при целочисленных весах вершин). О сложности этой задачи говорит то, что даже в частном случае, когда все $c_i = 1$, $i = 1, \dots, n$, не найден полиномиальный алгоритм, который бы гарантировал (сколь угодно большую) фиксированную относительную погрешность решения по функционалу для любого графа $D(V, E)$.

Обычно для решения задач нахождения МВ ВУМ используются методы «ветвей и границ», практическая эффективность которых во многом определяется алгоритмами получения оценок целевой функции.

Для задачи МВ ВУМ в связи с исследованиями по теории кодирования в [12, 13], [17] были предложены несколько эквивалентных оценок. Например, следующая: пусть \overline{m}_G – класс неотрицательно-определенных матриц $B = \{b_{ij}\}_{i,j=1}^n$, для которых $b_{ij} = 0$, если $i \neq j$; $(i, j) \in E$; $\sum_{i=1}^n b_{ii} \leq 1$; остальные элементы произвольные.

Тогда

$$\nu(G) = \max_{B \in \overline{m}_G} \sum_{i,j=1}^n \sqrt{c_i c_j} b_{ij} \geq \alpha_c(G). \quad (4.4)$$

В [4] показано, что описанный в данной статье двойственный подход к получению оценок в к.з. дает для задачи (4.1)–(4.3) оценку, совпадающую по величине с $\nu(G)$. Автор совместно с С. И. Стеценко разработал алгоритм, основанный на методе «ветвей и границ», в котором в качестве оценок используются описанные в данной статье оценки вида (2.3), и провели вычислительные эксперименты. Нахождение оценок производилось с помощью специальной модификации метода субградиентного типа с растяжением пространства – r -алгоритма (см. разд. 3). Эксперименты производились на графах с числом вершин от 30 до 60 при средней степени вершин 4, ребра генерировались случайным образом, веса представляли собой случайные целые числа, большие 20 и меньшие 40. Как видно из [4], размерность задачи по x равна $|V|$ – числу вершин, а по λ – числу ребер. Во всех экспериментах было получено

оптимальное решение, причем первая оценка отличалась от оптимального значения не более чем на 5%, благодаря чему количество вершин в дереве ветвлений было небольшим (от 3 – 5 при $|V| = 30$ до нескольких десятков при $|V| = 60$). Время счета на ЕС-1060 менялось от 1,5 – 2 мин. при $|V| = 30$ до 1 часа при $|V| = 60$. Эксперименты показали устойчивую работу алгоритма оценки и хорошее качество оценок.

5. Экстремальные задачи на классе неотрицательно-определенных матриц

Получение оценок в к.з. (1.1), (1.2) приводит к специальным экстремальным задачам на классе положительно-определенных матриц. К такого же типа задачам приводит вычисление оценки Ловаса (4.4), статистические задачи обработки результатов экзаменов [14, 15] и ряд других задач обработки результатов наблюдений, связанных с корреляционными матрицами. Так как множество неотрицательно-определенных матриц представляет собой выпуклое множество с негладкой границей, при решении такого рода задач по существу оказываются полезными методы недифференцируемой оптимизации. Попытки решения этих задач классическими методами нелинейного программирования наталкиваются на серьезные трудности [14], [15], [13].

Нами накоплен опыт решения ряда экстремальных задач на классе неотрицательно-определенных матриц в связи с проблемами оценки хроматического класса графов и хроматического числа графов специального вида ([3], [8], [20]). Вычислительные эксперименты показали, что r -алгоритм в сочетании с методом точных негладких штрафных функций, если целевая функция определена на всем пространстве, является достаточно надежным и эффективным средством решения такого рода задач.

Таким образом, недифференцируемая оптимизация открывает хорошие перспективы решения широкого круга практически важных экстремальных задач на классе неотрицательно-определенных матриц.

Литература

1. Зойтендейк Г. Методы возможных направлений. – М.: Изд-во иностр. лит., 1963.

2. Кюнци Г. П., КРЕЛЛЕ В. Нелинейное программирование. – М.: Сов. радио, 1963.
3. СТЕЦЕНКО С. И. Вычислительные эксперименты по решению квадратичных задач с булевыми переменными. – В кн.: Методы решения сложных задач математического программирования. (Сб. научных трудов ИК АН УССР.) Киев: ИК АН УССР, 1985.
4. СТЕЦЕНКО С. И., ШОР Н. З. Связь оценок Ловаса с двойственными оценками в квадратичных булевых задачах. – В кн.: Методы решения нелинейного и дискретного программирования. (Тр. семинара ИС по кибернетике АН УССР.) Киев: ИК АН УССР, 1984.
5. ШОР Н. З. Метод отсечения с растяжением пространства для решения задач выпуклого программирования // Кибернетика. – 1977. – № 1. – С. 94–95.
6. ШОР Н. З. Методы минимизации недифференцируемых функций и их приложения. – К.: Наук. думка, 1979. – 199 с.
7. ШОР Н. З., ГЕРШОВИЧ В. И. Метод эллипсоидов, его обобщения и приложения. // Кибернетика. – 1982. – №5. – С. 61–69.
8. ШОР Н. З., СТЕЦЕНКО С. И. Получение оценок для хроматического класса графов и использованием моделей булева квадратичного программирования // Тезисы Республиканского семинара по дискретной оптимизации (г. Ужгород). – К., 1985. – С. 135–136.
9. ШОР Н. З., ДАВЫДОВ А. С. О методе получения оценок в квадратичных экстремальных задачах с булевыми переменными // Кибернетика. – 1985. – № 2. – С. 48–50.
10. ШОР Н. З., АЛИЕВ И. Т. Об одной монотонной модификации r -алгоритма. – В кн.: Методы решения сложных задач математического программирования. (Сб. научных трудов ИК АН УССР.) Киев: ИК АН УССР, 1985. – С. 45–50.
11. ЮДИН Д. Б., НЕМИРОВСКИЙ А. С. Информационная сложность и эффективные методы решения выпуклых экстремальных задач. // Экономика и мат. методы. – 1976. – Т. 12. – № 2.
12. ELIESE M., RODEMICH E. R., RUMSEY JR. H. G. The Lovász bound and generalization. // Journal of Combinatorics, Information and System Sciences. – 1978. – V. 3.

13. GRÖTSCHEL M., LOVASZ L., SCHRIYVER A. Polynomial algorithms for perfect graphs: (Rep.) // Institut für ökonometrie und operations research. – WP 81176-OR. – Bonn, 1981. – 40 p.
14. FLETCHER R. A nonlinear programming problem in Statistics (educational testing) // SIAM J. Sci. Stat. Comput. – 1981. – V. 2. – № 3.
15. FLETCHER R. Semi-definite matrix constraints in optimization // SIAM J. of Control and Optimization. – 1985. – V. 23. – № 4.
16. KORNER F., RICHTER CL. Zur effektiven lösung von booleschen quadratischen optimierungs problemen // Numer. Math. – 1982. – V. 40.
17. LOVASZ L. On the Shannon Capacity of a graph // IEEE Trans. Inform. Theory – 1979. – V. 25. – № 1.
18. PARDALOS P. M., GLICK J. H., ROSEN G. B. Global minimization of indefinite quadratic problems. // Report CS-85-31. – The Pennsylvania State University. Department of Computer Sciences, 1985.
19. WOLFE P. The simplex method for quadratic programming // Econometrica. – 1959. – № 27.
20. SHOR N. Z., STECENKO S. J. Quadratic boolean problems and Lovász's bounds // Abstracts JFIP conference on system modelling and optimization. – Budapest, 1985.
21. KIWIEL K. C. Methods of Descent for Nondifferentiable Optimization. – Lecture Notes in Mathematics. – Berlin: Springer-Verlag, 1985. – 1133:362.

Поступила в редакцию 21.04.86

Об одном подходе к получению глобальных экстремумов в полиномиальных задачах математического программирования

Н. З. Шор

Кибернетика. — 1987. — № 5. — С. 102–106.

Описан способ получения оценок снизу целевой функции в полиномиальных задачах математического программирования путем сведения к квадратичным экстремальным задачам. Для случая минимизации полиномов от одной переменной доказывается, что соответствующие оценки являются точными.

Рассмотрим евклидово пространство E^n с векторами $z = \{z_1, z_1, \dots, z_n\}$. Под полиномиальной задачей математического программирования будем понимать задачу следующего вида: найти

$$\inf P_0(z) \tag{1}$$

при ограничениях

$$P_i(z) = 0, \quad i = 1, \dots, m, \tag{2}$$

где $P_0(z)$, $P_i(z)$, $i = 1, \dots, m$ — полиномиальные функции от z .

То, что ограничения (2) заданы в форме равенств, существенно не ограничивает общность задачи, так как любое полиномиальное неравенство вида $R(z) \leq 0$ можно свести к полиномиальному равенству $R(z) + t^2 = 0$, где t — дополнительная переменная.

Вводя новые переменные и используя квадратичные подстановки вида $z_i^2 = y_i$, $z_{jk} = z_j z_k$, $y_i^2 = v_i$ и т. п., можно снизить степень полиномов в (1), (2) до квадратичной, рассматривая их как функции от расширенного множества переменных, при этом появляются новые квадратичные равенства, соответствующие указанным выше подстановкам. Таким образом, любую задачу вида (1)–(2) можно свести к квадратичной экстремальной задаче: найти

$$\inf_x K_0(x), \quad x \in E^{\bar{n}}, \quad \bar{n} \geq n, \tag{3}$$

при ограничениях:

$$K_i(x) = 0, \quad i = 1, \dots, \bar{m}; \quad \bar{m} \geq m. \quad (4)$$

где $K_\nu(x)$, $\nu = 0, 1, \dots, \bar{m}$ – квадратичные функции.

В работе [1] для получения оценок снизу оптимального значения целевой функции предложен метод, который будем называть двойственным, так как он использует множители Лагранжа.

Рассмотрим функцию Лагранжа: $L(x, u) = K_0(x) + \sum_{i=1}^{\bar{m}} u_i K_i(x) = (A(u)x, x) + (l(u), x) + c(u)$. Здесь $A(u)$ – симметричные матрицы $\bar{n} \times \bar{n}$, $l(u)$ – векторы размерности \bar{n} , $c(u)$ – константы, зависящие от вектора множителей Лагранжа $u = \{u_1, \dots, u_{\bar{m}}\}$. Пусть Ω ($\bar{\Omega}$) – множество таких значений $u = \{u_1, \dots, u_{\bar{m}}\}$, при которых матрица $A(u)$ положительно определенная (неотрицательно определенная). При $u \in \Omega$ $\min_x L(x, u) = \psi(u)$ достигается в некоторой точке $x(u)$, являющейся решением линейной системы уравнений

$$2A(u)x + l(u) = 0. \quad (5)$$

Определим $\psi(u) = L(x(u), u)$ при $u \in \Omega$. Для любого допустимого \bar{x} $L(\bar{x}, u) = K_0(\bar{x})$, поэтому $\psi(u) \leq K_0(\bar{x})$ для произвольного u . Отсюда $\psi^* = \sup_{u \in \Omega} \psi(u) \leq K_0(\bar{x})$ для произвольного допустимого \bar{x} , т. е. $\psi^* \leq f^*$,

где f^* – оптимальное значение целевой функции в задаче (3)–(4) (если система уравнений (4) несовместна, будем считать, что $f^* = +\infty$), $\psi(u)$ – вогнутая функция на $\bar{\Omega}$, $\bar{\Omega}$ – выпуклое, либо пустое множество. В последнем случае будем считать, что $\psi^* = -\infty$.

Задача нахождения $\psi^* = \sup_{u \in \Omega} \psi(u)$ относится к выпуклому программированию. Для нее разработаны довольно эффективные алгоритмы, в частности, хорошие практические результаты дал r -алгоритм со специальной регулировкой шагового множителя [2]. Задачи вида (1)–(2) и соответствующие им (3)–(4) могут быть невыпуклыми и многоэкстремальными. Поэтому нельзя в общем случае гарантировать равенство $\psi^* = f^*$. Одним из подходов к получению глобального оптимума может быть метод ветвей и границ с использованием указанных выше двойственных оценок. Примером такого подхода может служить алгоритм нахождения максимального взвешенного независимого множества вершин графа, описанный в [3]. Однако метод ветвей и границ не для всех задач достаточно эффективен. Интересно исследовать такие классы полиномиальных невыпуклых задач, для которых $\psi^* = f^*$. Отметим, что

при переходе от полиномиальных задач к квадратичным можно использовать различные квадратичные подстановки и в зависимости от этого будет меняться ψ^* . Кроме того, можно генерировать новые ограничения, которые являются алгебраическими следствиями прежних, оставляя область допустимых решений без изменения. При этом двойственные оценки не убывают, а в некоторых случаях возрастают, и увеличиваются шансы совпадения ψ^* и f^* .

Легко доказывается следующая теорема.

Теорема 1. Если $\psi^* = \sup_{u \in \Omega} \psi(u)$ достигается на множестве Ω , то $\psi^* = f^*$.

Доказательство. Градиент функции $\psi(u)$ при $u \in \Omega$ совпадает с вектором невязок

$$g_{\psi}(u) = \{K_i(x(u))\}_{i=1}^{\overline{m}}.$$

Если максимум $\psi(u)$ достигается в некоторой точке $u^* \in \Omega$, то $K_i(x(u^*)) = 0$, $i = 1, \dots, \overline{m}$, т. е. $x(u^*)$ — допустимая точка, при этом $f^* \leq K_0(x(u^*)) = \psi(u^*) = \psi^*$. Но ψ^* — нижняя оценка для f^* . Отсюда следует, что $f^* = \psi^*$, что и требовалось доказать.

Таким образом, если $\psi^* < f^*$, то супремум $\psi(u)$ достигается на границе области Ω . Пусть u^* — точка на $\overline{\Omega} \setminus \Omega$, в любой окрестности которой при заданном произвольно $\varepsilon > 0$ найдутся точки $u \in \Omega$, для которых справедливо $\psi^* - \psi(u) < \varepsilon$. При этом ранг матрицы $A(u^*)$ должен совпадать с рангом присоединенной матрицы $(A(u^*)|l(u^*))$, т. е. система (5) должна иметь при $u = u^*$ решение. Поясним сказанное на примерах.

I. Найти $\min [(Ax, x) + (c, x)]$ при ограничениях $(x, x) - 1 = 0$, $x \in E^n$.

Приведем матрицу A к диагональному виду путем ортогонального преобразования. Получаем в новых переменных $y = \{y_1, \dots, y_n\}$ следующую задачу: найти $\min K_0(y)$, где

$$K_0(y) = \sum_{i=1}^n \lambda_i y_i^2 + \sum_{i=1}^n a_i y_i, \quad (6)$$

при ограничениях

$$K_1(y) = \sum_{i=1}^n y_i^2 - 1 = 0. \quad (7)$$

Пусть $\lambda_1, \dots, \lambda_n$ — собственные числа матрицы A , записанные в неубывающем порядке с учетом их кратности, и минимальное собственное число имеет кратность 1. Рассмотрим функцию Лагранжа

задачи (6)–(7):

$$L(y, u) = K_0(y) + uK_1(y).$$

При $u > -\lambda_1$ $L(y, u)$ является положительно определенной. $L(y, u)$ можно переписать в следующем виде:

$$L(y, u) = \sum_{i=1}^n (\lambda_i + u) \left(y_i + \frac{a_i}{2(\lambda_i + u)} \right)^2 - \sum_{i=1}^n \frac{a_i^2}{4(\lambda_i + u)} - u.$$

При $u + \lambda_1 > 0$

$$\psi(u) = \min_x L(x, u) = -\frac{1}{4} \sum_{i=1}^n \frac{a_i^2}{\lambda_i + u} - u.$$

Если $a_1 \neq 0$, то при $u > -\lambda_1$, $u \rightarrow -\lambda_1$, $\psi(u) \rightarrow -\infty$, при $u \rightarrow +\infty$, $\psi(u) \rightarrow -\infty$. Таким образом, $\max \psi(u)$ достигается в области положительной определенности: $-\lambda_1 < u^* < +\infty$, и ему соответствует глобальный экстремум задачи (6)–(7). Если $a_1 = 0$, то $\max_u \psi(u)$ может достигаться как на границе области положительной определенности ($u + \lambda_1 = 0$), так и внутри, при этом $\psi^* = f^*$. Аналогичные рассуждения проходят и в случае кратных собственных чисел матрицы A .

II. Рассмотрим задачу минимизации полинома 4-й степени от одной переменной

$$\min(x^4 + ax^2 + bx).$$

Произведем подстановку: $x^2 - y = 0$. Получаем следующую квадратичную задачу:

$$\min(y^2 + ay + bx)$$

при ограничении $x^2 - y = 0$. Функция Лагранжа имеет следующий вид (при $u > 0$):

$$\begin{aligned} L(x, y, u) &= y^2 + ay + bx + u(x^2 - y) = \\ &= \left(y + \frac{a-u}{2} \right)^2 + u \left(x + \frac{b}{2u} \right)^2 - \left(\frac{a-u}{2} \right)^2 - \frac{b^2}{4u}. \end{aligned}$$

При $u > 0$ (в области положительной определенности по y, x функции Лагранжа) $\psi(u) = -\left(\frac{a-u}{2} \right)^2 - \frac{b^2}{4u}$. При $b \neq 0$, $u \rightarrow +0$, $\psi(u) \rightarrow -\infty$.

При $u \rightarrow +\infty$, $\psi(u) \rightarrow -\infty$. Значит, $\max_{u>0} \psi(u)$ достигается в точке u^* , $0 < u^* < +\infty$ области $\Omega : u > 0$ и ему будет соответствовать единственный глобальный оптимум $x = -\frac{b}{2u^*}$. При $b = 0$, $a \geq 0$, имеем $u^* = a$, $\psi(u^*) = 0$, $x^* = 0$. При $b = 0$, $a < 0$, $\sup_{u>0} \psi(u)$ достигается в точке на границе Ω , $u^* = 0$, при этом $y^* = x^{*2} = -\frac{a}{2}$, $\psi^* = -\frac{a^2}{4}$, что соответствует двум различным глобальным минимумам: $x^* = \pm\sqrt{-\frac{a}{2}}$. Таким образом, во всех случаях $\sup_{u>0} \psi(u) = \psi^* = f^*$.

Оказывается, что результат, полученный для многочленов 4-й степени, можно обобщить на многочлены от одной переменной произвольной четной степени. Рассмотрим многочлен четной степени $2n$, $n \geq 1$, от переменной x_1 , у которого коэффициент при старшей степени равен 1:

$$P_{2n}(x_1) = x_1^{2n} + \sum_{k=1}^{2n} a_{2n-k} x_1^{2n-k}. \quad (8)$$

Введем обозначения: $x_k = x_1^k$; $k = 0, 1, \dots, n$. В этих обозначениях (заметим, что $x_0 = 1$) задача минимизации $P_{2n}(x_1)$ превращается в квадратичную экстремальную задачу следующего вида:

найти минимум

$$K_{2n}(x) = x_n^2 + \sum_{k=1}^{n-1} a_{2n-k} x_n x_{n-k} + \sum_{i=0}^n a_{n-i} x_{n-i} \quad (9)$$

при ограничениях:

$$x_p x_q - x_r x_s = 0, \quad p + q = r + s \leq 2n - 2, \quad p \geq q, \quad p > r \geq s, \quad (10)$$

где p, q, r, s – неотрицательные целые числа. Выражение, стоящее в левой части равенств вида (10), обозначим $R(p, q; r, s)$, множество различных допустимых четверок $(p, q; r, s)$ в (10) обозначим Q_{2n} .

Замечание. Очевидно, что среди равенств вида (10) часть является алгебраически избыточными, некоторые равенства являются даже линейными следствиями других, именно их можно отбросить без ущерба для дальнейших рассуждений. Мы сознательно сохраняем все множество выражений $R(p, q; r, s)$ для упрощения обозначений и доказательств.

Каждому равенству вида $R(p, q; r, s) = 0$ сопоставим множитель Лагранжа $\lambda(p, q; r, s)$. Совокупность множителей Лагранжа образует вектор λ .

Пусть $x = (x_1, \dots, x_n)$, $x_0 = 1$.

Образуем функцию Лагранжа квадратичной задачи (9)–(10):

$$L(x, \lambda) = K_{2n}(x) + \sum_{(p, q; r, s) \in Q_{2n}} \lambda(p, q; r, s) R(p, q; r, s).$$

Пусть $\Omega(P_{2n})$ ($\bar{\Omega}(P_{2n})$) – множество таких векторов λ , при которых квадратичная по x функция $L(x, \lambda)$ является положительно (неотрицательно) определенной. Функция $\psi(\lambda) = \min_x L(x, \lambda)$ определена при $\lambda \in \Omega(P_{2n})$. Если супремум $\psi(\lambda)$ при $\lambda \in \bar{\Omega}(P_{2n})$ совпадает с $f^* = \min_{x_1} P(x_1)$, то будем говорить, что P_{2n} обладает ω -свойством. Если этот супремум достигается на $\lambda^* \in \Omega(P_{2n})$, то будем говорить, что полином P_{2n} обладает сильным ω -свойством, при этом (см. теорему 1) $\psi(\lambda^*)$ совпадает с f^* , т. е. P_{2n} обладает ω -свойством.

Покажем, что ω -свойство сохраняется при сдвиге начала координат. Итак, справедлива

Теорема 2. (О сдвиге). Если $P(x_1)$ обладает ω -свойством, то и полином $P(x_1) = P(x_1 + a)$ обладает ω -свойством при произвольном a .

Доказательство. Введем переменную $\bar{x}_1 = x_1 - a$. Тогда $\bar{P}(\bar{x}_1) = P(x_1)$. Таким образом, переход от полинома P к \bar{P} сводится к подстановке: $x_1 = \bar{x}_1 + a$ в $P(x_1)$. Если определить $\bar{x}_k = \bar{x}_1^k$, $k = 1, \dots, n$, то для x_k получим следующие выражения:

$$\bar{x}_k = (\bar{x}_1 + a)^k = \sum_{i=0}^k C_k^i a^i \bar{x}_{k-i}. \quad (11)$$

Для продолжения доказательства нам понадобятся две леммы.

Лемма 1. Пусть задана система равенств вида

$$y_i - y_j = 0; \quad i, j = 1, \dots, k; \quad i \neq j. \quad (12)$$

Тогда любая линейная форма $l(y) = \sum_{i=1}^k c_i y_i$ при условии $\sum_{i=1}^n c_i = 0$ может быть записана в виде линейной комбинации левых частей равенств (12).

Доказательство. При $k = 1$ лемма тривиальна. Для $k > 1$ легко ее доказать индукцией от n к $n + 1$, используя простое тождество

$$\sum_{i=1}^{n+1} c_i y_i = \sum_{i=1}^{n-1} c_i y_i + (c_n + c_{n+1}) y_n + c_{n+1} (y_{n+1} - y_n).$$

Лемма 2. Любое выражение вида $x_{\bar{s}} x_{\bar{q}} - x_{\bar{p}} x_{\bar{r}}$; $\bar{s} + \bar{q} = \bar{p} + \bar{r}$, где \bar{s} , \bar{q} , \bar{p} , \bar{r} — целые неотрицательные числа, при подстановках вида (11) переходят в выражения от переменных $\bar{x}_0, \bar{x}_1, \dots, \bar{x}_n$, представляющие собой линейную комбинацию выражений вида $\bar{x}_s \bar{x}_q - \bar{x}_p \bar{x}_r$, $s + q = p + r$; s, q, p, r — целые неотрицательные числа.

Доказательство. Используем равенства (11):

$$\begin{aligned} x_{\bar{s}} x_{\bar{q}} - x_{\bar{p}} x_{\bar{r}} &= \left(\sum_{i=0}^{\bar{s}} a^i C_{\bar{s}}^i \bar{x}_{\bar{s}-i} \right) \left(\sum_{j=0}^{\bar{q}} a^j C_{\bar{q}}^j \bar{x}_{\bar{q}-j} \right) - \\ &\quad - \left(\sum_{i=0}^{\bar{p}} a^i C_{\bar{p}}^i \bar{x}_{\bar{p}-i} \right) \left(\sum_{j=0}^{\bar{r}} a^j C_{\bar{r}}^j \bar{x}_{\bar{r}-j} \right) = \\ &= \sum_{t=0}^{\bar{s}+\bar{q}} a^t \left(\sum_{i,j:i+j=t} C_{\bar{s}}^i C_{\bar{q}}^j \bar{x}_{\bar{s}-i} \bar{x}_{\bar{q}-j} - \sum_{i,j:i+j=t} C_{\bar{p}}^i C_{\bar{r}}^j \bar{x}_{\bar{p}-i} \bar{x}_{\bar{r}-j} \right). \end{aligned}$$

Как следует из биномиальных тождеств и равенства $\bar{s} + \bar{q} = \bar{p} + \bar{r}$,

$$\sum_{i,j:i+j=t} C_{\bar{s}}^i C_{\bar{q}}^j - \sum_{i,j:i+j=t} C_{\bar{p}}^i C_{\bar{r}}^j = C_{\bar{s}+\bar{q}}^t - C_{\bar{p}+\bar{r}}^t = 0.$$

Используя лемму 1 многократно при различных t , получаем, что выражение $x_{\bar{s}} x_{\bar{q}} - x_{\bar{p}} x_{\bar{r}}$ представимо в виде линейной комбинации выражений вида $\bar{x}_s \bar{x}_q - \bar{x}_p \bar{x}_r$, $s + q = p + r$, s, q, p, r — целые неотрицательные числа. Лемма доказана.

Продолжим доказательство теоремы.

Полином $P(x_1)$ обладает ω -свойством. Это значит, что найдется вектор множителей Лагранжа λ^* такой, что функция Лагранжа $L(x, \lambda)$ квадратичной задачи (9)–(10) при $\lambda = \lambda^*$ является неотрицательно определенной по x и $\psi(\lambda^*) = f^*$, где $f^* = \min_{x_1} P(x_1)$. Подставим в $L(x, \lambda^*)$ вместо компонент вектора x выражения из (11). (Формулы (11) можно рассматривать как переход от одной системы координат к другой.) Получим квадратичную относительно \bar{x} функцию $\bar{L}(\bar{x}, \lambda^*)$, и так

как положительно (неотрицательно) определенная квадратичная функция при неособых линейных преобразованиях координат переходит в положительно (неотрицательно) определенную функцию, то $\overline{L}(\overline{x}, \lambda^*)$ будет неотрицательно определенной. Заметим, что области значений $L(x, \lambda^*)$ и $\overline{L}(\overline{x}, \lambda^*)$ совпадают, откуда $\overline{\psi}^* = \min_{\overline{x}} \overline{L}(\overline{x}, \lambda^*) = f^*$. С другой стороны, используя лемму 2, приходим к выводу, что выражение $\overline{L}(\overline{x}, \lambda^*)$ может быть записано в форме функции Лагранжа квадратичной задачи, соответствующей нахождению $\min \overline{P}(\overline{x}_1)$:

$$\min \overline{K}(\overline{x}) \quad (13)$$

при ограничениях

$$\overline{x}_p \overline{x}_q - \overline{x}_r \overline{x}_s = 0, \quad (p, q; r, s) \in Q_{2n}, \quad (14)$$

где $\overline{K}(\overline{x})$ – результат подстановки в $K_{2n}(x)$ (9) вместо компонент вектора x выражений (11); при этом соответствующие множители Лагранжа линейно выражаются через компоненты λ^* . Таким образом, для задачи (13), (14) существует вектор множителей Лагранжа $\overline{\lambda}^*$, при котором соответствующая квадратичная функция от \overline{x} неотрицательно определена и ее минимум совпадает с f^* значением минимума полиномов $P(x_1)$ и $\overline{P}(\overline{x}_1)$. Значит, полином \overline{P} также обладает ω -свойством. Теорема 2 о сдвиге доказана.

Перейдем к доказательству основного результата.

Теорема 3. *Любой полином $P_{2n}(x_1)$ вида (8) четной степени обладает ω -свойством.*

Доказательство. Для $n = 2$ ω -свойство доказано (см. пример 2). Проведем доказательство теоремы индукцией по n . Пусть ω -свойство справедливо для полиномов вида (8) степени $2n$. Докажем его справедливость для полиномов степени $2(n+1)$. По теореме 2 (о сдвиге) ω -свойство сохраняется при сдвиге аргумента, поэтому можно предположить, не уменьшая общности, что глобальный минимум находится в 0 и значение полинома в 0 равно 0. Отсюда будет следовать, что у рассматриваемого полинома $P_{2n+2}^0(x_1)$ коэффициенты при младших степенях $a_0 = a_1 = 0$ и $a_2 \geq 0$, т. е. он представим в виде $P_{2n+2}^0(x_1) = x_1^2 \cdot P_{2n}(x_1)$, причем полином $P_{2n}(x_1)$ принимает неотрицательные значения на всей оси. По предположению индукции для $P_{2n}(x_1)$ выполняется ω -свойство, т. е. для соответствующей квадратичной задачи найдется вектор множителей Лагранжа λ^* , при котором функция Лагранжа $L_{2n}(x, \lambda^*)$ будет принимать неотрицательные значения при

любом x , т. е. может быть записана в виде $L_{2n}^*(x) = \sum_{i=1}^n l_i(x) + r^2$, где $l_i(x)$ – линейные функции от (x_1, \dots, x_n) , r^2 – неотрицательная константа.

Рассмотрим выражение $x_1^2 L_{2n}^*(x)$. Ему соответствует квадратичная форма

$$L_{2n+2}^*(\bar{x}) = \sum_{i=1}^n \bar{l}_i^2(\bar{x}) + r^2 x_1^2,$$

где $\bar{l}_i(\bar{x})$ – линейная функция от переменных x_1, \dots, x_n, x_{n+1} , следующим образом получающаяся из l_i : если $l_i(x) = \sum_{j=1}^n l_{ij} x_j + l_{i0}$, то

$$\bar{l}_i(\bar{x}) = \sum_{j=0}^n l_{ij} x_{j+1}.$$

Легко видеть, что $L_{2n+2}^*(\bar{x})$ получается из функции Лагранжа квадратичной задачи, соответствующей минимизации полинома $P_{2n+2}^0(x_1)$ при тех же значениях множителей Лагранжа, которые образуют вектор λ^* , однако эти значения относятся к «преобразованным» ограничениям: ограничениям вида $x_p x_q - x_r x_s = 0$, относящимся к задаче минимизации $P_{2n}(x_1)$, будут соответствовать ограничения $x_{p+1} x_{q+1} - x_{r+1} x_{s+1} = 0$ в задаче минимизации $P_{2n+2}^0(x_1)$. Минимальное значение $L_{2n+2}^*(\bar{x})$ достигается при $\bar{x} = 0$ и равно 0. Таким образом, при определенных значениях множителей Лагранжа квадратичной задачи, соответствующей минимизации полинома $P_{2n+2}^0(x_1)$, получается точная оценка, т. е. полином $P_{2n+2}^0(x_1)$ обладает ω -свойством. По теореме о сдвиге этим же свойством будет обладать произвольный полином $P_{2n+2}(x_1)$ степени $2n + 2$ с коэффициентом, равным 1, при старшем члене. Теорема 3 доказана.

Сделаем несколько замечаний.

1. Сведение задачи минимизации $P_{2n}(x_1)$ к квадратичной задаче неоднозначно, это связано с неоднозначностью представления x_1^k , например, $x_1^5 = x_1 x_4 = x_2 x_3 = x_5$ и т. п. В зависимости от принятого представления изменяются оптимальные множители Лагранжа, однако ω -свойство от конкретного представления не зависит.

2. Минимальное число ограничений, которые нужно учитывать при формулировке эквивалентной квадратичной задачи при минимизации $P_{2n}(x_1)$, равно $(n - 1)$, так как нужно определить переменные x_2, \dots, x_n . Остальные ограничения избыточны, т. е. их добавление не сужает области допустимых решений. Роль избыточных ограничений состоит в расширении числа двойственных переменных функции Лагранжа, что

приводит, вообще говоря, к более точным оценкам. Добавление ограничения, которое является линейной комбинацией имеющихся, не отражается на точности двойственных оценок, так как «вклад» этого ограничения в функцию Лагранжа эквивалентен определенному изменению множителей Лагранжа при имеющихся ограничениях.

3. Избыточность числа ограничений приводит, как правило, к неоднозначности вектора оптимальных множителей Лагранжа λ^* . По геометрическому смыслу оптимальных множителей Лагранжа – это коэффициенты разложения по градиентам ограничений ортогонального к многообразию, вырезаемого ограничениями, антиградиента целевой функции в оптимальной точке. При неоднозначности разложения совокупность допустимых векторов коэффициентов образует линейное многообразие. Пересечение этого многообразия с $\bar{\Pi}(P)$ и дает множество оптимальных векторов Лагранжа.

Весьма правдоподобным кажется предположение, что при минимизации полиномиальной функции многих переменных $P(x_1, \dots, x_n)$ путем построения эквивалентной квадратичной задачи при использовании фиксированного множества избыточных ограничений получается точная двойственная оценка значения глобального минимума (аналог теоремы 3). Если это предположение удастся доказать, то в арсенале методов нахождения глобального минимума полиномиальных функций от нескольких переменных появится метод, основанный на двойственных оценках. Этот метод можно в различных формах комбинировать с методом ветвей и границ и методом релаксации ограничений. Еще больший интерес представляет обобщение наметченной в данной статье теории на общий класс полиномиальных задач с ограничениями вида (1), (2). Можно ли в общем случае дать способ генерации избыточной системы ограничений, чтобы гарантировать, что двойственная оценка окажется точной? В настоящее время идет экспериментальное исследование эффективности предложенной методики двойственных оценок на ряде классов экстремальных задач теории графов и невыпуклых квадратичных задач с линейными ограничениями [3].

Литература

1. Шор Н. З., Давыдов А. С. О методе получения оценок в квадратичных экстремальных задачах с булевыми переменными // Кибернетика. – 1985. – № 2. – С. 48–50.

2. ШОР Н. З. Методы минимизации недифференцируемых функций и их приложения. – К.: Наук. думка, 1979. – 199 с.
3. СТЕЦЕНКО С. И. Об алгоритмах решения некоторых задач выбора // Исследование методов решения экстремальных задач: сб. науч. тр. – Киев: ИК АН УССР, 1986. – С. 12–16.

Поступила в редакцию 14.07.86

Об одном классе оценок глобального минимума полиномиальных функций

Н. З. Шор

Кибернетика. — 1987. — № 6. — С. 9–11.

Задача минимизации полиномиальной функции сводится к квадратичной оптимизационной задаче, для которой строятся двойственные оценки. Приводятся необходимые и достаточные условия их точности.

Пусть E^n — n -мерное евклидово пространство, на котором определена полиномиальная вещественная функция $P(x) = P(x_1, \dots, x_n)$. Рассмотрим задачу нахождения $f^* = \inf_{x \in E^n} P(x_1, \dots, x_n)$. Если $f^* > -\infty$, т. е. $P(x)$ ограничена снизу, то назовем ее ОС-функцией. Так как полином нечетной степени от одной переменной может принимать сколь угодно большие по абсолютной величине положительные и отрицательные значения, то для того, чтобы $P(x)$ была ОС-функцией, необходимо, чтобы для любого $i \leq n$ старшие степени s_i переменных x_i были четными, а значения коэффициентов при них при любых значениях остальных переменных принимали неотрицательные значения. ОС-функциями должны также являться функции $P(x; x_i = c)$ от $(n-1)$ -переменной, которые получаются из $P(x)$, если придать произвольной переменной x_i ($i \leq n$) определенное значение $x_i = c$. Проверка того, что данная полиномиальная функция $P(x)$ является ОС-функцией, — далеко не тривиальная задача.

Пусть $s_i = 2l_i$; $i = 1, \dots, n$. Рассмотрим целочисленные векторы $\alpha = \{\alpha_1, \dots, \alpha_n\}$ с неотрицательными элементами и одночленные выражения вида

$$R[\alpha] = x^{\alpha_1} \dots x_n^{\alpha_n}, \quad \text{где } 0 \leq \alpha_i \leq l_i; \quad i = 1, \dots, n. \quad (1)$$

Имеем систему тождественных соотношений

$$R[\alpha_1^{(1)}, \dots, \alpha_n^{(1)}]R[\alpha_1^{(2)}, \dots, \alpha_n^{(2)}] - R[\alpha_1^{(3)}, \dots, \alpha_n^{(3)}]R[\alpha_1^{(4)}, \dots, \alpha_n^{(4)}] = 0 \quad (2)$$

(или $R[\alpha^{(1)}]R[\alpha^{(2)}] - R[\alpha^{(3)}]R[\alpha^{(4)}] = 0$) при $\alpha_i^{(1)} + \alpha_i^{(2)} = \alpha_i^{(3)} + \alpha_i^{(4)} \leq s_i$, $i = 1, \dots, n$. Любой полином $P(x) = P(x_1, \dots, x_n)$ со старшими степенями при x_i , равными $s_i = 2l_i$, $i = 1, \dots, n$, может быть неоднозначно, вообще говоря, записан в виде квадратичной функции от переменных $R[\alpha^{(\nu)}]$ (см. (1)) следующего вида:

$$P(x) = K(R, \lambda) = \sum_{i,j} c_{ij} R[\alpha^{(i)}]R[\alpha^{(j)}] + \\ + \sum_{(k,l;m,n)} \lambda_{k,l;m,n} \left(R[\alpha^{(k)}]R[\alpha^{(l)}] - R[\alpha^{(m)}]R[\alpha^{(n)}] \right),$$

где $R[\alpha^{(i)}]R[\alpha^{(j)}]$ – некоторое представление одночлена полинома $P(x)$ в виде произведения одночленов вида (1); c_{ij} – соответствующие коэффициенты; $\lambda_{k,l;m,n}$ – произвольные множители при левых частях тождественных соотношений (2). С другой стороны, $K(R, \lambda)$ можно рассматривать как функцию Лагранжа квадратичной экстремальной задачи: минимизировать

$$K(R) = \sum_{i,j} c_{ij} R[\alpha^{(i)}]R[\alpha^{(j)}]$$

при ограничениях (2), при этом $\lambda_{k,l;m,n}$ могут интерпретироваться как множители Лагранжа. Таким образом, задача безусловной минимизации полинома сведена к задаче минимизации квадратичной функции $K(R)$ при квадратичных ограничениях (2).

Пусть множество тех $\lambda = \{\lambda_{k,l;m,n}\}$, для которых квадратичная функция $K(R, \lambda)$ является неотрицательно определенной по R , не пусто. Обозначим его $\bar{\Lambda}^+$. Внутреннюю часть этого множества (совокупность тех λ , при которых $K(R, \lambda)$ положительно определена) обозначим Λ^+ . Заметим, что при $\lambda \in \Lambda^+$ $\inf_R K(R, \lambda)$ достигается в некоторой точке $R^*(\lambda)$. Легко видеть, что $\Lambda^+(\bar{\Lambda}^+)$ – выпуклые множества. Пусть

$$\psi(\lambda) = \inf_R K(R, \lambda). \quad (3)$$

Так как $K(R, \lambda)$ линейна по λ , то $\psi(\lambda)$ в области своего существования (в частности, на Λ^+) – вогнутая функция. Так как при любом λ и допустимом векторе R $K(R, \lambda) = K(R)$, то

$$\psi(\lambda) = \inf_R K(R, \lambda) \leq f^*. \quad (4)$$

Если $P(x)$ не является ОС-функцией, то Λ^+ пусто. Пусть $\bar{f} = \sup_{\lambda \in \bar{\Lambda}^+} \psi(\lambda)$.

Из (4) следует, что $\bar{f} \leq f^*$.

Если найдется $\lambda^* \in \bar{\Lambda}^+$ такой, что $\psi(\lambda^*) = f = f^* > -\infty$, то будем говорить, что $P(x)$ и соответствующая квадратичная задача вида (3), (2) обладает ω -свойством.

Если полином $P(x)$ обладает ω -свойством, то для получения оценки снизу для f^* с произвольной точностью нужно решать задачу выпуклого программирования: найти $\sup_{\lambda \in \bar{\Lambda}^+} \psi(\lambda) = -\inf_{\lambda \in \bar{\Lambda}^+} (-\psi(\lambda))$. В [1, 2] показано, что для решения задачи такого типа можно применять метод эллипсоидов. Практически хорошо работают при решении подобных задач другие методы недифференцируемой оптимизации, например, определенные модификации r -алгоритма.

Таким образом, для полиномиальных функций, обладающих ω -свойством, получаем неперереборный алгоритм оценки глобального экстремума. Отметим также, что если $\sup_{\lambda \in \bar{\Lambda}^+} \psi(\lambda)$ достигается на $\lambda^* \in \Lambda^+$, то градиент функции $\psi(\lambda)$ в этой точке, равный вектору невязок в ограничении (2), при $R = R(\lambda^*)$ равен 0, т. е. все ограничения (2) выполняются. Но это значит, что $\bar{K}(R(\lambda^*), \lambda^*) = K(R(\lambda^*)) = f^*$, т. е. $\bar{f} = f^*$ и ω -свойство выполняется. Таким образом, если ω -свойство не выполняется, то $\sup_{\lambda \in \bar{\Lambda}^+} \psi(\lambda)$ должен достигаться на границе области неотрицательной определенности.

Оказалось, что вопрос о том, обладает ли данная полиномиальная ОС-функция ω -свойством, тесно связан с возможностью ее представления в виде суммы квадратов полиномов. Докажем следующую теорему.

Теорема 1. *Для того, чтобы полином $P(x) = P(x_1, \dots, x_n)$, принимающий минимальное значение в точке 0 и равный при этом 0, обладал ω -свойством, необходимо и достаточно, чтобы он был представим в виде суммы квадратов полиномов.*

Доказательство. 1. Необходимость. Пусть полином обладает ω -свойством. Тогда найдется $\lambda^* \in \bar{\Lambda}^+$, при котором $\bar{K}(R, \lambda^*)$ неотрицательно определена как функция от R и $\bar{K}(0, \lambda^*) = 0$. Но любую неотрицательно определенную функцию можно представить в виде суммы квадратов линейных функций. Подставив в функцию $\bar{K}(R, \lambda^*)$, представленную в форме суммы квадратов, вместо переменных $R[\alpha^{(\nu)}]$ их выражения в форме одночленов, получим искомое представление $P(x)$ в форме суммы квадратов полиномов.

2. Достаточность. Пусть $P(x)$ представим в виде суммы квадратов полиномов $P_1(x), \dots, P_m(x)$:

$$P(x) = \sum_{i=1}^m P_i^2(x).$$

Полиномы $P_k(x)$, $k = 1, \dots, m$, не должны содержать свободных членов, иначе $P(0) > 0$, а это противоречит $P(0) = 0$. Подставив вместо каждого одночлена в полиномах $P_k(x)$ соответствующее $c_a R[\alpha]$, получим однородную неотрицательно определенную форму относительно R , минимум которой при $R = 0$ равен 0. Сравнив коэффициенты при $R[\alpha^{(i)}]R[\alpha^{(j)}]$ в представлении $P(x)$ в форме суммы квадратов $\overline{K}(R)$ с коэффициентом при $R[\alpha^{(i)} + \alpha^{(j)}]$ в представлении $P(x) = K(R) = \overline{K}(R, 0)$, найдем $\overline{\lambda}$, при котором $\overline{K}(R) = K(R, \overline{\lambda})$. Таким образом, найдется $\overline{\lambda}$ такой, что $\psi(\overline{\lambda}) = \min_R K(R, \overline{\lambda}) = K(0, \overline{\lambda}) = 0$. Значит, ω -свойство выполняется.

Теорема доказана.

Пусть $\min P(x_1, \dots, x_n) = f^*$ и достигается при $x = x^*$.

Рассмотрим полином $P_0(x) = P(x - x^*) - f^*$. $P_0(x)$ принимает свое минимальное значение в точке 0 и это значение равно 0.

Графики полиномов $P_0(x)$ и $P(x)$ совпадают при соответствующем сдвиге в пространстве x и по f . Легко показать, что сдвигу в пространстве координат x соответствует линейное невырожденное преобразование в пространстве координат $R[\alpha]$. Так как при таких преобразованиях сохраняется неотрицательная (положительная) определенность квадратичных функций, то и сохраняется ω -свойство, а также возможность представления функций в форме суммы квадратов. Таким образом, из теоремы 1 вытекает более общая.

Теорема 2. *Для того, чтобы ограниченный снизу полином $P(x)$ обладал ω -свойством, необходимо и достаточно, чтобы полином $P(x) - f^*$ был представим в виде суммы квадратов полиномов.*

К сожалению, не всякая полиномиальная функция, принимающая неотрицательные значения, представима в виде суммы квадратов полиномов. Этот вопрос рассматривался еще в 1888 году молодым в то время Д. Гильбертом [3]. Он изучал представимость в форме суммы квадратов однородных полиномов (форм) четной степени m с числом переменных n (m – максимальная сумма степеней переменных, входящих в одночлен данного полинома; для однородных полиномов сумма степеней переменных, входящих в произвольный одночлен, равна m).

Гильберт показал, что при $n=3$, $m \geq 6$ и $n \geq 4$, $m \geq 4$ найдутся неотрицательные формы, не представимые в виде суммы квадратов форм.

Лишь для некоторых общих классов форм, а именно:

- 1) $m = 2$, n – произвольное число (квадратичные формы);
- 2) $n = 2$; m – четное произвольное, т. е. произвольные формы от двух переменных (им соответствуют произвольные полиномы от одной переменной);

3) $m = 4$, $n = 3$, т. е. произвольные биквадратные формы с тремя переменными (им соответствуют произвольные полиномы 4-й степени с двумя переменными) – вопрос о представимости в виде суммы квадратов неотрицательных форм решается положительно.

Отсюда возникла следующая задача, известная как 17-я проблема Гильберта [4]. Рассматривается поле рациональных функций от вещественных переменных. Если рациональная функция в области своего определения неотрицательна, то всегда ли она представима в виде суммы квадратов рациональных функций? Окончательный положительный ответ на этот вопрос был дан в работе Э. Артина в 1927 г. [5].

Для неотрицательных полиномиальных форм $P(x)$ это означает, что всегда можно найти такую представимую в виде суммы квадратов неотрицательную форму $P_0(x)$, что $P_0(x)P(x)$ представим в виде суммы квадратов полиномов. Интересным остается вопрос о конструктивном способе выбора $P_0(x)$.

После краткого отступления вернемся к рассмотрению некоторых важных частных случаев экстремальных полиномиальных задач, для которых ω -свойство справедливо.

Системы полиномиальных уравнений. Пусть задана система уравнений

$$P_i(x) \equiv P_i(x_1, \dots, x_n) = 0; \quad i = 1, \dots, m. \quad (5)$$

Рассмотрим $P(x) = \sum_{i=1}^m P_i^2(x)$.

Если система (5) совместна, то $f^* = \min P(x) = 0$; $P(x) - f^* = P(x)$ представим в виде суммы квадратов, т. е. $P(x)$ обладает ω -свойством. Множество решений системы (5) в случае ее нелинейности может быть пустым, континуальным или дискретным. Лишь в исключительных случаях решение единственно. Поэтому при минимизации $P(x)$ путем сведения к квадратичной задаче оптимальная оценка будет достигаться, как правило, на границе Λ^+ при вырожденной матрице квадратичной формы, что не позволит непосредственно получить хотя бы одно из решений. Для получения конкретного решения нужно использовать дополнительные средства для отделения корней (например, дополнительные неравенства на переменные, использование ε -возмущения коэффициентов функции $P(x)$ для получения приближения к решению и т. п.).

Интересным является вопрос – всегда ли в случае несовместности

системы (5) при минимизации

$$P(x) = \sum_{i=1}^m P_i^2(x) \quad \sup_{\lambda \in \Lambda^+} \psi(\lambda) > 0?$$

Задача о линейной дополнителности. Данная задача [6] ставится следующим образом: определить вектор $x = \{x_1, \dots, x_n\} \in E^n$, удовлетворяющий системе неравенств $x \geq 0$; $Ax + b \geq 0$ (A – матрица $n \times n$, b – n -мерный вектор), и такой, что $(x, Ax + b) = 0$. Пусть a_i , $i = 1, \dots, n$, – строки матрицы A . Введем дополнительные переменные v_i , $i = 1, \dots, n$, и рассмотрим задачу: найти

$$f^* = \min_{x \geq 0; v \geq 0} \left\{ \sum_{i=1}^n [(a_i, x) + b_i - v_i]^2 + c \sum_{i=1}^n x_i v_i \right\}, \quad (6)$$

где $v = \{v_1, \dots, v_n\}$, $c > 0$. Легко видеть, что $f^* \geq 0$ и равно 0 тогда и только тогда, когда соответствующая задача о линейной дополнителности имеет решение. Задача (6) путем подстановок $x_i = y_i^2$; $v_i = \omega_i^2$, $i = 1, \dots, n$, сводится к задаче безусловной минимизации функции $P(y, \omega)$, являющейся полиномом 4-й степени от компонент y , ω и представимой в виде суммы квадратов. Таким образом, если задача линейной дополнителности имеет решение, то $P(y, \omega)$ обладает ω -свойством. Задача о линейной дополнителности является частным случаем задачи минимизации произвольной квадратичной функции при линейных ограничениях. Если эти ограничения имеют форму $x \geq 0$, то можно добавить квадратичные ограничения $x_i \cdot x_j \geq 0$, $i, j = 1, \dots, n$, являющиеся следствием ограничений $x \geq 0$, применить к расширенной задаче метод множителей Лагранжа и построить двойственные оценки, следуя описанной в статье методике. Хотя эти оценки не всегда являются точными, однако они обычно оказываются более точными, чем оценки, полученные путем линеаризации вогнутых составляющих квадратичной функции [7].

Частный случай теоремы 2 для полиномов от одной переменной приводится в [8].

Литература

1. Шор Н. З., Давыдов А. С. О методе получения оценок в квадратичных экстремальных задачах с булевыми переменными // Кибернетика. – 1985. – № 2. – С. 48–50.

2. ШОР Н. З. Квадратичные оптимизационные задачи // Техн. кибернетика. – 1987. – № 1. – С. 128–139.
3. HILBERT D. Über die darstellung definiten Formen als Summen von Formen quadraten // Math. Ann. – 1988. – 22. – P. 342–350.
4. Проблемы Гильберта / Под ред. П. С. Александрова. – М.: Наука, 1969. – 238 с.
5. ARTIN E. Über die Zerlegung definiten Funktionen in Quadrate. – 1927. – 5. – S. 100–115.
6. STONE R. Linear complementarity problems with an invariant number of Solutions // Math. Programming – 1968. – 34. – № 3. – P. 265–291.
7. PARDALOS P. M., GLICK J. H., ROSEN J. B. Global minimisation of indefinite quadratic problems. – S. l. 1985. – 14 p. – (Reprint / Pennsylv. State University. – NCS–85–31).
8. ШОР Н. З. Об одном подходе к получению глобальных экстремумов в полиномиальных задачах математического программирования // Кибернетика. – 1987. – № 5. – С. 102–106.

Поступила в редакцию 05.01.87

Задачи минимизации матричных функций и недифференцируемая оптимизация

Н. З. Шор

Обозрение 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 ограничениях

$$\Phi_\rho(A(u)) \leq 0, \quad \rho = 1, \dots, r. \quad (2)$$

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 обработаны модификации этих алгоритмов, в которых удается с помо-

щью 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 $\mathcal{A}(G)$ 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 разделе 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 матрица A — из класса симметричных вещественных матриц S_n

размерности $n \times n$:

$$\psi_k(A) = \max_{Y \in M_n^k} \left\{ \text{tr}(AYY^T) \right\}, \quad k \leq n, \quad (5)$$

где $\text{tr}Z$ означает след (сумму диагональных элементов) матрицы Z , M_n^k — класс таких прямоугольных матриц Y размера $n \times k$, что $Y^TY = I_k$ (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 ее 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 $YY^T \in \text{co}\{ZZ^T\}$, $Z \in M_n^k$, то формула останется верной (co — 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 x , 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 внутренности $M_{k,n}^{C_0}$. 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, $|V_i| = m_k$, $i = 1, \dots, k$. Недиагональные 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 smoothed 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. Следовательно, 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 квадратичную форму

$$K_A(y) = (Ay, y) = \sum_{i,j=1}^n a_{ij} y_i y_j.$$

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

$$-\varphi(x) \leq 0 \quad (14)$$

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 были не очень успешны [14]–[16], в них ограничение на неотрицательную определенность матрицы сводилось к ограничению на неотрицательность минимального собственного числа матрицы, и применялась стандартная методика нелинейного программирования. В данном 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

$$x_i x_j = 0 \quad \text{для всех } \{i, j\} \in E \quad (19)$$

и булевости переменных

$$x_k^2 - x_k = 0 \quad \text{для всех } k \in V, \quad (20)$$

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, dom ψ 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 двойствен-

ную оценку (21). Подобная оценка может быть получена для произвольной задачи квадратичного типа, имеющей 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

$$\varphi(H) = \frac{\left(\sqrt{\lambda_1(H)}\right)^{3/2}}{\sqrt{\lambda_n(H)} \lambda_n [-(A^T H + H A)]},$$

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

$$\text{dom } \psi = \text{dom } f \cap \{x : f(x) > 0\} \quad (32a)$$

при условии, что правая часть (32a) непуста.

Легко проверить, что составляющие функции $\psi(H)$ удовлетворяют условию леммы. Если $H \in G$, то $\lambda_1(H)$ и $\lambda_n(C)$ принимают положительные значения. Отсюда следует, что функция $\psi(H)$ является вогнутой при $H \in G$.

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

$$\min [- \ln \lambda_n(H) - 2 \ln \lambda_n [-(A^T H + H A)]], \quad (33)$$

$$\lambda_1(H) \leq 1, \quad (34)$$

$$H \succ 0; \quad -(A^T H + H A) \succ 0. \quad (35)$$

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 (33) 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 $\text{po} M^*$ — 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

$$\min_{K \succ 0} (-\ln \det K) \quad (36)$$

при ограничениях

$$(K a_i, a_i) \leq 1; \quad i = 1, \dots, k. \quad (37)$$

Как показано в [44], эта задача сводится к задаче безусловной минимизации негладкой выпуклой функции:

$$\varphi(K) = -\ln \det K + n \max\{0, \max_{1 \leq i \leq k} \{(K a_i, a_i) - 1\}\}, \quad K \succ 0. \quad (37a)$$

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]:

$$\max_u \left\{ \ln \det(AU A^T) : U = \text{diag } u; \quad \sum_{i=1}^k u_i = n; \quad u \geq 0 \right\}, \quad (38)$$

где $u = \{u_1, \dots, u_k\}$ суть множители Лагранжа, относящиеся к ограничениям (37); 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 (36)–(37), 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 (36)–(37) (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 максимального по объему вписанного эллипсоида с центром в точке 0 сводится к задаче вида (37а), которую мы умеем решать.

Если же центр вписанного эллипсоида не фиксирован, то получаем следующую задачу нелинейного программирования [44], [29]: найти

$$\min_{Q \succ 0} (-\ln \det Q) \quad (39)$$

при ограничениях

$$\sqrt{(Qa_i, a_i)} + (a_i, b) - 1 \leq 0, \quad i = 1, \dots, k. \quad (40)$$

При подстановке $P = \sqrt{Q}$, $P \succ 0$, имеем задачу выпуклого программирования: найти

$$\min_{P \succ 0} (-2 \ln \det P), \quad (39a)$$

$$|Pa_i| + (a_i, b) - 1 \leq 0, \quad i = 1, \dots, k, \quad (40a)$$

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]

$$(Qa_i, a_i) \leq (1 - (b, a_i))^2 \quad (41)$$

заменяются линейными неравенствами вида

$$(Qa_i, a_i) \leq (1 - (b, a_i)) (1 - (\bar{c}, a_i)) \quad i = 1, \dots, k, \quad (42)$$

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, решении 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 $A_0, A_1, \dots, \dots, A_5$ 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., HOFFMAN 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 BuI., 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, Ithaca, 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.

Р А З Д Е Л П

Двойственные оценки Шора

В раздел включены статьи:

- 2.1. Использование модификации r -алгоритма для нахождения
глобального минимума полиномиальных
функций (*Шор, Стецюк, 1997*) 94
- 2.2. Использование r -алгоритма в задачах полуопределенного
программирования (*Шор, Стецюк, 1997*) 123
- 2.3. Нахождение глобальных минимумов полиномиальных функций с .
использованием двойственных квадратичных
оценок (*Шор, Стецюк, Крылов, 1997*) 130
- 2.4. Dual Solution of Quadratic-Type Problems by
 r -algorithm (subroutine DSQTPr) (*Shor, Stetsyuk, 2001*) 148
- 2.5. Lagrangian bounds in multiextremal polynomial and discrete
optimization problems (*Shor, Stetsyuk, 2002*) 150

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.

Рассмотрены свойства квадратичных двойственных оценок для полиномиальных функций, алгоритмы нахождения этих оценок, основанные на методах негладкой оптимизации. Приведены результаты тестовых экспериментов для экстремальных задач, которые описываются в виде моделей нелинейного программирования квадратичного типа. Показано, что использование избыточных ограничений в ряде случаев существенно улучшает двойственные оценки.

Определение глобального минимума полинома от нескольких переменных – весьма сложная задача уже для полиномов четвертой степени при числе переменных более пяти. 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. 1), алгоритмы нахождения этих оценок, основанные на методах негладкой оптимизации (разд. 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. 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 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 i ($i = 1, \dots, n$), а их тотальная степень не должна превышать $\frac{d(P)}{2}$, т.е. если

$$P(y) = \sum_{k=1}^N P_k^2(y),$$

то $P_k(y) \in \pi(n; l; \frac{d(P)}{2})$, $k = 1 \dots, N$; $l = (l_1, \dots, l_n)$.

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 $\overline{P}(y) \in \overline{\pi}(n; 2l; d(P))$ соответствует квадратичная форма

$$\overline{P}(y) \sim \sum_{\alpha \in A} c_{\overline{P}}[\alpha] R[\beta^{(1)}] R[\beta^{(2)}],$$

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 $\overline{P}(y)$.

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

$\overline{M}[n; l; \frac{d(P)}{2}]$, квадратичный полином от переменных $R[\beta]$, стоящий в левой части тождества (2), обозначим $K_{\beta^{(3)}, \beta^{(4)}}(\rho)$ (здесь ρ – вектор переменных $\{R(\beta)\}_{\beta \in \overline{M}[n, l, \frac{d(P)}{2}]}$).

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 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 $x_i^j x_2^k x_3^l$ 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 $\text{Lin}Y_P(0) = k < 3$ (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 $P(\bar{y}) = 0$. Пусть $\overline{P}(y) = P_{\bar{t}}(y)$. 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 в точке $u^* \in \Omega$, 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 за пределы 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

$$\bar{u} = \left\{ (u_i^{(k)})_{i=1, \dots, k}^*; \overbrace{0, \dots, 0}^{(m-k) \text{ раз}} \right\}.$$

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 регуляризацией 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 a -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 вычислительных 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 перемен-

ных (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 $x_0^* = x_0$ 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 последовательных субградиентов в преобразованном пространстве аргументов $Y_k = A_k X$, т.е. $r_k = g_{\varphi_k}(y_k) - g_{\varphi_k}(y_{k-1}) = B_k^T(g_f(x_k) - g_f(x_{k-1}))$).

$$1.3. \xi_k = \begin{cases} 0, & \text{если } \|r_k\| \leq \varepsilon_0, \\ \frac{r_k}{\|r_k\|}, & \text{если } \|r_k\| > \varepsilon_0. \end{cases} \quad (\text{нормировка вектора } r_k \text{ для выпол-}$$

нения 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 $\varepsilon_g = 10^{-6}$ 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 $\tilde{h}_l = h_k q_2^t, t = \lfloor \frac{l}{n_h} \rfloor$ и $\lfloor \cdot \rfloor$ — 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 точки

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 $r(\alpha)$ -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 $r(\alpha)$ -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 $r(\alpha)$ -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

$$\text{supergrad}_u \lambda_n[Q(u)] = \sum_{i,j}^n \frac{\partial^2 K_l(x)}{\partial x_i \partial x_j} \xi_i \xi_j, l = 1, \dots, m, \quad (19)$$

где $\xi = \{\xi_r\}_{r=1}^n$ – 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); $n_{\tilde{\psi}}$ – 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 $\overline{P}(x_1, x_2, x_3)$ 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 $y^* = (y_1^*, y_2^*)$:

$$\varepsilon_1^* = \max \left\{ \left| \frac{\partial P(y_1, y_2)}{\partial y_1}(y_1^*, y_2^*) \right|, \left| \frac{\partial P(y_1, y_2)}{\partial y_2}(y_1^*, y_2^*) \right| \right\}. \quad (20)$$

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_r^* 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_r^* приближается к границе области Ω . Результаты вычислений для полинома $P_\gamma(y_1, y_2)$ 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 $\varepsilon_2^* = \max \left\{ \frac{\varepsilon_1^*}{|y_1^*|^3}, \frac{\varepsilon_1^*}{|y_2^*|^3} \right\}$ – 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_r^* приближается 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

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

является довольно устойчивой (для всех тестовых задач из табл. 2. срабатывал критерий останова по отклонению аргумента).

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 различных $\varepsilon = \{0.1, 0.01, 0.001\}$ для полинома $P_{a_1, a_2}(y_1, y_2)$ 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

Табл. 2. Тесты для полинома $P_\gamma(y_1, y_2)$ при $\gamma \rightarrow 2$.

γ	$iter(n_{\bar{y}}, n_\lambda)$	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)	---	---	---	---	$-1.08 * 10^{-2}$

Табл. 3. Тесты для полинома $P_{a_1, a_2}(y_1, y_2)$.

a_1	a_2	$iter(n_{\hat{y}}, n_{\lambda})$	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

задачи (21)–(23), соответствующей графу G_1^n , равно $n/2$, если n – четное, и $\frac{n \cos(\pi/n)}{1 + \cos(\pi/n)}$, если n – нечетное. Учитывая, что для графа G_1^n 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 $\lfloor \frac{n}{2} \rfloor$ переменных (здесь $\lfloor \cdot \rfloor$ – 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 $\lfloor \frac{n}{2} \rfloor$ -ой переменной расстояние равно $\lfloor \frac{n}{2} \rfloor$.

Вычислительные затраты модификации $r(\alpha)$ -алгоритма (в терминах числа итераций, общего количества вычислений суперградиента $\psi(u)$ и количества вычислений суперградиента $\psi(u)$ 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 $\psi(u_r^*)$, 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 $\lfloor \frac{n-2}{2} \rfloor$ раз

$u_0 = (1, \overbrace{1, \dots, 1}^{\lfloor \frac{n-2}{2} \rfloor})$ – 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]

$$\nu_1^*(G) = \min_y \lambda_1(A(y)), \quad (24)$$

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)

где $\lambda_1(A)$ – максимальное собственное число матрицы A , $A(y)$ – класс симметричных матриц, у которых (i, j) -й элемент принимает произвольное 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, между которыми

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)	–	–	–

расстояние по Хемингу равно единице, для второй переменной расстояние равно двум, для третьей переменной – трем. Оптимальное значение оценки $\nu_1^*(G_2^{64}) = 16/3$. Нахождение 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 количество вычислений функции и субградиента. По итерациям $r(\alpha)$ -алгоритма в табл. 5 приведено (в графе 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 эк-

спериментов 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 $\tilde{\psi}(u)$, вычисляемого по формуле (18), может стремиться к нулю, а вычисляемого по формуле (19) к нулю не стремится. В то же время величина шага h_k одинакова для обоих направлений, что приводит к неравномерным сдвигам по этим направлениям и препятствует останову алгоритма 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 про-

странства можно использовать и одноранговые операторы из [21], которые обеспечивают преобразование по типу r -алгоритмов и позволяют проводить внешнюю аппроксимацию множества экстремумов монотонно уменьшающимся по объему эллипсоидом. 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. ЖУРВЕНКО 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.

Поступила 13.05.97

Использование r -алгоритма в задачах полуопределенного программирования

Н. З. Шор, П. И. Стецюк

*Праці міжнародної конференції „Питання оптимізації обчислень“
(6–8 жовтня 1997 р., м. Київ) – С. 330–335.*

Большое число докладов на 16-м Международном симпозиуме по математическому программированию (24–29 августа 1997г.) было посвящено моделям полуопределенного программирования (semidefinite programming) и развитию методов внутренних точек применительно к решению такого рода задач. Как бы в тени оказалась возможность решения задач полуопределенного программирования с помощью методов негладкой оптимизации, ведь условие неотрицательной определенности некоторой симметричной матрицы A эквивалентно тому, что минимальное собственное число этой матрицы $\lambda_n(A) \geq 0$ ($n \times n$ – размерность матрицы A). Но $\lambda_n(A)$ – вогнутая недифференцируемая функция элементов матрицы, т.е. если элементы матрицы $A(x) = \{a_{ij}\}_{i,j=1}^n$ являются линейными функциями от вектора варьируемых параметров $x \in R^m$, то условие $A(x) \geq 0$ эквивалентно выпуклому негладкому ограничению $\varphi(x) = -\lambda_n(A(x)) \leq 0$.

По нашему мнению, многие задачи полуопределенного программирования целесообразно рассматривать как частный случай задач недифференцируемой оптимизации и применять для их решения практически эффективные общие методы минимизации негладких выпуклых функций.

Основной метод, который мы применяем для решения задач негладкой оптимизации, – это метод субградиентного типа с растяжением пространства в направлении разности двух последовательных субградиентов, так называемый r -алгоритм [1]. Этот метод отличается высокой эффективностью при решении самых разнообразных задач: 1) при реализации схем декомпозиции по переменным и по ограничениям; 2) при использовании негладких штрафных функций для нахождения оптимального решения в общих задачах выпуклого программирования; 3) для нахождения двойственных Лагранжевых квадратичных оценок в невыпуклых и дискретных задачах оптимизации и т.п. Упомянутые в п. 3) задачи, как правило, являются задачами полуопределенного программирования.

Как показывает опыт решения разнообразных задач негладкой оптимизации с помощью r -алгоритма, для получения относительной погрешности (по целевой функции) не более $\varepsilon > 0$ при минимизации функции от n переменных достаточно провести $O(n \ln(\frac{1}{\varepsilon}))$ итераций r -алгоритма, каждая итерация требует в среднем не более трех обращений к подпрограмме вычисления значения целевой функции и субградиента в определенных точках и порядка $O(n^2)$ арифметических операций, связанных с линейными преобразованиями векторов и матриц (см. [2]). Отметим, что r -алгоритм легко распараллеливается (см. [3]). В силу этих свойств r -алгоритм может оказаться вполне конкурентоспособным при сравнении с методами внутренних точек.

Мы провели ряд тестовых экспериментов, связанных с решением задач полуопределенного программирования, возникающих при нахождении двойственных оценок функционала в ряде экстремальных задач на графах, таких, как максимальный разрез графа (MaxCut), оптимальная бисекция графа (BiSection) и минимальное взвешенное по ребрам разреза разбиение вершин графа на k частей с заданным числом вершин в каждой части (MinPartK). Все эксперименты проводились с использованием одной из версий $r(\alpha)$ -алгоритма, описанной в [2]. Результаты экспериментов показали хорошие качества $r(\alpha)$ -алгоритма: достаточно высокую скорость сходимости, устойчивость вычислений вблизи границы области определения минимизируемой функции, возможность получения оценок с высокой точностью.

Более детально рассмотрим применение $r(\alpha)$ -алгоритма для задачи MaxCut.

Для графа $G(V, E)$ задача MaxCut состоит в таком разбиении множества вершин V ($|V| = n$) на два непересекающихся подмножества V_1 и V_2 , чтобы сумма весов тех ребер из E , концевые вершины которых находятся в разных подмножествах, была максимальна. Согласно [4], MaxCut представляется следующей задачей нелинейного программирования квадратичного типа:

$$\nu_w(V_1, V_2) = \max \left(1/2 \sum_{(i,j) \in E} w_{ij} - 1/4 \sum_{(i,j) \in E} w_{ij} y_i y_j \right), \quad (1)$$

$$y_i^2 = 1, \quad i = 1, \dots, n, \quad (2)$$

где w_{ij} – веса ребер $(i, j) \in E$.

Двойственная верхняя Лагранжева квадратичная оценка φ_{up} для

задачи (1)–(2) получается путем решения следующей задачи: найти

$$\varphi_{up} = 1/2 \sum_{(i,j) \in E} w_{ij} + 1/4 \psi^*, \quad (3)$$

где

$$\psi^* = \min_u \psi(u) = \min_u \left(\sum_{i=1}^n u_i + t \max(0, -\lambda_n(B_w(u))) \right). \quad (4)$$

Здесь t – значение штрафного множителя ($t \geq n$), симметричная матрица $B_w(u)$ размером $n \times n$, такая, что

$$B_w(u) = \text{diag}(u) + W, \quad (5)$$

где W – симметричная матрица размером $n \times n$, элементы которой равны w_{ij} , если $(i, j) \in E$, и нулю, если $(i, j) \notin E$.

Функция $\psi(u)$ – выпуклая и негладкая, для нахождения ее минимума можно применять $r(\alpha)$ -алгоритм. Используя нормированный собственный вектор, соответствующий собственному числу $\lambda_n(B_w(u))$, который требуется при вычислении субградиента $\psi(u)$, можно указать нетрудоемкую по вычислениям эвристическую процедуру построения допустимого целочисленного решения задачи (1)–(2) (см. [4]). Будем использовать эту процедуру при каждом вычислении субградиента в процессе минимизации и то разбиение множества вершин V на подмножества V_1 и V_2 , при котором реализуется максимум функционала (1), будем считать нижней оценкой для задачи MaxCut (φ_{lo}).

Для экспериментов использовались тестовые задачи для графа Петерсона (10 вершин и 15 ребер) с нумерацией вершин как в [4]. Значения весовых коэффициентов ребер для тестовых примеров P_1, \dots, P_9 приведены в табл. 1.

Результаты работы $r(\alpha)$ -алгоритма по итерациям ($iter$) и числу вычислений субградиента $\psi(u)$ (n_ψ) при нахождении оценок для MaxCut применительно к тестовым примерам P_1, \dots, P_9 приведены в табл. 2. Использовались следующие параметры $r(\alpha)$ -алгоритма: $\alpha = 2$, $h_0 = 1.0$, $q_1 = 0.9$, $n_h = 3$, $q_2 = 1.1$. В качестве начальной стартовой точки исполь-

зовалась $u_0 = \{\overbrace{1, \dots, 1}^{(\frac{n}{2})}, \overbrace{0, \dots, 0}^{(\frac{n}{2})}\}$. Останов $r(\alpha)$ -алгоритма производился при выполнении одного из условий: либо $\|u_k - u_{k-1}\| \leq \varepsilon_u = 10^{-3}$, либо $\varphi_{up} - \varphi_{lo} < 1$.

Из табл. 2 мы видим, что для семи случаев были получены точные решения для MaxCut, так как достигнутые верхние оценки отличаются

Таблица 1 Тестовые примеры для графа Петерсона

(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

Таблица 2 MaxCut. Штрафной множитель $t = n = 10$

$Test$	$iter(n_\psi)$	φ_{up}	φ_{lo}	V_1	V_2
P_1	4(6)	12.71280	12.0	1,2,4,6,7,8	3,5,9,10
P_2	3(9)	35.70171	35.0	2,4,5,8,9,10	1,3,6,7
P_3	5(13)	40.91025	40.0	2,5,7,8,10	1,3,4,6,9
P_4	2(9)	43.81359	43.0	1,3,6,7	2,4,5,8,9,10
P_5	49(70)	52.76558	51.0	1,2,4,6,10	3,5,7,8,9
P_6	3(11)	34.97500	34.0	2,3,5,6,7,10	1,4,8,9
P_7	49(74)	40.13301	39.0	2,3,5,6,7,10	1,4,8,9
P_8	3(12)	41.75555	41.0	1,3,5,7,8,9	2,4,6,10
P_9	17(31)	52.99566	52.0	1,2,4,6,7,8	3,5,9,10

от эвристически построенных допустимых решений не более, чем на единицу, при этом потребовалось от трех до семнадцати итераций $r(\alpha)$ -алгоритма. Для случаев P_5 и P_7 верхняя оценка получена с точностью 10^{-3} .

Близкие по качеству результаты получены для задачи BiSection.

Задача BiSection для графа $G(V, E)$ формулируется при четных $|V|$ и соответствует задаче MaxCut с условием, чтобы число вершин в подмножествах V_1 и V_2 было одинаково. Ей соответствует задача нелинейного программирования квадратичного типа (1)–(2), дополненная либо ограничением

$$\sum_{i=1}^n y_i = 0, \quad (6)$$

либо его квадратичным аналогом

$$\left(\sum_{i=1}^n y_i \right)^2 = 0. \quad (7)$$

Нахождение верхней Лагранжевой квадратичной оценки для BiSection можно проводить различными способами: а) выразив y_n через переменные y_1, \dots, y_{n-1} , используя (6); б) ставя в соответствие ограничению (7) свой множитель Лагранжа; в) используя для учета ограничения (7) штрафной множитель s , выбирая его порядка величины $\sum_{(i,j) \in E} |w_{ij}|$.

Для последнего из вариантов схема решения задачи BiSection аналогична получению решения для MaxCut, с тем отличием, что матрица $B_w(u)$ заменяется матрицей $\tilde{B}_w(u) = sI_n + B_w(u)$, где I_n – матрица размером $n \times n$, состоящая из всех единиц, а матрица $B_w(u)$ такая же, как и в (5).

Результаты экспериментов по нахождению оценок в задаче BiSection для тестовых примеров P_1, \dots, P_9 по варианту (в) приведены в табл. 3. Здесь начальная стартовая точка, параметры и критерии останова $r(\alpha)$ -алгоритма такие же, как и ранее.

Задача MinPartK для графа $G(V, E)$ состоит в нахождении такого разбиения вершин графа G на k частей с заданным числом вершин в каждой части (m_1, \dots, m_k) , чтобы сумма разрезов была минимальна.

Пусть $m_1 \geq \dots \geq m_k$ и $\sum_{i=1}^k m_i = n$. Тогда оценка снизу для MinPartK соответствует

$$\pi_k(G) \geq -\frac{1}{2}\psi^*, \quad (8)$$

Таблица 3 BiSection. $t = n = 10$, $s = 5n = 50$

<i>Test</i>	<i>iter</i> (n_ψ)	φ_{up}	φ_{lo}	V_1	V_2
P_1	41(46)	12.50045	11.0	1,2,5,7,8	3,4,6,9,10
P_2	7(16)	34.91024	34.0	2,3,6,7,10	1,4,5,8,9
P_3	2(8)	40.84215	40.0	2,5,7,8,10	1,3,4,6,9
P_4	44(68)	41.63351	40.0	1,3,6,7,10	2,4,5,8,9
P_5	35(52)	52.52589	51.0	1,2,4,6,7	3,5,8,9,10
P_6	32(43)	34.14082	33.0	1,4,5,8,9	2,3,6,7,10
P_7	17(32)	39.99753	39.0	1,2,4,8,9	3,5,6,7,10
P_8	11(21)	40.99564	40.0	2,4,6,8,10	1,3,5,7,9
P_9	7(21)	52.92904	52.0	1,2,4,6,7	3,5,8,9,10

где

$$\psi^* = \min_u \sum_{j=1}^k m_j \lambda_j(W + \text{diag}(u)) \quad (9)$$

при ограничении

$$\sum_{i=1}^n u_i = - \sum_{(i,j) \in E} w_{ij}. \quad (10)$$

Здесь $\lambda_j(W + \text{diag}(u))$ – j -е по порядку собственное число матрицы $(W + \text{diag}(u))$ размером $n \times n$, $\lambda_1 \geq \lambda_2 \geq \dots \lambda_k \geq \dots \lambda_n$. Матрица W формируется так же, как и для (5).

Для нахождения оценок в MinPartK можно применять r -алгоритм путем сведения задачи (9)–(10) к безусловной минимизации выпуклой негладкой функции, используя для учета ограничения (10) штрафной множитель s , который должен быть порядка $\max_{i,j} |w_{ij}|$.

В качестве тестовых экспериментов для задачи MinPartK рассматривались два способа разбиения вершин графа Петерсона: $k = 2$ ($m_1 = 5$, $m_2 = 5$) и $k = 3$ ($m_1 = 4$, $m_2 = 3$, $m_3 = 3$). Начальная стартовая точка и параметры $r(\alpha)$ -алгоритма такие же, как и в двух предыдущих расчетах. Для получения нижних оценок с точностью до 10^{-4} потребовалось во всех примерах от 52 до 67 итераций $r(\alpha)$ -алгоритма при первом способе и от 48 до 57 итераций – при втором. При $s=10$ для всех тестовых примеров ограничение (10) выполнялось с высокой точностью: $10^{-8} - 10^{-6}$.

Литература

1. ШОР Н.З. Методы минимизации недифференцируемых функций и их приложения. – Киев: Наукова думка, 1979. – 199с.
2. ШОР Н.З., СТЕЦЮК П.И. Использование модификации r -алгоритма для нахождения глобального минимума полиномиальных функций // Кибернетика и систем. анализ. – 1997. – № 4. – С. 28–49.
3. O. DAMBERG AND A. MIGDALAS. A data parallel space dilation algorithm for the concentrator location problem // DIMACS Series in Discr. Math., 22:57-80, DMS 1995.
4. ШОР Н.З., БЕРЕЗОВСКИЙ О.А. Новые алгоритмы решения задачи о максимальном разрезе графа // Кибернетика и систем. анализ. – 1995. – № 2. – С. 100-106.

Нахождение глобальных минимумов полиномиальных функций с использованием двойственных квадратичных оценок

Н. З. Шор, П. И. Стецюк, С. В. Крылов

Вестник Международного Соломонова Университета

„Математические методы в кибернетике“, – 2000. – № 4. –

С. 217–233.

В статье [1] рассматривались задачи нахождения глобального минимума полиномиальных функций от одной или нескольких переменных путем сведения их к задачам квадратичного типа, построения функций Лагранжа и вычисления двойственных оценок, которые являются оценками снизу для оптимального значения целевой функции первоначальной задачи. Для ряда классов полиномиальных задач генерировались функционально избыточные ограничения в форме квадратичных равенств и неравенств, которые не изменяя области допустимых решений первоначальной задачи, давали возможность модифицировать функцию Лагранжа (при этом увеличивалось число множителей Лагранжа) и получать более точные двойственные оценки, а в некоторых случаях, оценки, совпадающие с точным значением глобального минимума полинома.

Получение двойственных оценок ψ^* связано с нахождением максимума вогнутых функций $\psi(u)$ от вектора множителей Лагранжа $u = \{u_1, \dots, u_m\}$, т.е.

$$\psi^* = \sup_{u \in \text{dom}\psi} \psi(u),$$

где $\psi(u) = \inf_x L(x, u)$, L – функция Лагранжа, $\text{dom}\psi$ – область определения вогнутой функции $\psi(u)$. Для максимизации $\psi(u)$ на $\text{dom}\psi$ предложена некоторая модификация r -алгоритма, которая описана ниже в разделе 1. В случае, когда $\psi^* = \psi(u^*)$ достигается во внутренней точке u^* области $\text{dom}\psi$ ($u^* \in \text{intdom}\psi$), то ψ^* – совпадает с оптимальным значением целевой функции первоначальной задачи. При этом r -алгоритм работает достаточно эффективно. Если же ψ^* достигается на границе области $\text{dom}\psi$, то это может быть по разным причинам: а) ψ^* строго меньше оптимального значения первоначальной задачи; б) u^* – неоднозначно. В этих случаях нужно контролировать максимизирующую

последовательность $\{u^{(k)}\}_{k=0}^{\infty}$ и в случае выхода $u^{(k)}$ за пределы $\text{dom}\psi$ переключать r -алгоритм.

Содержание статьи следующее. В разделе 1 описан алгоритм для нахождения двойственных оценок в задаче квадратичного типа, основанный на модификации r -алгоритма. В разделе 2 описан подход к минимизации полиномов от одной и нескольких переменных, использующий сведение к задачам квадратичного типа с применением функционально избыточных ограничений. В разделе 3 описаны вычислительные эксперименты по нахождению двойственных квадратичных оценок для задачи нахождения глобального минимума полиномов 6-й и 4-й степени от одной и четырех переменных соответственно.

1. Алгоритм нахождения двойственных оценок в задаче квадратичного типа

Рассмотрим задачу минимизации квадратичного типа:

найти

$$q^* = \inf_{x \in E^n} Q_0(x) \quad (1)$$

при ограничениях

$$Q_i(x) = 0, \quad i = 1, \dots, m. \quad (2)$$

Здесь $Q_\nu(x)$ – 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, $\nu = 0, 1, \dots, m$.

Оценки снизу для q^* могут быть получены путем лагранжевой релаксации. Пусть $u = (u_1, \dots, u_m) \in R^m$ – вектор множителей Лагранжа задачи (1),(2). Рассмотрим функцию Лагранжа

$$L(x, u) = Q_0(x) + \sum_{i=1}^m u_i Q_i(x)$$

и функцию

$$\psi(u) = \inf_x L(x, u) = \inf_x [(K(u)x, x) + (b(u), x) + c(u)],$$

где

$$K(u) = K_0(x) + \sum_{i=1}^m u_i K_i(x),$$

$$b(u) = b_0 + \sum_{i=1}^m u_i b_i,$$

$$c(u) = c_0 + \sum_{i=1}^m u_i c_i.$$

Пусть $\lambda_1(K) \geq \lambda_2(K) \geq \dots \geq \lambda_n(K)$ – собственные числа симметричной матрицы K размера $n \times n$, упорядоченные в порядке убывания. Обозначим: Ω^+ – множество u , у которых $\lambda_n[K(u)] > 0$; Ω^- – множество u , для которых $K(u)$ имеет отрицательные собственные числа; Ω^0 – множество u , у которых $\lambda_n[K(u)] = 0$.

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^*

$$\psi^* = \sup_{u \in \text{dom}\psi} \psi(u),$$

которая называется двойственной квадратичной оценкой задачи (1), (2). Если $\psi^* = +\infty$, то это означает, что система ограничений (2) несовместна.

Если ψ^* достигается на $u^* \in \Omega^+$, то

$$\psi^* = \psi(u^*) = q^* = Q_0(x(u^*)),$$

где $x(u^*)$ – решение системы (3) при $u = u^*$. В противном случае ψ^* достигается на границе области Ω^+ , при этом может существовать положительный "разрыв" ("gap")

$$\Delta^* = q^* - \psi^* > 0.$$

Один из способов уменьшения этого разрыва связан с введением функционально избыточных ограничений, добавление которых оставляет множество оптимальных решений первоначальной задачи неизменным (при этом может увеличиться и множество переменных). Однако их введение меняет функцию Лагранжа, что может в некоторых случаях уменьшить разрыв между оптимальным значением q^* целевой функции и лагранжевой (двойственной) оценкой ψ^* (см.[2, 1]).

Для нахождения ψ^* нужно сконструировать алгоритмы, дающие возможность решать следующие подзадачи:

- (i) проверять, является ли область $\text{dom}\psi$ не пустой, и если $\text{dom}\psi$ не пуста, находить точку $\bar{u} \in \text{dom}\psi$;
- (ii) достаточно эффективно максимизировать $\psi(u)$ на $\text{dom}\psi$;
- (iii) после каждой итерации алгоритма отслеживать, не вышли ли мы за пределы $\text{dom}\psi$, и если это произошло, то алгоритм должен снова найти допустимую точку в $\text{dom}\psi$.

Подзадача (i) сводится к нахождению $\lambda^* = \max_u \lambda_n[K(u)]$, где $\lambda_n(A)$ означает минимальное собственное число симметричной матрицы A размерности $n \times n$. Если $\lambda^* \leq 0$, то $\text{dom}\psi$ пусто. Если в процессе максимизации $\lambda_n[K(u)]$ получено \bar{u} такое, что $\lambda_n[K(\bar{u})] > 0$, то точка $\bar{u} \in \text{dom}\psi$. Так как $\lambda_n[K(u)]$ 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 [3, 4]), который практически является одним из наиболее эффективных методов негладкой оптимизации.

Этот же метод можно применять для решения подзадач (ii) и (iii), учитывая что $\psi(u)$ – вогнутая функция от u . При их решении проверка выхода текущей точки \bar{u} за пределы $\text{dom}\psi$ осуществляется путем разложения матрицы $K(\bar{u})$ по Холецкому. Если $\bar{u} \notin \text{dom}\psi$, то включается ветвь алгоритма, решающая первую подзадачу. В случае, если $\psi(u)$ имеет единственную точку максимума $u^* \in \text{dom}\psi$, то после определенного числа итераций r -алгоритма, когда длины шагов становятся достаточно малыми, текущие точки максимизирующей последовательности находятся внутри области $\text{dom}\psi$, и процесс максимизации $\psi(u)$ ведет себя так, как-будто $\psi(u)$ определена на всем R^m .

Осложнения возникают тогда, когда точка максимума функции $\psi(u)$ либо определяется неоднозначно, либо находится на границе области $\text{dom}\psi$. Первый случай типичен для квадратичных задач, когда мы вводим избыточные ограничения, чтобы улучшить двойственные оценки. Второй случай встречается довольно часто, например, в задачах, где двойственная оценка не совпадает с оптимальным значением задачи (1),(2), а также, когда оптимальная точка задачи (1),(2) – не единственная.

Пусть оптимальное значение вектора двойственных переменных u^* находится на границе области $\text{dom}\psi$. Тогда при приближении \bar{u} к u^* , $\bar{u} \in \text{dom}\psi$, решение системы (3) становится все менее устойчивым. Для регуляризации решения можно использовать ε -возмущения диагональ-

ных элементов матриц $K(u)$, рассматривая $\tilde{K}_\varepsilon(u) = K(u) + \varepsilon I$, где $\varepsilon > 0$, I - единичная матрица. Такой способ приводит к монотонным вариантам r -алгоритмов, для которых в ряде случаев проблему представляет сходимость к экстремальной точке из точек, которые близки к границе области $\text{dom}\psi$.

Другой способ преодоления этой неустойчивости состоит в следующем. Рассмотрим функцию:

$$\tilde{\psi}(u) = \begin{cases} \psi(u), & \text{если } u \in \Omega^+ (\text{т.е. } \lambda_n[K(u)] > 0) \\ \lambda_n[K(u)], & \text{если } u \notin \Omega^+ (\text{т.е. } \lambda_n[K(u)] \leq 0). \end{cases} \quad (4)$$

и будем использовать хорошо зарекомендовавший себя на практике вариант r -алгоритма с адаптивной регулировкой шагового множителя, взяв в качестве суперградиента вогнутой функции на каждом шаге суперградиент от функции $\tilde{\psi}(u)$ в соответствующей точке, т.е. суперградиент от функции $\psi(u)$ если $u \in \Omega^+$ и суперградиент от функции $\lambda_n[K(u)]$ если $u \notin \Omega^+$. Такой способ можно рассматривать как модификацию $r(\alpha)$ -алгоритма. Он дает возможность решать все три указанные выше подзадачи (i), (ii) и (iii) и, кроме того, позволяет избежать осложнений, связанных как с неоднозначностью u^* так и с принадлежностью u^* границе $\text{dom}\psi$.

Этот способ оказался достаточно эффективным (см. [5, 2]). Применительно к нахождению двойственных оценок в задачах квадратичного типа (1),(2) суперградиент функции $\tilde{\psi}(u)$, определенной согласно (4), вычисляется по формуле:

$$\text{supergrad}_u \tilde{\psi}(u) = \begin{cases} \text{supergrad}_u \psi(u), & \text{если } \lambda_n[K(u)] > 0, \\ \text{supergrad}_u \lambda_n[K(u)], & \text{если } \lambda_n[K(u)] \leq 0, \end{cases} \quad (5)$$

Здесь

$$\text{supergrad}_u \psi(u) = \{Q_1(x(\bar{u}), \dots, Q_m(x(\bar{u}))\}^T, \quad (6)$$

где $x(\bar{u})$ - решение системы уравнений (3), и

$$\text{supergrad}_u \lambda_n[K(u)] = \{(\xi, K_1 \xi), \dots, (\xi, K_m \xi)\}^T, \quad (7)$$

где ξ - собственный вектор, относящийся к минимальному собственному числу $\lambda_n[K(u)]$.

Для нахождения двойственных оценок в задаче квадратичного типа (1),(2) по указанной выше схеме реализована исследовательская программа DQuadE. При вычислении суперградиента $\tilde{\psi}(u)$ по формулам (5)-(7) были использованы следующие фортрановские подпрограммы:

для решения системы уравнений (3) с симметричной положительно определенной матрицей $K(u)$ – подпрограммы *DPPFA* (реализует факторизацию матрицы $K(u)$, если $K(u)$ – положительно определена, и в случае, если $K(u)$ не является положительно определенной, сигнализирует об этом) и *DPPSL* (находит решение системы, используя факторизованную программой *DPPFA* матрицу $K(u)$) из пакета LINPACK; для нахождения минимального собственного числа $\lambda_n[K(u)]$ матрицы $K(u)$ и соответствующего этому числу собственного вектора ξ – подпрограмма *EIGEN* из [6].

С помощью программы DQuadE проведен ряд тестовых экспериментов по нахождению двойственных оценок в квадратичных задачах, к которым сводятся некоторые классы полиномиальных задач.

2. Двойственный подход к минимизации полиномов от нескольких переменных

Рассмотрим класс полиномов $P_n^{(2l)}$ тотальной степени $2l$ от n переменных $x = \{x_1, \dots, x_n\}$ и систему одночленов $Q_n^{(l)}$ с коэффициентом 1 тотальной степени, не превышающей l , от n переменных. Каждый одночлен с коэффициентом 1 тотальной степени, не превышающей $2l$, может быть представлен (не обязательно однозначно) как произведение двух одночленов из системы $Q_n^{(l)}$. Каждому одночлену из этого семейства сопоставим переменную $z[\alpha_1, \alpha_2, \dots, \alpha_n]$, где $\alpha_1, \alpha_2, \dots, \alpha_n$ – степени, с которыми переменная x_k входит в соответствующий одночлен ($\alpha_1, \alpha_2, \dots, \alpha_n \geq 0$; $\sum_{1 \leq k \leq n} \alpha_k \leq l$).

Рассмотрим четыре переменных $z_1(\alpha_1^{(1)}, \dots, \alpha_n^{(1)})$, $z_2(\alpha_1^{(2)}, \dots, \alpha_n^{(2)})$, $z_3(\alpha_1^{(3)}, \dots, \alpha_n^{(3)})$ и $z_4(\alpha_1^{(4)}, \dots, \alpha_n^{(4)})$. Если $\alpha_r^{(1)} + \alpha_r^{(2)} = \alpha_r^{(3)} + \alpha_r^{(4)}$ для всех r : $1 \leq r \leq n$, то мы можем записать квадратичное равенство:

$$z_1 z_2 - z_3 z_4 = 0 \quad (8)$$

Каждый полином тотальной степени $2l$ может быть представлен в виде

$$P_n^{(2l)}(x) = Q_0(z) = \sum_{\beta \in R} c_\beta z_\beta^{(1)} z_\beta^{(2)}, \quad (9)$$

где R – множество всех одночленов полинома $P_n^{(2l)}(x)$, $z_\beta^{(1)} z_\beta^{(2)}$ – некоторое представление одночлена β , c_β – соответствующий коэффициент.

В новых переменных z мы получаем задачу с квадратичной целевой функцией вида (9) и квадратичными ограничениями вида (8).

В [1] доказана следующая теорема.

Теорема. *Если задача состоит в нахождении глобального минимума полинома $P_n^{(2l)}(x)$ и с использованием z -переменных она свелась к задаче минимизации квадратичной функции (9) при всевозможных ограничениях вида (8), то двойственная квадратичная оценка ψ^* совпадает с P^* – глобальным минимумом $P_n^{(2l)}(x)$ тогда и только тогда, когда неотрицательный полином $\overline{P(x)} = P_n^{(2l)}(x) - P^*$ разлагается в сумму квадратов полиномиальных функций.*

Рассмотрим несколько классов задач нахождения глобального минимума полиномов.

а) Полиномы четной степени $2l$ от одной переменной с коэффициентом 1 при старшем члене:

$$P^{(2l)}(x_1) = x_1^{2l} + \sum_{i=1}^{2l-1} c_{2l-i} x_1^{2l-i}; \quad l \geq 1. \quad (10)$$

Для того, чтобы полином (10) свести к квадратичному виду, определим переменные z_1, \dots, z_l следующим образом

$$z_k = x_1^k, \quad k = 1, \dots, l,$$

что дает следующие квадратичные ограничения-равенства:

$$z_k - z_{k-1}z_1 = 0, \quad k = 2, \dots, l.$$

Запишем также систему функционально избыточных ограничений:

$$z_1 z_{m-1} - z_k z_{m-k} = 0$$

для $2 \leq k \leq \frac{m}{2}$; $4 \leq m \leq l$; и

$$z_l z_m - z_{l-k} z_{m+k} = 0$$

для $1 \leq m \leq l-2$; $l-k \geq m+k$; $k \geq 1$.

Например, при $2l = 6$ имеем ограничения

$$z_2 - z_1^2 = 0; \quad z_3 - z_1 z_2 = 0; \quad z_3 z_1 - z_2^2 = 0$$

при $2l = 8$ имеем ограничения

$$\begin{aligned} z_2 - z_1^2 &= 0; & z_3 - z_1 z_2 &= 0; & z_4 - z_1 z_3 &= 0 \\ z_1 z_3 - z_2^2 &= 0; & z_1 z_4 - z_2 z_3 &= 0; & z_2 z_4 - z_3^2 &= 0. \end{aligned}$$

Так как любой неотрицательный полином четной степени одной переменной разлагается в сумму квадратов полиномов, то оценка

$$\psi^* = \min_{x_1} P^{(2l)}(x_1) = P^{(2l)}(x_1^*) = P^*.$$

Маргинальная функция $\psi(u)$ функции Лагранжа $L(z, u)$ при полной системе квадратичных ограничений и единственном глобальном минимуме достигает своего максимума внутри области $\text{dom}\psi$. Если имеется несколько глобальных минимумов, то оптимальное значение ψ^* функции $\psi(u)$ достигается на границе $\text{dom}\psi$.

б) Рассмотрим класс $P_n^{(4)}$ полиномов тотальной 4-й степени от n переменных, $n \geq 2$. Запишем множество всевозможных квадратичных равенств вида (8), включающий полный набор функционально избыточных ограничений. Пусть z_1, \dots, z_n – набор первоначальных переменных, введем дополнительные переменные: $z_{ii} = z_i^2$, $i = 1, \dots, n$, $z_{ij} = z_i z_j$, $\forall (i, j)$, $i < j$. Рассматривая произвольный одночлен с коэффициентом 1, тотальной степени меньше или равной 4, получим следующие функционально избыточные квадратичные равенства для различных i_1, i_2, i_3, i_4 ; $i_1 < i_2 < i_3 < i_4$:

$$z_{i_1 i_2} z_{i_3 i_4} - z_{i_1 i_3} z_{i_2 i_4} = z_{i_1 i_2} z_{i_3 i_4} - z_{i_1 i_4} z_{i_2 i_3} = 0. \quad (11a)$$

Ограничения типа (11a) можно написать для любого сочетания четырех индексов из n , и это дает нам $2C_n^4$ дополнительных ограничений. Если ровно два индекса из 4-х совпадают, например $i_1 = i_2$, то получим следующее равенство:

$$z_{i_1 i_1} z_{i_3 i_4} - z_{i_1 i_3} z_{i_1 i_4} = 0. \quad (11b)$$

Для одночленов вида $x_i^2 x_j^2$, $i_1 < i_2$ имеем избыточное равенство:

$$z_{i_1 i_1} z_{i_2 i_2} - z_{i_1 i_2}^2 = 0. \quad (11c)$$

Для кубических одночленов вида $x_{i_1} x_{i_2} x_{i_3}$, $i_1 < i_2 < i_3$, имеем соотношения:

$$z_{i_1} z_{i_2 i_3} - z_{i_2} z_{i_1 i_3} = 0; \quad z_{i_1} z_{i_2 i_3} - z_{i_3} z_{i_1 i_2} = 0. \quad (11d)$$

Если $i_1 = i_2$, то

$$z_{i_1} z_{i_1 i_3} - z_{i_3} z_{i_1 i_1} = 0. \quad (11e)$$

Таким образом полное число дополнительных ограничений, полученных из равенств (11a)–(11e) равно

$$M(n) = 2C_n^4 + 3C_n^3 + C_n^2 + 2C_n^3 + 2C_n^2,$$

а число дополнительных переменных равно $\frac{n(n+1)}{2}$.

В [7] Д. Гильберт доказал теорему о том, что каждый неотрицательный полином 4-й степени от 2-х переменных разлагается в сумму квадратов других полиномов. В то же время существуют неотрицательные полиномы 4-й степени от трех и более переменных, которые не представимы в виде суммы квадратов других полиномов. Это означает, что для $n \geq 3$ даже использование полного набора функционально избыточных ограничений не гарантирует того, что соответствующая квадратичная оценка совпадает с глобальным минимумом полинома 4-й степени P^* , даже если $P^* > -\infty$. С другой стороны мы можем предложить алгоритм последовательного улучшения оценок, выбрав некоторое начальное приближение для $\{z_i^{(0)}\}_{i=1}^n$, а затем улучшить значения z_{i_1} и z_{i_2} , $i_1 < i_2$, получив глобальный минимум по этим переменным при фиксации остальных переменных. Этот процесс можно повторять многократно с различными парами индексов, получая новые вектора $\{z_i^{(k)}\}_{i=1}^n$.

Вычислительные эксперименты, которые мы проводили, были связаны с нахождением двойственных оценок в задачах квадратичного типа, соответствующих нахождению глобального минимума как полинома четной степени от одной переменной, так и полинома 4-й степени от четырех переменных (на примере так называемой задачи невыполнимости). Подробное описание тестовых экспериментов дано ниже.

3. Численные эксперименты по нахождению глобальных минимумов полиномов

Первая серия экспериментов по нахождению глобального минимума полиномов связана с полиномом шестой степени от одной переменной

$$P_6(x_1) = x_1^6 + a_5x_1^5 + a_4x_1^4 + a_3x_1^3 + a_2x_1^2 + a_1x_1, \quad (12)$$

глобальный минимум которого обозначим $P_6^*(x_1)$. Введем переменные $x_2 = x_1^2$ и $x_3 = x_1^3 = x_1x_2$. Тогда задача нахождения минимума полинома $P_6(x_1)$ сводится к эквивалентной задаче квадратичного программирования относительно переменных x_1, x_2, x_3 :

найти

$$q^* = \min_x \{Q_0(x) = a_4 x_2^2 + a_5 x_2 x_3 + x_3^2 + a_1 x_1 + a_2 x_2 + a_3 x_3\} \quad (13)$$

при ограничениях:

$$Q_1(x) = x_1^2 - x_2 = 0, \quad (14)$$

$$Q_2(x) = x_1 x_2 - x_3 = 0, \quad (15)$$

где $x = (x_1, x_2, x_3)$.

Оптимальную двойственную лагранжеву оценку для задачи (13)–(15) обозначим ψ_1^* . В ряде случаев (например, когда $a_4 < 0$) ψ_1^* будет тривиальной. Для того чтобы получить нетривиальную двойственную квадратичную оценку введем дополнительное функционально избыточное ограничение:

$$Q_3(x) = x_2^2 - x_1 x_3 = 0. \quad (16)$$

Ограничение (16) не меняет смысла первоначальной задачи нахождения глобального минимума полинома $P_6(x_1)$, так как x_2^2 и $x_1 x_3$ равны x_1^4 . Ограничение $Q_3(x)$ совместно с ограничениями $Q_1(x)$ и $Q_2(x)$ задает полную систему функционально избыточных квадратичных ограничений для полинома 6-й степени от одной переменной (см. раздел 2). Согласно теореме из раздела 2 оптимальная двойственная лагранжева оценка ψ_2^* для задачи квадратичного типа (13)–(16) дает точную оценку снизу для глобального минимума полинома $P_6(x_1)$, т.е. всегда справедливо равенство: $P_6^*(x_1) = \psi_2^*$.

Численные эксперименты связаны с нахождением двойственных квадратичных оценок для задач (13)–(15) и (13)–(16) при различных значениях коэффициентов a_1, a_2, a_3, a_4, a_5 для полинома $P_6(x_1)$. Для тестовых расчетов использовалась программа DQuadE с выбором следующих значений параметров для $r(\alpha)$ -алгоритма: коэффициент растяжения пространства $\alpha = 2$ и параметры адаптивной регулировки шагового множителя $h_0 = 1.0, q_1 = 1.0, n_h = 3, q_2 = 1.1$. В качестве начальной стартовой точки выбиралась $u_0 = (1, \dots, 1)$. Работа алгоритма прекращалась на k -ом шаге, если выполнялось условие останова: $\|u_k - u_{k-1}\| \leq 10^{-6}$.

Первая группа тестовых экспериментов для полинома $P_6(x_1)$ состояла в нахождении оценок его минимума для ряда значений коэффициента a_4 и для произвольных значений коэффициентов $a_1, a_2, a_3, a_5 \in 0 \vee 1$. Результаты численных экспериментов приведены в табл. 1 и 2. Здесь ψ_1^*, ψ_2^* – найденные с помощью программы DQuadE приближения оптимальных двойственных оценок для задач квадратичного типа (13)–(15) и (13)–(16) (соответственно).

Таблица 1.

a_1, a_2, a_3, a_5	$a_4 = 0.01$		$a_4 = 1$	
	ψ_2^*	ψ_1^*	ψ_2^*	ψ_1^*
0,0,0,0	.000000	.000000	.000000	.000000
0,0,0,1	-.062249	$-\infty$.000000	.000000
0,0,1,0	-.246059	-.246059	-.047811	-.047811
0,0,1,1	-1.212015	$-\infty$	-.168157	-.168157
0,1,0,0	.000000	.000000	.000000	.000000
0,1,0,1	.000000	$-\infty$.000000	.000000
0,1,1,0	.000000	-.160000	.000000	.000000
0,1,1,1	.000000	$-\infty$.000000	.000000
1,0,0,0	-.579984	-1.162552	-.430895	-.430895
1,0,0,1	-.990080	$-\infty$	-.513046	-.513046
1,0,1,0	-1.091714	-1.458970	-.685694	-.685694
1,0,1,1	-2.420464	$-\infty$	-1.035148	-1.035148
1,1,0,0	-.238818	-.238818	-.211822	-.211822
1,1,0,1	-.270041	$-\infty$	-.220560	-.220560
1,1,1,0	-.433171	-.443941	-.297365	-.297365
1,1,1,1	-1.045353	$-\infty$	-.364906	-.364906

Из табл. 1 видим, что при положительном значении $a_4 = 0.01$ оптимальная двойственная оценка для задачи (13)–(15) в половине случаев – тривиальной и только в одном случае является точной оценкой глобального минимума полинома $P_6(x_1)$. При отрицательных значениях коэффициента a_4 оптимальная оценка ψ_1^* всегда будет тривиальной, так как для задачи (13)–(15) в этом случае невозможно обеспечить неотрицательную определенность матрицы $K(u_1, u_2)$ (см.[1]). В то же время при коэффициенте $a_4 = 1$ квадратичные оценки ψ_1^* и ψ_2^* совпадают для всех тестовых примеров, и, следовательно, оптимальная лагранжева оценка для задачи квадратичного типа (13)–(15) дает точные значения глобального минимума полинома $P_6(x_1)$ при $a_4 = 1$ и указанных в табл. 1 значениях коэффициентов a_1, a_2, a_3, a_5 .

В табл. 2 приведены результаты нахождения оптимальной оценки ψ_2^* при четырех значениях (в том числе и отрицательных) коэффициента a_4 . Приведенные в табл. 2 значения ψ_2^* являются глобальным минимумом полинома $P_6(x_1)$ для заданных значений коэффициентов a_1, \dots, a_5 . Точность нахождения глобального минимума можно оценить по тем

Таблица 2.

a_1, a_2, a_3, a_5	$\psi_2^*(a_4 = -2)$	$\psi_2^*(a_4 = -1)$	$\psi_2^*(a_4 = 0)$	$\psi_2^*(a_4 = 2)$
0,0,0,0	-1.185185	-1.148148	.000000	.000000
0,0,0,1	-6.875730	-1.755830	-.066980	.000000
0,0,1,0	-3.084179	-1.035279	-.250000	-.011044
0,0,1,1	-11.808634	-4.640625	-1.231800	-.019511
0,1,0,0	.000000	.000000	.000000	.000000
0,1,0,1	-4.258825	-.136038	.000000	.000000
0,1,1,0	-1.462974	-.043660	.000000	.000000
0,1,1,1	-8.796754	-2.489647	.000000	.000000
1,0,0,0	-2.361420	-1.031139	-.582356	-.361701
1,0,0,1	-8.527442	-3.107123	-1.000000	-.392031
1,0,1,0	-4.406037	-2.125353	-1.098012	-.510897
1,0,1,1	-13.569233	-6.150644	-2.443518	-.613653
1,1,0,0	-1.048048	-.332656	-.239214	-.196174
1,1,0,1	-5.857333	-1.367189	-.271215	-.200901
1,1,1,0	-2.712784	-1.000000	-.435720	-.250000
1,1,1,1	-10.517725	-3.933006	-1.060246	-.267485

наборам коэффициентов, для которых аналитическое нахождение минимума легко сделать. Так, для набора $a_1 = a_2 = a_3 = a_5 = 0$ и $a_4 < 0$ значение глобального минимума равно $\frac{4}{27}(a_4)^3$, что с точностью до всех значащих цифр совпадает с приведенным в таблице. Это же касается некоторых других наборов коэффициентов (например $a_1 = a_2 = a_4 = a_5 = 0$, что соответствует полиному $x_1^6 + a_3 x_1^3$). Поэтому результаты табл. 2 и те из результатов табл. 1, которые связаны с ψ_2^* , можно считать точными (с точностью до всех указанных значащих цифр) значениями глобального минимума полинома $P_6(x_1)$. Их можно использовать как ориентир при проверке и тестировании методов нахождения глобального минимума. Вторая группа экспериментов для полинома $P_6(x_1)$ связана с проверкой численной устойчивости программы DQuadE, когда оптимальное решение двойственной задачи к задаче (13)–(16) находится на границе области положительной определенности матрицы $K(u_1, u_2, u_3)$. В качестве тестового выбран следующий полином:

$$P_6(x) = x^6 - 2tx^4 + (t^2 - \varepsilon)x^2 = x^2((x^2 - t)^2 - \varepsilon), \quad (17)$$

т.е. полином типа (12) с коэффициентами $a_1 = a_3 = a_5 = 0$, $a_4 = -2t$,

Таблица 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

$a_2 = t^2 - \varepsilon$. При $t > 0$, $\varepsilon > 0$ и $t^2 - \varepsilon > 0$ для полинома (17) существует два глобальных минимума:

$$x_1^* = \sqrt{\frac{2t}{3} + \frac{\sqrt{t^2 + 3\varepsilon}}{3}}; \quad x_2^* = -\sqrt{\frac{2t}{3} + \frac{\sqrt{t^2 + 3\varepsilon}}{3}},$$

где полином принимает отрицательное значение, а также точка локального минимума $x^* = 0$, в которой полином $P_6(x)$ принимает значение, равное нулю. Более того, когда значение ε стремится к нулю, то и значения полинома в точках x_1^* и x_2^* стремятся к нулю и в пределе при $\varepsilon = 0$ полином $P_6(x)$ имеет три локальных минимума, каждый из которых является его глобальным минимумом.

Результаты по нахождению точной двойственной оценки для глобального минимума полинома $P_6(x)$ при различных значениях параметра t и ε представлены в табл. 3.

Здесь: ψ_2^* – найденные приближения оптимальных двойственных оценок для задачи квадратичного типа (13)–(16), которая соответствует

задаче нахождения глобального минимума полинома $P_6(x)$; P^* – вычисленное аналитически точное значение глобального минимума полинома $P_6(x)$; u_1^* , u_2^* , u_3^* – найденные приближения оптимальных двойственных переменных (множителей Лагранжа) для ограничений (14), (15) и (16) (соответственно).

Из табл. 3 видно, что для всех тестовых примеров ψ_2^* в точности соответствует глобальному минимуму полинома $P_6(x)$. Для всех примеров оптимальные значения двойственных переменных u_1^* , u_2^* , u_3^* находятся на границе области положительной определенности матрицы $K(u)$. На самом деле, матрица $K(u)$ для задачи (13)–(16), соответствующей нахождению минимума полинома $P_6(x)$, имеет следующий вид:

$$K(u) = K(u_1, u_2, u_3) = \begin{pmatrix} u_1 & u_2/2 & -u_3/2 \\ u_2/2 & u_3 - 2t & 0 \\ -u_3/2 & 0 & 1 \end{pmatrix}.$$

Учитывая, что $u_2^* = 0$, имеем

$$K(u_1^*, 0, u_3^*) = \begin{pmatrix} u_1^* & 0 & -u_3^*/2 \\ 0 & u_3^* - 2t & 0 \\ -u_3^*/2 & 0 & 1 \end{pmatrix}$$

и собственные числа матрицы $K(u_1^*, 0, u_3^*)$ будут следующими

$$\lambda_1 = u_3^* - 2t; \quad \lambda_{2,3} = \frac{(u_1^* + 1) \pm \sqrt{(u_1^* - 1)^2 + (u_3^*)^2}}{2}.$$

Для $\varepsilon > 0$, что соответствует двум глобальным минимумам полинома $P_6(x)$, собственные числа матрицы $K(u_1^*, 0, u_3^*)$ удовлетворяют $\lambda_1 > 0$, $\lambda_2 > 0$ и $\lambda_3 = 0$, а для $\varepsilon = 0$, что соответствует трем глобальным минимумам, удовлетворяют $\lambda_2 > 0$ и $\lambda_1 = \lambda_3 = 0$. Точность нахождения оптимальных двойственных переменных легко оценить по тому, насколько точно выполнено соотношение $u_3^* = 2\sqrt{u_1^*}$, которое должно выполняться для $\lambda_3 = 0$.

Из приведенных результатов можно сделать вывод, что программа DQuadE является устойчивой по нахождению оптимальных двойственных оценок в случае, когда решение находится на границе области положительной определенности матрицы $K(u)$. Более того, анализ самого решения u^* позволяет указать и общее число глобальных минимумов полинома в том случае, когда ψ^* совпадает со значением глобального минимума полинома.

Вторая серия экспериментов связана с нахождением оптимальных двойственных квадратичных оценок для задачи с полиномом 4-й степени от четырех переменных, которая возникает при решении так называемых взвешенных задач выполнимости типа MAX-W-4-SAT, т.е. задач максимизации суммы весов выполненных дизъюнктивных скобок, в каждой из которых не более четырех литералов, полученных из $n = 4$ пропозициональных ("булевых") переменных (см. [8]). В качестве тестового примера рассматривалась задача:

$$f^* = \min_x \{ w_5 x_1 x_2 x_3 x_4 - \sum_{i=1}^4 w_i x_i - w_5 \} \quad (18)$$

при ограничениях:

$$x_i^2 - x_i = 0, \quad i = 1, \dots, 4. \quad (19)$$

Задаче (18)–(19) соответствует следующая целевая функция для задачи MAX-W-4-SAT:

$$F = w_1 x_1 + w_2 x_2 + w_3 x_3 + w_4 x_4 + w_5 (\overline{x_1} \vee \overline{x_2} \vee \overline{x_3} \vee \overline{x_4}).$$

Задачу (18)–(19) можно свести к квадратичной, вводя разное количество дополнительных переменных. Для того, чтобы одночлен $x_1 x_2 x_3 x_4$ преобразовать в квадратичный, достаточно введения 6-ти переменных: $z_1 = x_1$, $z_2 = x_2$, $z_3 = x_3$, $z_4 = x_4$, $z_5 = x_1 x_2$, $z_6 = x_3 x_4$. При этом целевая функция будет следующей:

$$Q_0(z) = w_5 z_5 z_6 - w_1 z_1 - w_2 z_2 - w_3 z_3 - w_4 z_4 - w_5.$$

Сведение задачи (18)–(19) к задаче квадратичного типа можно получить, вводя 10 переменных, т.е. к уже имеющимся 6-ти переменным добавить переменные $z_7 = x_1 x_3$, $z_8 = x_1 x_4$, $z_9 = x_2 x_3$, $z_{10} = x_2 x_4$ или даже весь полный набор из 14 переменных (т.е. все возможные одночлены степени не выше 2), добавив еще и переменные $z_{11} = x_1^2$, $z_{12} = x_2^2$, $z_{13} = x_3^2$, $z_{14} = x_4^2$.

Способ сведения задачи (18)–(19) к задаче квадратичного типа будет определять свою оптимальную двойственную квадратичную оценку, точность которой зависит от количества построенных избыточных ограничений в задаче квадратичного типа. В табл. 4 для некоторых наборов коэффициентов w_1 , w_2 , w_3 , w_4 , w_5 приведены оптимальные двойственные квадратичные оценки в зависимости от числа переменных n и числа ограничений m в задаче квадратичного типа, соответствующей задаче

Таблица 4.

n	m	w_1, w_2, w_3, w_4, w_5			
		$(3,3,3,3,2)$ $f^* = -12$	$(1,3,3,3,2)$ $f^* = -11$	$(1,1,3,3,2)$ $f^* = -9$	$(1,1,1,3,2)$ $f^* = -7$
6	6	$-\infty$	$-\infty$	$-\infty$	$-\infty$
6	8	-12.20089	-11.0	-9.01945	-7.14869
10	10	$-\infty$	$-\infty$	$-\infty$	$-\infty$
10	16	-12.20089	-11.0	-9.01945	-7.14869
10	18	-12.03458	-11.0	-9.0	-7.00719
10	24	-12.20089	-11.0	-9.01945	-7.14869
10	26	-12.03458	-11.0	-9.0	-7.00719
14	14	$-\infty$	$-\infty$	$-\infty$	$-\infty$
14	24	-12.20089	-11.0	-9.01945	-7.14869
14	44	-12.07233	-11.0	-9.0	-7.04666
14	44	-12.00043	-11.0	-9.0	-7.0
14	64	-12.00043	-11.0	-9.0	-7.0

(18)–(19). Здесь f^* – точные значения минимума для соответствующих наборов коэффициентов w_1, w_2, w_3, w_4, w_5 .

Для случая шести переменных ($n = 6$) рассматривались следующие ограничения – 2 ограничения из определений: $z_5 - z_1 z_2 = 0$, $z_6 - z_3 z_4 = 0$; 6 "булевых" ограничений: $z_i^2 - z_i = 0$ для $i = 1, \dots, 6$. "Булевые" ограничения для z_5 и z_6 являются функционально избыточными (т.е. логически следуют из предыдущих), однако, как видно из таблицы, без них ($m = 6$) получение нетривиальной двойственной квадратичной оценки невозможно.

Для случая $n = 10$ рассматривались такие ограничения: 6 ограничений из определений; 10 "булевых" ограничений (4 последних – избыточные); 2 избыточных ограничения из разложения $x_1 x_2 x_3 x_4$ (4-я степень); 8 избыточных ограничений из разложения одночленов $x_1 x_2 x_3$, $x_1 x_2 x_4$, $x_1 x_3 x_4$, $x_2 x_3 x_4$ (3-я степень). Итого 26 ограничений, 16 из которых избыточные. В табл. 4 даны значения оптимальных двойственных квадратичных оценок для 10 ограничений (без избыточных), 16 ограничений (добавлены все "булевые"), 18 ограничений (без разложений одночленов 3-й степени), 24 ограничений (без разложений одночлена 4-й степени), 26 (полный набор ограничений в случае десяти переменных).

Таблица 5.

w_1, w_2, w_3, w_4, w_5	f^*	ψ^*	$iter$
4,4,4,4,2	-16.0	-16.0	359
5,5,5,5,2	-20.0	-20.0	314
6,6,6,6,2	-24.0	-24.0	292
7,7,7,7,2	-28.0	-28.0	283
9,9,9,9,2	-36.0	-36.0	264
1,4,4,4,2	-14.0	-14.0	343
1,5,5,5,2	-17.0	-17.0	303
1,6,6,6,6	-20.0	-20.0	305
1,7,7,7,2	-23.0	-23.0	303
1,9,9,9,2	-29.0	-29.0	284

Для случая $n = 14$ рассматривались следующие ограничения: 10 ограничений из определений; 14 "булевых" ограничений, из которых 10 избыточных; 20 ограничений из неоднозначности разложения одночленов 4-й степени (все избыточные); 20 ограничений из неоднозначности разложения одночленов 3-й степени (все избыточные). Таким образом, всего 64 ограничения, 50 из которых избыточные. В табл. 4 представлены результаты в следующем порядке: для 14 ограничений (без избыточных), 24 ограничения (добавлены все "булевые"), 44 ограничения (без разложений одночленов 4-й степени), 44 ограничения (без разложений одночленов 3-й степени), 64 (полный набор ограничения для $n = 14$).

Для некоторых наборов w_1, w_2, w_3, w_4, w_5 , когда получена точная оценка для задачи (18)–(19), в табл. 5 представлены затраты по числу итераций ($iter$) для $r(\alpha)$ -алгоритма при нахождении оптимальных двойственных оценок в задаче квадратичного типа с 14 переменными и полным набором ограничений $m = 64$.

Как видно из табл. 5, нахождение двойственных оценок затрачено меньше, чем $6m = 6 \times 64 = 384$ итераций, что вполне соответствует оценке скорости сходимости $r(\alpha)$ -алгоритма.

В заключение отметим, что результаты экспериментов показали существенное улучшение квадратичных оценок при введении избыточных ограничений, высокую точность и устойчивость предложенного алгоритма для вычисления двойственных квадратичных оценок.

Литература

1. ШОР Н.З. Двойственные оценки в квадратичных оптимизационных задачах // Вестник Международного Соломонова университета, 1999. – № 1. – С. 92–108.
2. SHOR N.Z. Nondifferentiable optimization and polynomial problems. – Kluwer Academic Publishers. – 1998. – 394 p.
3. ШОР Н.З. Методы минимизации недифференцируемых функций и их приложения. – Киев: Наукова думка, 1979. – 199с.
4. ШОР Н.З., ЖУРБЕНКО Н.Г. Метод минимизации, использующий операцию растяжения пространства в направлении разности двух последовательных градиентов // Кибернетика. – 1971. – № 3. – С. 51–59.
5. ШОР Н.З., СТЕЦЮК П.И. Использование модификации r -алгоритма для нахождения глобального минимума полиномиальных функций // Кибернетика и систем. анализ. – 1997. – № 4. – С. 28–49.
6. Сборник научных программ на ФОРТРАНе. Руководство для программиста. Вып. 2. Матричная алгебра и линейная алгебра. – М.: Статистика, 1974. – 223 с.
7. HILBERT D. Ueber die Darstellung definitiver Formen als Summen von Formen quadraten // Math. Ann. – Leipzig, 1888. – Bd. 22. – P. 342–350.
8. VATTI R., PROTASI M. Approximate Algorithms and Heuristics for MAX-SAT // Handbook of Combinatorial Optimization (Vol. 1), D.Z. Du and P.M.Pardalos (Eds.), pp. 77–148, 1998, Kluwer Academic Publishers.

Dual Solution of Quadratic-Type Problems by r -algorithm (subroutine DSQTPr)

Naum Z. Shor and Petro I. Stetsyuk

Abstracts of Second International Workshop "Recent Advances in Non-Differentiable Optimization" (October 1-4, 2001, Kyiv, Ukraine).

General quadratic-type problems can be represented in the form

$$\begin{aligned} \text{to find} \quad & Q^* = Q_0(x^*) = \inf_{x \in R^n} Q_0(x), \\ \text{subject to} \quad & Q_i(x) = 0, \quad i = 1, \dots, m_1, \\ & Q_i(x) \leq 0, \quad i = m_1 + 1, \dots, m_1 + m_2, \end{aligned} \quad (1)$$

where quadratic functions $Q_i(x) = (K_i x, x) + (b_i, x) + c_i$ are given by real symmetric $n \times n$ matrices K_i , $i = 0, 1, \dots, m$, $m = m_1 + m_2$. Particularly, if $K_i \equiv 0$, then $Q_i(x)$ are linear functions.

Finding dual (Lagrangian) bound for the problem (1) can be reduce to the following problem (see [1]):

$$\text{to find } \psi^* = \psi(u^*) = \sup_{u \in U, K(u) \geq 0} \inf_{x \in R^n} \left\{ Q_0(x) + \sum_{i=1}^m u_i Q_i(x) \right\} \quad (2)$$

Here $u \in R^m$ is a vector of Lagrangian multipliers corresponding to constraints of (1); $U = \{u \in R^m : u_i \geq 0, i = m_1 + 1, \dots, m\}$; $K(u) \succeq 0$ means semidefiniteness of the family of symmetric matrices $K(u) = K_0 + \sum_{i=1}^m u_i K_i$.

For finding optimal ψ^* and u^* for the problem [2] the Fortran subroutine DSQTPr was realized using the modification of r -algorithm (see [2]). If $K(u^*) \succ 0$, then subroutine DSQTPr finds optimal solution Q^* and x^* for the problem (1).

The report will be include the description and characteristics of the subroutine DSQTPr and results of some test experiments for quadratic-type problems: convex ($m_1 = 0, K_0 \succ 0, K_i \succeq 0$ for all $i = 1, \dots, m$) and nonconvex ones.

Bibliography

1. SHOR NAUM Z. Nondifferentiable optimization and polynomial problems. Kluwer Academic Publishers, 1998, 394 p.
2. SHOR N.Z. AND STETSYUK P.I. The use of modification of r -algorithm for finding a global minimum of polynomial functions Kibernetika i Sistemnyi Analiz, 1997, # 4, 28–49.

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 bound techniques. The Lagrangian bounds may be improved for some important examples by adding in a model the 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 (the 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 dilation in the direction of difference of two successive subgradients. Below we give a 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:

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

Let $f(x)$ be a minimized convex function defined 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} , $g_f(x_k)$, 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 current number of steps and $q = \frac{1}{2}$. So, as a rule the convergence of r -algorithm is 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 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 at \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 $u_3 > \frac{a_5^2}{4}$ and u_1 is 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 of 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 by 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):

$$z_k - z_r z_{k-r} = 0 \text{ for } k \leq m, \quad r \leq \frac{k}{2}; \quad (16)$$

$$z_{k-m} z_m - z_r z_{k-r} = 0 \text{ for } m < k \leq 2m; \quad r \leq \frac{k}{2}; \quad k - r \neq m. \quad (17)$$

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 $K_0^{(h)}(z_1, \dots, z_m)$ representing 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 bound $\psi^*(h)$ 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

$$P(x) = \sum_{\alpha} c_{\alpha} R[\alpha] \text{ in new variables, } \quad 0 \leq \alpha_i \leq 2l_i, i = 1, \dots, n. \quad (19)$$

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 to $\alpha_1(\alpha)$, 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 $\overline{P}(x) = P(x) - P^*$ 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 (homogeneous 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 into a sum of squares, but to find such decomposition if it exists using the described above algorithm of finding Lagrangian 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

$$\alpha_w(G) = \max(w, x), \quad x = \{x_1, \dots, x_n\}, \quad (22)$$

$$x_i + x_j \leq 1 \text{ for all } (i, j) \in E, \quad (23)$$

$$x_k \in \{0, 1\} \text{ for all } k \in V. \quad (24)$$

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:

$$\alpha(G) \leq \bar{\chi}(G).$$

Similarly 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's 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 gives 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 $\psi(u) = \sup_x L(x, u)$ 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), $\max_{Y \in M_{n,k}} \text{tr}[A(u)Y^{(t)}]$, 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 by 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 obtained 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-otherswise; 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:

$$0 \leq y_e \leq 1. \quad (42)$$

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: $P_B(G) = \{y \in R^E | y \text{ satisfies all inequalities (42) and (43)}\}$ 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

$$W(u) = W + \text{diag } u,$$

($\text{diag } u$ 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

$$\lambda_{\min}^- = \begin{cases} 0, & \text{if } \lambda_{\min} \geq 0, \\ \lambda_{\min}, & \text{if } \lambda_{\min} < 0, \end{cases}$$

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:

$$\text{maximize } f(y) = \frac{1}{4} \left(\sum_{(i,j) \in E} w_{ij} - \sum_{i,j \in E} w_{ij} y_i y_j \right) \quad (48)$$

subject to constraints:

$$y_i^2 - 1 = 0, \quad \forall i \in V = \{1, \dots, n\}, \quad (49)$$

$$\sum_{i=1}^n y_i = 0 \quad (50)$$

(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):

$$\text{maximize } \bar{f}(\tilde{y}) = \frac{1}{4} \left(\sum_{(i,j) \in E} w_{ij} - \sum_{i,j \in E} [w_{ij} y_i y_j - \sum_{i=1}^{n-1} w_{in} (y_i y_n + y_n y_i)] \right) \quad (51)$$

subject to constraints:

$$y_i^2 - 1 = 0, \quad \forall i \in \{1, \dots, n-1\}, \quad (52)$$

$$\left(\sum_{i=1}^{n-1} y_i \right)^2 - 1 = 0. \quad (53)$$

One can construct Lagrange function for the problem (51)–(53) in the form:

$$L(\tilde{y}, u) = -\bar{f}(\tilde{y}) + \sum_{i=1}^{n-1} u_i (y_i^2 - 1) + u_n \left[\left(\sum_{i=1}^{n-1} y_i \right)^2 - 1 \right]$$

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

$$\psi(u) = -\frac{1}{4} \bar{W} + \sum_{i=1}^n u_i, \quad \text{if } u \in \text{dom } \psi,$$

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 to at least one, then \bar{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:

$$x_1^* = \sqrt{\frac{2t}{3} + \frac{\sqrt{t^2 + 3\varepsilon}}{3}}; \quad x_2^* = -\sqrt{\frac{2t}{3} + \frac{\sqrt{t^2 + 3\varepsilon}}{3}}.$$

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 $\lambda_2 > 0$, $\lambda_1 = \lambda_3 = 0$.

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 $P_6(x)$

Described 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 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 minimum 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 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^*

$2m$	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"

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. We obtain the quadratic-type problem (44)–(45), with a new constraint

$$\sum_{i=1}^n x_i = 0, \quad (60)$$

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"

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: $\alpha = 2$, $h_0 = 1.0$, $q_1 = 0.9$, $n_h = 3$, $q_2 = 1.1$ and $u_0 = \{1, 1, 1, 1, 1, 0, 0, 0, 0, 0\}$ for the 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 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 is 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 Universität Halle-Wittenberg, Mathematisch-Naturwissenschaftliche Reihe, pp. 114–115,.
6. BERGE, C. (1962), Sur une conjecture relative au probleme des codes optimaux, *Communication*, 13eme assemblee generale 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 solving of mathematical programming problem (interior point method: algorithm), 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: Proced. 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.

Р А З Д Е Л III

Новые семейства функционально избыточных ограничений в квадратичных булевых задачах

В раздел включены статьи:

- 3.1. О функционально избыточных ограничениях для булевых
оптимизационных задач квадратичного типа (*Стецюк, 2005*) .. 198
- 3.2. Новые модели квадратичного типа для задачи о максимальном ..
взвешенном разрезе графа (*Стецюк, 2006*) 204
- 3.3. Об уточнении лагранжевых двойственных оценок в бинарных ...
и булевых квадратичных задачах (*Стецюк, Пардалос, 2006*) ... 221
- 3.4. Об уточнении верхних оценок Н.З. Шора для числа
независимости графа (*Стецюк, 2007*) 230
- 3.5. Бинарный квадратичный многогранник и его
аппроксимации (*Стецюк, Золотых, 2010*) 232
- 3.6. О внешней аппроксимации бинарного и булевого
квадратичных многогранников (*Золотых, Стецюк, 2010*) 245

О функционально избыточных ограничениях для булевых оптимизационных задач квадратичного типа¹

П. И. Стецюк

Кибернетика и системный анализ. — 2005. — № 6. — С. 168—172.

Предложен способ построения функционально избыточных квадратичных ограничений в булевых задачах квадратичного типа. Он основан на расширении множества булевых переменных и формировании функционально избыточных ограничений для связи новых и старых переменных.

Функционально избыточные квадратичные ограничения играют важную роль при использовании техники лагранжевых двойственных квадратичных оценок в многоэкстремальных задачах квадратичного типа [1, 2]. Функционально избыточные ограничения в форме квадратичных уравнений или неравенств не изменяют оптимального значения исходной задачи квадратичного типа. Но добавление таких ограничений к исходной задаче приводит к новой функции Лагранжа для модифицированной задачи квадратичного типа, и может оказаться, что новая лагранжева двойственная квадратичная оценка лучше, чем лагранжева двойственная квадратичная оценка для исходной задачи.

С помощью булевых оптимизационных задач квадратичного типа формулируются такие экстремальные задачи на графах, как максимальный разрез графа, оптимальная бисекция графа, максимальное независимое множество вершин графа. Для этих задач в [1, 2, 3] исследованы различные семейства функционально избыточных квадратичных ограничений и показано, что их использование в ряде случаев улучшает лагранжевые двойственные квадратичные оценки. Как правило, эти случаи связаны со специальными типами графов. Методика формирования семейств функционально избыточных квадратичных ограничений ориентирована на специфику конкретной задачи и неудобна при исследовании этой же задачи для более общих случаев.

Оказывается, этого неудобства можно избежать, используя простой прием "расширения" множества булевых переменных. Он дает возмо-

¹Работа выполнена при частичной финансовой поддержке гранта UM2-2547-KV-03(CRDF Cooperative Grants Program).

жность генерировать функционально избыточные квадратичные ограничения, не принимая во внимание специфику задачи, для которой была сформулирована модель квадратичного типа. Более того, специфические особенности конкретной задачи могут оказаться простым следствием расшифровки содержательного смысла того или иного семейства функционально избыточных квадратичных ограничений.

О таком "автомате" для генерации функционально-избыточных ограничений в задачах квадратичного типа и пойдет речь в данной работе. Сделаем это на примере простейшей задачи квадратичного типа, которая связана с минимизацией квадратичной функции от n переменных x_1, \dots, x_n , каждая из которых может принимать значение либо $+1$, либо -1 . Задачу в булевых переменных 0-1 легко свести к задаче в булевых переменных ± 1 . Заметим, что ничего не изменится, если рассмотреть задачу квадратичного типа от булевых переменных в общем виде, т.е. при наличии квадратичных ограничений равенств или неравенств. Следовательно, рассмотренная ниже схема позволяет учитывать функционально избыточные ограничения, которые построены в [1, 2, 3] для ряда экстремальных задач на графах.

Итак, пусть имеется задача квадратичного типа в следующей формулировке:

$$f^* = f(x^*) = \min \left\{ f(x) = \sum_{i,j:1 \leq i < j \leq n} a_{ij} x_i x_j + \sum_{i=1}^n b_i x_i + c \right\} \quad (1)$$

при ограничениях:

$$x_i^2 = 1, i = 1, \dots, n. \quad (2)$$

Задачу (1)–(2) условимся называть задачей А.

Пусть ψ_A^* – лагранжева двойственная квадратичная оценка, соответствующая задаче А. Нахождение ψ_A^* сводится к задаче максимизации негладкой вогнутой функции, которая определена на параметрическом семействе неотрицательно определенных матриц, и может быть найдена с помощью современных методов недифференцируемой оптимизации (см. [1, 2, 3]). В общем случае $\psi_A^* \leq f^*$ и она служит нижней оценкой глобального минимума функции $f(x)$. Предположим, что $\psi_A^* < f^*$, иначе нам нечего будет улучшать.

Рассмотрим эту же задачу для расширенного множества переменных, где дополнительные переменные y_{ij} для каждого i и j такого, что $1 \leq i < j \leq n$ имеют такой вид: $y_{ij} = x_i x_j$. Обозначение дополнительных переменных y_{ij} (а не x_{ij}) сделано для удобства дальнейшего изложения

материала статьи. В терминах расширенного множества переменных задачу А можно сформулировать такой задачей квадратичного типа:

$$f^* = f(x^*) = \min \left\{ f(x) = \sum_{i,j:1 \leq i < j \leq n} a_{ij} x_i x_j + \sum_{i=1}^n b_i x_i + c \right\} \quad (3)$$

при ограничениях:

$$x_i^2 = 1, \quad i = 1, \dots, n, \quad (4)$$

$$y_{ij} = x_i x_j, \quad \forall i, j : 1 \leq i < j \leq n, \quad (5)$$

$$y_{ij}^2 = 1, \quad \forall i, j : 1 \leq i < j \leq n. \quad (6)$$

Задачу (3)–(6) условимся называть задачей В. Заметим, что целевую функцию в задаче В мы оставили такой же, как и в задаче А. Хотя с учетом дополнительных переменных ее можно сделать линейной функцией от переменных x и y , т.е.

$$F(x, y) = \sum_{i,j:1 \leq i < j \leq n} a_{ij} y_{ij} + \sum_{i=1}^n b_i x_i + c,$$

а также какой-либо комбинацией функций $f(x)$ и $F(x, y)$:

$$\Phi(x, y) = \lambda f(x) + (1 - \lambda) F(x, y).$$

Однако, для самих значений лагранжевых двойственных квадратичных оценок это не принципиально, хотя может оказаться существенным для алгоритмов нахождения этих оценок.

Пусть ψ_B^* – лагранжева двойственная квадратичная оценка, соответствующая задаче В. Очевидно, что $\psi_B^* = \psi_A^*$ и в силу сделанного предположения $\psi_B^* < f^*$. Это естественно, так как пока никакой дополнительной информации в задачу В по отношению к задаче А мы не привнесли. Однако это легко сделать, учитывая, что переменные y и x связаны между собой. Используя их связь, для задачи В можно построить два семейства функционально избыточных квадратичных ограничений. И если такие ограничения (или некоторые из них) добавить к задаче В, то это может улучшить лагранжеву двойственную квадратичную оценку, а также нижнюю оценку для f^* и в задаче А.

Опишем оба семейства функционально избыточных ограничений для задачи В. Первое семейство связано с различного типа ограничениями, которые можно построить для трех переменных x_i , x_j и x_k . Пусть индексы i , j и k различные. Тогда из цепочки равенств

$$x_i x_j x_k = (x_i x_j) x_k = y_{ij} x_k = (x_i x_k) x_j = y_{ik} x_j = (x_j x_k) x_i = y_{jk} x_i$$

следует группа функционально избыточных ограничений:

$$\begin{aligned} y_{ij}x_k - y_{ik}x_j &= 0, \\ y_{ij}x_k - y_{jk}x_i &= 0, \quad \forall i, j, k : 1 \leq i < j < k \leq n. \\ y_{ik}x_j - y_{jk}x_i &= 0, \end{aligned} \quad (7)$$

В частном случае, когда два индекса совпадают, имеем такую группу функционально избыточных ограничений:

$$\begin{aligned} x_i - y_{ij}x_j &= 0, \\ x_j - y_{ij}x_i &= 0, \quad \forall i, j : 1 \leq i < j \leq n. \end{aligned} \quad (8)$$

Группа ограничений (8) есть следствие равенств

$$\begin{aligned} x_i &= x_i x_j^2 = (x_i x_j) x_j = y_{ij} x_j, \\ x_j &= x_j x_i^2 = (x_i x_j) x_i = y_{ij} x_i, \end{aligned}$$

которые справедливы при любых i и j , таких что $1 \leq i < j \leq n$.

Второе семейство функционально избыточных ограничений связано со всевозможными ограничениями для четырех переменных x_i , x_j , x_k и x_l . Пусть все индексы различные. Тогда из цепочки равенств

$$\begin{aligned} x_i x_j x_k x_l &= (x_i x_j)(x_k x_l) = y_{ij} y_{kl} \\ &= (x_i x_k)(x_j x_l) = y_{ik} y_{jl} \\ &= (x_i x_l)(x_j x_k) = y_{il} y_{jk} \end{aligned}$$

следует такая группа функционально избыточных ограничений:

$$\begin{aligned} y_{ij} y_{kl} - y_{ik} y_{jl} &= 0, \\ y_{ij} y_{kl} - y_{il} y_{jk} &= 0, \quad \forall i, j, k, l : 1 \leq i < j < k < l \leq n. \\ y_{ik} y_{jl} - y_{il} y_{jk} &= 0. \end{aligned} \quad (9)$$

В случае, когда два индекса в четверке (i, j, k, l) совпадают, имеем следующую группу функционально избыточных ограничений:

$$\begin{aligned} y_{ij} - y_{ik} y_{jk} &= 0, \\ y_{ik} - y_{ij} y_{jk} &= 0, \quad \forall i, j, k : 1 \leq i < j < k \leq n. \\ y_{jk} - y_{ij} y_{ik} &= 0, \end{aligned} \quad (10)$$

Ограничения (10) являются следствием равенств

$$y_{ij} = x_i x_j = x_i x_j x_k^2 = (x_i x_k)(x_j x_k) = y_{ik} y_{jk},$$

$$y_{ik} = x_i x_k = x_i x_k x_j^2 = (x_i x_j)(x_j x_k) = y_{ij} y_{jk},$$

$$y_{jk} = x_j x_k = x_j x_k x_i^2 = (x_i x_j)(x_i x_k) = y_{ij} y_{ik},$$

которые справедливы при любых i, j и k , таких что $1 \leq i < j < k \leq n$.

Соответственно, добавляя к задаче В группы ограничений (7)–(10) (или некоторые отдельные ограничения из этих групп), получаем расширенные за счет функционально избыточных квадратичных ограничений задачи квадратичного типа. Соответствующие этим расширенным задачам лагранжевые двойственные квадратичные оценки могут быть лучше, чем ψ_B^* .

Проиллюстрируем такое улучшение на примере задачи А от трех переменных x_1, x_2 и x_3 , связанной с целевой функцией

$$f(x) = x_1 x_2 + x_1 x_3 + x_2 x_3 + x_1.$$

Глобальный минимум в задаче А равен -2 ($f^* = -2$). Лагранжевы двойственные квадратичные оценки ψ_A^* и ψ_B^* равны соответственно -2,50000. Добавление шести ограничений вида (8) улучшает лагранжеву двойственную квадратичную оценку и делает ее равной -2,22222. Добавление еще трех ограничений вида (7) делает оценку равной -2,04468. Если к задаче В прибавить три ограничения вида (10), то лагранжева двойственная квадратичная оценка становится точной и равна -2,00000. Двойственные оценки находились с помощью программы DSQTPR [4].

Следовательно, с помощью достаточно простого приема "расширения множества переменных" и использования полученных семейств функционально избыточных ограничений в ряде случаев можно улучшить двойственные лагранжевы квадратичные оценки. Понятно, что не всегда это приведет к успеху, даже если к задаче В добавить полный набор функционально избыточных ограничений. Но, процесс "расширения множества переменных" можно продолжить, применив описанную выше технику уже к задаче В с полным набором функционально избыточных ограничений. Это приведет к тому, что в "расширенное" множество булевых переменных будут входить произведения всех возможных троек $y_{ijk} = x_i x_j x_k$, всех возможных четверок $y_{ijkl} = x_i x_j x_k x_l$ и соответствующие им семейства функционально избыточных ограничений. Если этот процесс повторять и далее, то придем к переменным y , которые будут характеризовать произведения пяти, шести переменных и т.д. Вполне возможно, что подобная техника может привести к задаче квадратичного типа, для которой лагранжева двойственная квадратичная оценка совпадет с f^* .

Литература

1. SHOR N.Z. Nondifferentiable Optimization and Polynomial Problems, – Dordrecht, Kluwer. – 1998. – 394 p.
2. ШОР Н.З. Роль избыточных ограничений в улучшении двойственных оценок для полиномиальных оптимизационных задач // Кибернетика и системный анализ. 1998. – N4. – С.106-121.
3. SHOR N.Z., STETSYUK P.I. Lagrangian bounds in multiextremal polynomial and discrete optimization problems // Journal of Global Optimization, 2002, **23**. P. 1–41.
4. SHOR N.Z., 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", (October, 1-4, 2001, Kyiv, Ukraine). – P. 36.

Поступила 18.08.2005

Новые модели квадратичного типа для задачи о максимальном взвешенном разрезе графа¹

П. И. Стецюк

Кибернетика и системный анализ. — 2006. — № 1. — С. 63–75.

Предложены новые модели квадратичного типа для улучшения верхних оценок в задаче о максимальном взвешенном разрезе графа. Эти модели получены с помощью двух оригинальных способов построения функционально избыточных квадратичных ограничений. Из предложенных моделей вытекает известная линейная модель и рекомендации для построения ее "усиленных" аналогов.

Введение.

Пусть $G = (V, E)$ – взвешенный неориентированный граф со множеством вершин $V = \{1, \dots, n\}$ и множеством ребер E . Вес ребра $(i, j) \in E$ задан произвольным вещественным числом w_{ij} и с учетом неориентированности графа ребра (i, j) и (j, i) отождествляются, т.е. $w_{ij} = w_{ji}$. Без ограничения общности будем считать, что граф G – полный, так как ребрам дополнительного графа \bar{G} можно присвоить нулевой вес. Во избежание недоразумений в тех случаях когда нам понадобится неполный граф G , будем явно это указывать.

Задача о максимальном взвешенном разрезе графа (MaxCut-задача) состоит в следующем. Пусть V_1 и V_2 – непустые и непересекающиеся подмножества вершин графа G , такие что $V_1 \cup V_2 = V$. Разрезом графа G ($Cut(G)$) будет совокупность тех ребер, конечные вершины которых лежат в разных подмножествах. Вес разреза $Cut(G)$ определяется как сумма весов его ребер. MaxCut-задача заключается в разбиении множества вершин V на два таких подмножества V_1^* и V_2^* , чтобы соответствующий этому разбиению вес разреза $Cut(G)$ был максимальным. Максимальный вес разреза $Cut(G)$ обозначим $mc(G, w)$.

В общем случае MaxCut-задача является NP-трудной [1, с. 113]. В силу этого представляют интерес вопросы, связанные с уточнением верхних оценок для $mc(G, w)$. В настоящей работе рассматриваются два

¹Работа выполнена при частичной финансовой поддержке гранта UM2-2547-KV-03(CRDF Cooperative Grants Program).

типа верхних оценок для $mc(G, w)$: оценки, получаемые усовершенствованием моделей квадратичного типа, и оценки, получаемые с помощью линейных моделей (моделей линейного программирования).

Предложено три новые модели квадратичного типа для MaxCut-задачи, позволяющие находить более точные верхние оценки, чем простейшая модель квадратичного типа (описана в разделе 1). Все модели получены из простейшей модели двумя разными способами построения функционально избыточных квадратичных ограничений. Первые две модели (разделы 2 и 3) – одним способом, а третья модель (раздел 4) – другим. Из предложенных моделей вытекает известная линейная модель. Кроме того, третья модель дает рецепт для построения «усиленных» аналогов линейной модели.

1. Простейшая модель квадратичного типа.

Пусть вершине $i \in V$ ($i \in \{1, 2, \dots, n\}$) соответствует бинарная переменная $x_i \in \{-1, 1\}$, такая что

$$x_i = \begin{cases} -1, & \text{если } i \in V_1, \\ +1, & \text{если } i \in V_2. \end{cases}$$

Пусть $w_{ij} = w_{ji} = 0$, если $(i, j) \notin E$.

Простейшая модель квадратичного типа для MaxCut-задачи дана в [2, 3]. Она формулируется с помощью задачи квадратичного типа:

$$mc(G, w) = \max \left\{ f(x) = \frac{1}{2} \sum_{(i,j):i<j} w_{ij} (1 - x_i x_j) \right\} \quad (1)$$

при ограничениях:

$$x_i^2 = 1, \quad \forall i = 1, \dots, n. \quad (2)$$

Модель (1)–(2) назовем А-моделью. А-Модель является многоэкстремальной задачей нелинейного программирования, в которой целевая функция (1) и ограничения (2) – квадратичные «однородные» функции, т.е. не содержат линейных членов от переменных x_i , $i = 1, \dots, n$. Факт «однородности» квадратичной функции (1) понадобится для двух разных (не связанных по переменным) форм записи целевой функции в новых моделях.

Пусть $\psi_A^*(G, w)$ – оптимальная лагранжева двойственная квадратичная оценка, соответствующая А-модели. Здесь и далее символом ψ^* обозначим оптимальную лагранжеву двойственную квадратичную

оценку, т.е. наилучшую из класса лагранжевых двойственных квадратичных оценок [2], которая получается в результате использования техники множителей Лагранжа для моделей квадратичного типа. Для оптимизационной задачи квадратичного типа на максимум нахождение оценки ψ^* сводится к задаче минимизации выпуклой недифференцируемой функции, определенной на параметрическом (линейно зависящем от множителей Лагранжа) семействе неположительно определенных симметричных матриц. Оценка ψ^* может быть найдена с помощью методов недифференцируемой оптимизации [2].

При произвольных значениях весов ребер графа G оценку $\psi_A^*(G, w)$ можно найти за полиномиальное время с любой наперед заданной точностью. Однородность квадратичных функций (1) и (2) в A-модели позволяет вычислить $\psi_A^*(G, w)$ проще [3, 4], чем для общего случая оптимизационной задачи квадратичного типа [2].

Оценка $\psi_A^*(G, w)$ обладает свойством: $\psi_A^*(G, w) \geq mc(G, w)$, она является верхней оценкой для оптимального веса разреза $Cut(G)$. Однако, эффективность оценки $\psi_A^*(G, w)$ в ряде случаев недостаточна. Так, например, для графа K_3 (3-клика) с весами ребер, равными единице, $\psi_A^*(K_3) = 9/4 = 2,25$. Оптимальное значение разреза для этого же графа $mc(K_3) = 2$. Таким образом, относительная точность оценки $\psi_A^*(K_3)$ определяется величиной

$$\frac{\psi_A^*(K_3) - mc(K_3)}{mc(K_3)} = (2.25 - 2)/2 = 0.25/2 = 1/8$$

и составляет 12,5%.

В связи с этим естественно возникает вопрос: можно ли улучшить относительную точность двойственных лагранжевых квадратичных оценок для MaxCut-задачи по сравнению с оценкой $\psi_A^*(G, w)$? Один из инструментов изложен в [2] и связан с техникой использования функционально избыточных квадратичных ограничений. Такие ограничения (в форме квадратичных уравнений или неравенств) не изменяют оптимального значения исходной задачи квадратичного типа. Однако добавление функционально избыточных квадратичных ограничений к исходной задаче приводит к тому, что оценка ψ^* для модифицированной задачи, которой соответствует другая функция Лагранжа, может оказаться лучше, чем оценка ψ^* для исходной задачи.

Покажем, что техника использования функционально избыточных квадратичных ограничений приводит нас к новым моделям квадратичного типа для MaxCut-задачи. Соответствующие этим моделям двойственные лагранжевые квадратичные оценки в ряде случаев лучше, чем

оценка $\psi_A^*(G, w)$. При построении новых моделей для MaxCut-задачи используются два способа для генерации функционально избыточных квадратичных ограничений. Оба способа могут применяться для общего случая задачи квадратичного типа с бинарными (а значит и булевыми) переменными.

Первый способ, изложенный в [5], связан с использованием четырех типов функционально избыточных квадратичных ограничений для «расширенных» (по отношению к исходной модели) моделей квадратичного типа. Такие «расширенные» модели получены за счет введения новых переменных вида $y_{ij} = x_i x_j$, а четыре типа функционально избыточных квадратичных ограничений (в [5] они объединены в два семейства) описывают различные виды «избыточных» квадратичных зависимостей между новыми переменными y и старыми переменными x . Далее нам понадобится только одна группа ограничений из второго семейства функционально избыточных ограничений, а именно ограничения (10) из [5]. Для новых переменных $y_{ij} = x_i x_j$ эта группа функционально избыточных квадратичных ограничений является следствием таких очевидных равенств

$$y_{ij} = x_i x_j = x_i x_j x_k^2 = (x_i x_k)(x_j x_k) = y_{ik} y_{jk}, \quad (3)$$

$$y_{ik} = x_i x_k = x_i x_k x_j^2 = (x_i x_j)(x_j x_k) = y_{ij} y_{jk}, \quad (4)$$

$$y_{jk} = x_j x_k = x_j x_k x_i^2 = (x_i x_j)(x_i x_k) = y_{ij} y_{ik}, \quad (5)$$

которые с учетом (2) справедливы при произвольных индексах i, j и k .

Второй новый способ связан с использованием «нечетного» количества бинарных переменных. Идея этого способа подробно описана в разд. 4. Он позволяет строить семейства функционально избыточных ограничений-неравенств для произвольных «нечетных» подмножеств из полного множества бинарных переменных исходной задачи. Использование таких функционально избыточных ограничений дает рекомендации, как для MaxCut-задачи можно строить «расширенные» (по количеству ограничений) линейные модели, которые имеют более сильные утверждения о точности линейных оценок, чем известный результат из [6].

2. Новая модель квадратичного типа.

Пусть $y_{ij} = x_i x_j$ — новые переменные для всех i и j , таких что $1 \leq i < j \leq n$. Количество новых переменных y равно $\frac{n(n-1)}{2}$.

Новую модель для MaxCut-задачи сформулируем с помощью задачи квадратичного типа:

$$mc(G, w) = \max \left\{ f(x) = \frac{1}{2} \sum_{(i,j): i < j} w_{ij} (1 - x_i x_j) \right\} \quad (6)$$

при ограничениях:

$$x_i^2 = 1, \quad i = 1, \dots, n, \quad (7)$$

$$y_{ij} = x_i x_j, \quad \forall i, j : 1 \leq i < j \leq n, \quad (8)$$

$$y_{ij}^2 = 1, \quad \forall i, j : 1 \leq i < j \leq n, \quad (9)$$

$$\begin{aligned} y_{ik} y_{jk} - y_{ij} &= 0, \\ y_{ij} y_{jk} - y_{ik} &= 0, \quad \forall i, j, k : 1 \leq i < j < k \leq n. \\ y_{ij} y_{ik} - y_{jk} &= 0, \end{aligned} \quad (10)$$

Модель (6)–(10) назовем В-моделью. Она является «расширенным» (за счет добавления новых переменных и новых ограничений) аналогом А-модели. Здесь (6) и (7) в точности совпадают с (1) и (2). Новые ограничения имеют следующий смысл. Ограничения (8) связаны с введением новых переменных y . Ограничения (9) соответствуют «бинарности» переменных y и следуют из правила их введения, так как в силу (2) для любых i, j , таких что $1 \leq i < j \leq n$, всегда справедливо равенство

$$y_{ij}^2 = (x_i x_j)^2 = x_i^2 x_j^2 = 1.$$

Ограничения (10) задают функционально избыточные квадратичные ограничения, которые следуют из (3)–(5).

Пусть $\psi_B^*(G, w)$ – оптимальная лагранжева двойственная квадратичная оценка, соответствующая В-модели. Оценка $\psi_B^*(G, w)$, точно также как и оценка $\psi_A^*(G, w)$, является верхней оценкой для $mc(G, w)$, и кроме того $\psi_B^*(G, w) \leq \psi_A^*(G, w)$. Из-за использования функционально избыточных ограничений (10) в ряде случаев $\psi_B^*(G, w)$ оказывается более точной верхней оценкой для $mc(G, w)$, чем $\psi_A^*(G, w)$. Подтверждением этому служит пример MaxCut-задачи для графа K_3 с единичными весами ребер. Так, в этом случае $\psi_B^*(K_3) = 2$ и является точной верхней оценкой для $mc(K_3)$. Заметим, что относительная точность для для оценки $\psi_A^*(K_3)$ составляла 12,5%.

Наличие в В-модели функционально избыточных квадратичных ограничений (10) в ряде случаев способствует улучшению верхних оценок для $mc(G, w)$. Этому факту дадим два объяснения.

Первое, техническое, состоит в следующем. Если из В-модели убрать ограничения (10), то оптимальная двойственная лагранжева квадратичная оценка совпадает с $\psi_A^*(G, w)$, и никакого улучшения верхней оценки для $mc(G, w)$ не получим. Функционально избыточные квадратичные ограничения (10) описывают определенного рода квадратичные зависимости между всевозможными тройками переменных x_i , x_j и x_k , т.е. эти зависимости будут отражены в функции Лагранжа, соответствующей В-модели. Следовательно, оценка $\psi_B^*(G, w)$ может оказаться лучше, чем оценка $\psi_A^*(G, w)$, которая соответствует А-модели, где подобные зависимости для троек переменных x_i , x_j и x_k отсутствуют.

Второе объяснение использует содержательный смысл функционально избыточных квадратичных ограничений. Новая переменная $y_{ij} = x_i x_j$ означает не что иное, как бинарная «реберную» переменная, такую что

$$y_{ij} = \begin{cases} -1, & \text{если ребро } (i, j) \in Cut(G), \\ +1, & \text{если ребро } (i, j) \notin Cut(G). \end{cases} \quad (11)$$

Такое свойство переменной y_{ij} следует из содержательного смысла «вершинных» переменных x_i и x_j . На самом деле, если ребро (i, j) принадлежит разрезу $Cut(G)$, то вершины i и j принадлежат разным подмножествам, т.е. $i \in V_2$ и $j \in V_1$, либо $i \in V_1$ и $j \in V_2$. Это означает, что пара «вершинных» переменных x_i и x_j может принимать два возможных «бинарных» исхода: $(x_i, x_j) = (1, -1)$, либо $(x_i, x_j) = (-1, 1)$, откуда в обоих случаях имеем $y_{ij} = x_i x_j = -1$. Если ребро (i, j) не принадлежит разрезу $Cut(G)$, то обе вершины i и j принадлежат одному и тому же подмножеству: т.е. $i, j \in V_2$ либо $i, j \in V_1$. Это означает, что та же пара «вершинных» переменных будет: либо $(x_i, x_j) = (1, 1)$, либо $(x_i, x_j) = (-1, -1)$, откуда в обоих случаях имеем $y_{ij} = x_i x_j = 1$.

Далее, пусть в графе G выделен подграф $K_3(i, j, k)$ (3-клика), построенный для трех фиксированных вершин i, j, k и пусть y_{ij}, y_{ik}, y_{jk} — «реберные» переменные, соответствующие трем его ребрам (i, j) , (i, k) и (j, k) . Тройка «реберных» переменных (y_{ij}, y_{ik}, y_{jk}) для $K_3(i, j, k)$ может принимать всего четыре различных «бинарных» значения, а именно

$$(y_{ij}, y_{ik}, y_{jk}) = \begin{pmatrix} +1, +1, +1 \\ +1, -1, -1 \\ -1, +1, -1 \\ -1, -1, +1 \end{pmatrix}.$$

Первое значение означает, что ни одно из ребер $K_3(i, j, k)$, не попадает в разрез $Cut(G)$, или, иначе, все три вершины i, j и k принадлежат одному из подмножеств V_1 , либо V_2 . Оставшиеся три значения задают все

возможные комбинации для пар ребер неориентированного подграфа $K_3(i, j, k)$ (их будет $C_3^2 = \frac{3 \cdot 2}{2} = 3$) и означают что, разрезу могут принадлежать только два, но произвольных ребра из трех. Другими словами, одна (но произвольная) вершина из подграфа $K_3(i, j, k)$ принадлежит одному подмножеству, а две оставшиеся вершины – другому.

Легко убедиться, что все четыре приведенные «бинарных» значения для тройки «реберных» переменных (y_{ij}, y_{ik}, y_{jk}) удовлетворяют ограничениям (10) при соответствующих этой тройке индексах i, j и k . Следовательно, для MaxCut-задач функционально избыточные квадратичные ограничения (10) отражают простой факт. А именно, для всех возможных подграфов $K_3(i, j, k)$ из графа G разрезу $Cut(G)$ могут принадлежать либо два (произвольных) ребра подграфа $K_3(i, j, k)$, либо ни одно из ребер $K_3(i, j, k)$ не может принадлежать разрезу $Cut(G)$. Других вариантов нет. Очевидно, что если в графе G существует хотя бы один подграф $K_3(i, j, k)$, для которого ребра в оптимальном разрезе определяются однозначно, то тогда эта информация будет содержательной для В-модели и следует ожидать, что оценка $\psi_B^*(G, w) < \psi_A^*(G, w)$.

Итак, В-модель одновременно использует «вершинные» и «реберные» переменные, и их взаимосвязь выражена содержательным смыслом функционально избыточных ограничений (10). И хотя оценка $\psi_B^*(G, w)$ в ряде случаев более точная, но ее нахождение намного сложнее, чем нахождение оценки $\psi_A^*(G, w)$ для А-модели. Ниже проведем «упрощение» В-модели, чтобы уменьшить количество переменных, и проанализируем полученную в результате модель.

3. «Реберная» модель квадратичного типа и ее анализ

Упростить В-модель нам помогут следующие соображения.

Заметим, что целевая функция в В-модели осталась такой же, как и в А-модели. Однако учитывая «однородность» функции $f(x)$, ее можно записать как линейную функцию, которая зависит только от «реберных» переменных y , т.е.

$$f(x) = F(y) = \sum_{i,j:i < j} w_{ij} (1 - y_{ij}).$$

Если в В-модели целевой функцией вместо $f(x)$ сделать функцию $F(y)$, то это не повлияет на значение оптимальной лагранжевой двойственной квадратичной оценки, и она будет равна оценке $\psi_B^*(G, w)$. Однако в результате такой замены в В-модели связь между переменными x и y остается только в ограничениях (8), которые связаны с введением новых (реберных) переменных y . Следовательно, если из В-модели убрать

ограничения (8), то можно опустить и ограничения (7), которые станут несущественными. Действуя таким образом, можно из В-модели убрать переменные x и оставить только «реберные» переменные y .

Итак, новую «реберную» модель для MaxCut-задачи (С-модель) сформулируем следующей моделью квадратичного типа:

$$F^*(G, w) = \max \left\{ F(y) = \frac{1}{2} \sum_{(i,j): i < j} w_{ij} (1 - y_{ij}) \right\} \quad (12)$$

при ограничениях:

$$\begin{aligned} y_{ik}y_{jk} - y_{ij} &= 0, \\ y_{ij}y_{jk} - y_{ik} &= 0, \\ y_{ij}y_{ik} - y_{jk} &= 0, \end{aligned} \quad \forall i, j, k : 1 \leq i < j < k \leq n. \quad (13)$$

$$y_{ij}^2 = 1, \quad \forall i, j : i < j. \quad (14)$$

Очевидно, что $F^*(G, w) \geq mc(G, w)$ из-за игнорирования переменных x из В-модели. Это значит, что значение $F^*(G, w)$ можно использовать в качестве верхней оценки для $mc(G, w)$. Пусть $\psi_C^*(G, w)$ – оптимальная лагранжева двойственная квадратичная оценка, соответствующая С-модели. Оценка $\psi_C^*(G, w)$ – верхняя оценка для $F^*(G, w)$, а тем самым и для $mc(G, w)$, т.е. $\psi_C^*(G, w) \geq mc(G, w)$.

Рассмотрим качество оценки $\psi_C^*(G, w)$. Так, например, $\psi_C^*(K_3) = 2$ и она так же, как и $\psi_B^*(K_3)$, является точной верхней оценкой для $mc(K_3)$. Это не удивительно, поскольку далее покажем, что из С-модели можно получить линейную модель из [6], которая считается одним из сильных результатов для верхних оценок в MaxCut-задачах.

В терминах «бинарных» реберных переменных линейная модель из [6] формулируется с помощью задачи линейного программирования [7, с.37]:

$$f_{BM}^*(G, w) = \max \left\{ \frac{1}{2} \sum_{(i,j): i < j} w_{ij} (1 - y_{ij}) \right\} \quad (15)$$

при ограничениях:

$$\begin{aligned} y_{ij} + y_{ik} + y_{jk} &\geq -1, \\ y_{ij} - y_{ik} - y_{jk} &\geq -1, \\ -y_{ij} + y_{ik} - y_{jk} &\geq -1, \\ -y_{ij} - y_{ik} + y_{jk} &\geq -1, \end{aligned} \quad \forall i, j, k : 1 \leq i < j < k \leq n. \quad (16)$$

Здесь $f_{BM}^*(G, w)$ – оптимальное значение линейной программы (15)–(16) и линейная оценка сверху для $mc(G, w)$. В том случае, когда неполный граф G принадлежит семейству графов, нестягиваемых к K_5 (5-клика), то $f_{BM}^*(G, w) = mc(G, w)$ и, следовательно, $f_{BM}^*(G, w)$ есть точной верхней оценкой для $mc(G, w)$.

Заметим, что линейная модель в форме (15)–(16) носит скорее теоретический характер, чем практический. Так, например, практически реализуемый вариант этой линейной модели формулируется в терминах булевых (0-1) «реберных» переменных и для неполного графа G . Именно для такой модели в [8] на основе метода эллипсоидов построен полиномиальный алгоритм для специального случая MaxCut-задачи (G – слабо двудольный граф с положительными весами ребер).

Покажем, как из S -модели получается линейная модель (15)–(16). Целевая функция в обеих моделях одна и та же. Следовательно, требуется только показать, что из ограничений (13) и (14) следуют ограничения (16).

Для произвольной тройки индексов i, j, k , таких что $1 \leq i < j < k \leq n$ из ограничений (13) легко получить соотношения:

$$2y_{ik}y_{jk} + 2y_{ij}y_{jk} + 2y_{ij}y_{ik} - 2(y_{ij} + y_{ik} + y_{jk}) = 0, \quad (17)$$

$$2y_{ik}y_{jk} - 2y_{ij}y_{jk} - 2y_{ij}y_{ik} - 2(y_{ij} - y_{ik} - y_{jk}) = 0, \quad (18)$$

$$-2y_{ik}y_{jk} + 2y_{ij}y_{jk} - 2y_{ij}y_{ik} - 2(-y_{ij} + y_{ik} - y_{jk}) = 0, \quad (19)$$

$$-2y_{ik}y_{jk} - 2y_{ij}y_{jk} + 2y_{ij}y_{ik} - 2(-y_{ij} - y_{ik} + y_{jk}) = 0. \quad (20)$$

Данные равенства получены умножением на два четырех линейных комбинаций, построенных из трех ограничений из (13). Так, (17) получено в результате сложения всех трех ограничений, (18) – вычитания из первого второго и третьего ограничений, (19) – вычитания из второго первого и третьего ограничений, а (20) – результат вычитания из третьего второго и первого ограничений из (13).

Учитывая, что $y_{ij}^2 = 1$, $y_{ik}^2 = 1$ и $y_{jk}^2 = 1$, равенство (17) можно записать как

$$y_{ij}^2 + y_{ik}^2 + y_{jk}^2 + 2y_{ik}y_{jk} + 2y_{ij}y_{jk} + 2y_{ij}y_{ik} - 2(y_{ij} + y_{ik} + y_{jk}) - 3 = 0,$$

или

$$(y_{ij} + y_{ik} + y_{jk})^2 - 2(y_{ij} + y_{ik} + y_{jk}) + 1 - 4 = 0,$$

$$(y_{ij} + y_{ik} + y_{jk} - 1)^2 = 4.$$

Из последнего равенства следует, что

$$y_{ij} + y_{ik} + y_{jk} = 3, \quad \text{либо} \quad y_{ij} + y_{ik} + y_{jk} = -1.$$

Таким образом, для суммы реберных переменных y_{ij} , y_{ik} и y_{jk} можно записать линейное неравенство

$$y_{ij} + y_{ik} + y_{jk} \geq \min\{-1, +3\} = -1,$$

которое совпадает с первым из ограничений в (16).

Аналогичные рассуждения имеют место и для равенств (18)–(20). В результате имеем равенства:

$$(y_{ij} - y_{ik} - y_{jk} - 1)^2 = 4,$$

$$(-y_{ij} + y_{ik} - y_{jk} - 1)^2 = 4,$$

$$(-y_{ij} - y_{ik} + y_{jk} - 1)^2 = 4,$$

из которых, так же как и выше, следуют оставшиеся три линейных ограничения из (16).

Итак, путем некоторого «огрубления» С-модели получена линейная модель (15)–(16). Это означает, что соответствующие новым В-модели и С-модели верхние оценки для MaxCut-задачи должны совмещать преимущества линейной модели из [6] и простейшей модели квадратичного типа (А-модели). Так, например, для графа K_5 (5-клика) с единичными весами ребер А-модель дает более точную оценку ($\psi_A^*(K_5) = 25/4$), чем линейная модель ($f_{BM}^*(K_5) = 20/3$). Оценка $\psi_B^*(K_5) = 25/4$ и оценка $\psi_C^*(K_5) = 25/4$, а, следовательно, они совпадают с более точной оценкой $\psi_A^*(K_5)$. В то же время для графа C_5 (пятивершинный цикл) с единичными весами ребер ситуация с точностью до наоборот. Здесь линейная модель дает лучше оценку ($f_{BM}^*(C_5) = 4$), чем А-модель ($\psi_A^*(C_5) = \frac{5}{2}(1 + \cos \frac{\pi}{5}) \approx 4,522542$). Оценка $\psi_B^*(C_5) = 4$, оценка $\psi_C^*(C_5) = 4$, а, следовательно, они также совпадают с более точной оценкой $f_{BM}^*(C_5)$.

Отметим, что к линейной модели (15)–(16) можно прийти от модели квадратичного типа, использующей другой тип функционально избыточных квадратичных ограничений. Такую модель рассмотрим ниже. Она объединяет лучшие характеристики линейной модели и простейшей модели квадратичного типа (А-модели). Кроме того, появляется возможность «расширения» линейных моделей (построенных по типу линейной модели (15)–(16)), но с более сильными утверждениями о точности линейных оценок.

4. Новая «вершинная» модель квадратичного типа

В основу данной модели положена простая идея для построения функционально избыточных квадратичных ограничений в моделях квадратичного типа с «бинарными» переменными, основанная на использовании «нечетного» количества бинарных переменных. Она состоит в следующем.

Рассмотрим «нечетной» длины последовательность $b_1, b_2, \dots, b_{2k+1}$ из $(2k+1)$ бинарных (± 1) чисел. Здесь $k = 1, 2, 3, \dots$ — натуральные числа. Очевидно, что сумма всех членов такой последовательности $S_{2k+1} = b_1 + b_2 + \dots + b_{2k+1}$ не может быть равна нулю, из каких бы бинарных чисел эта последовательность не была составлена. Связано это с тем, что S_{2k+1} получается в результате суммирования нечетного количества ± 1 . Самые близкие к нулю целые значения суммы S_{2k+1} такие: $S_{2k+1} = 1$, когда ровно $(k+1)$ произвольных членов последовательности равны $+1$, а оставшиеся k членов равны -1 ; $S_{2k+1} = -1$, когда ровно k членов последовательности равны $+1$ и оставшиеся $(k+1)$ ее членов равны -1 .

Это означает, что для произвольного нечетного количества бинарных переменных $x_1, x_2, \dots, x_{2k+1}$ всегда справедливо следующее неравенство:

$$\left(\sum_{i=1}^{2k+1} l_i x_i \right)^2 = (l_1 x_1 + l_2 x_2 + \dots + l_{2k+1} x_{2k+1})^2 \geq 1, \quad (21)$$

где $l_1, l_2, \dots, l_{2k+1}$ — произвольная последовательность, состоящая из бинарных чисел (± 1) . Для фиксированного k максимально возможное количество таких последовательностей равно $2^{2k+1} = 2 \cdot 4^k$. Однако, последовательность $l_1, l_2, \dots, l_{2k+1}$ и последовательность $-l_1, -l_2, \dots, -l_{2k+1}$ (т.е. получена из первой умножением на -1) дают одно и то же неравенство в форме (1D). Следовательно, для фиксированного k полное семейство неравенств вида (21) будет включать 4^k квадратичных неравенств.

Неравенство (21) естественно использовать для построения функционально избыточных квадратичных ограничений в моделях квадратичного типа с бинарными переменными. Алгоритм здесь очень простой. Пусть имеется исходная модель квадратичного типа с n бинарными переменными, которые описаны с помощью ограничений $x_i^2 = 1, i = 1, \dots, n$. Новую модель квадратичного типа из исходной можно получить путем добавления функционально избыточных ограничений в форме (21) для всех или некоторых «нечетных» подмножеств переменных из полного множества бинарных переменных. Самая полная модель

(включающая все возможные ограничения в форме (21)) содержит $N(s)$ функционально избыточных ограничений:

$$N_s = \sum_{k=1}^K 4^k C_n^{2k+1}, \text{ где } K = \left\lfloor \frac{n-1}{2} \right\rfloor.$$

Это равносильно использованию полных семейств функционально избыточных ограничений-неравенств в форме (21) для всех возможных «нечетных» подмножеств переменных из полного множества n бинарных переменных. Новым моделям будут сопутствовать более точные оценки ψ^* , чем оценка ψ^* для исходной модели (это зависит от выбора нечетных подмножеств).

Квадратичным неравенствам (21) можно придать и другую форму записи. Учитывая, что $l_i^2 = 1$ и $x_i^2 = 1$ для всех $i = 1, \dots, 2k+1$, правую часть неравенства (21) можно записать следующим образом:

$$\left(\sum_{i=1}^{2k+1} l_i x_i \right)^2 = \sum_{i=1}^{2k+1} l_i^2 x_i^2 + 2 \sum_{i=1}^{2k+1} \sum_{j>i}^{2k+1} l_i l_j x_i x_j = 2k+1 + 2 \sum_{i=1}^{2k+1} \sum_{j>i}^{2k+1} l_i l_j x_i x_j.$$

Тогда неравенство (21) можно переписать так:

$$2k+1 + 2 \sum_{i=1}^{2k+1} \sum_{j>i}^{2k+1} l_i l_j x_i x_j \geq 1,$$

$$2 \sum_{i=1}^{2k+1} \sum_{j>i}^{2k+1} l_i l_j x_i x_j \geq -2k,$$

откуда следует запись функционально избыточного квадратичного ограничения в такой форме:

$$\sum_{i=1}^{2k+1} \sum_{j>i}^{2k+1} l_i l_j x_i x_j \geq -k. \quad (22)$$

Неравенство (22) оказывается удобным при переходе от моделей квадратичного типа к линейным моделям, так как оно уже адаптировано под использование реберных переменных $y_{ij} = x_i x_j$, а именно:

$$\sum_{i=1}^{2k+1} \sum_{j>i}^{2k+1} l_i l_j y_{ij} \geq -k. \quad (23)$$

При определенном выборе «нечетных» подмножеств новым моделям квадратичного типа должны соответствовать более точные оценки ψ^* , чем оценка ψ^* для исходной модели. Далее покажем, к чему приводит использование максимально возможного количества функционально избыточных квадратичных ограничений (в форме (21), либо в форме (22)) для MaxCut-задачи.

Так, для $k = 1$ (соответствует использованию тройки «реберных» переменных x_i, x_j и x_k) полное семейство функционально избыточных ограничений в форме (21) содержит четыре неравенства

$$\begin{aligned} (+x_i + x_j + x_k)^2 &\geq 1, \\ (-x_i + x_j + x_k)^2 &\geq 1, \\ (+x_i - x_j + x_k)^2 &\geq 1, \\ (+x_i + x_j - x_k)^2 &\geq 1. \end{aligned} \quad (24)$$

Если в А-модель добавить ограничения (24), переписав их в форме (22), и сделать это для всех возможных «троек» вершинных переменных, то получим модель квадратичного типа (D-модель) вида

$$mc(G, w) = \max \left\{ f(x) = \frac{1}{2} \sum_{(i,j): i < j} w_{ij} (1 - x_i x_j) \right\} \quad (25)$$

при ограничениях:

$$x_i^2 = 1, \quad \forall i = 1, \dots, n. \quad (26)$$

$$\begin{aligned} x_i x_j + x_i x_k + x_j x_k &\geq -1, \\ -x_i x_j - x_i x_k + x_j x_k &\geq -1, \\ -x_i x_j + x_i x_k - x_j x_k &\geq -1, \\ x_i x_j - x_i x_k - x_j x_k &\geq -1, \end{aligned} \quad \forall i, j, k : 1 \leq i < j < k \leq n. \quad (27)$$

Пусть D-модели соответствует двойственная лагранжева квадратичная оценка $\psi_D^*(G, w)$. Заметим, что для $mc(G, w)$ она реализует верхнюю оценку, наиболее точную из двух оценок $\psi_A^*(G, w)$ и f_{BM}^* , как это имело место и для моделей В и С.

От D-модели легко перейти к линейной модели (15)–(16). Для этого достаточно целевую функцию $f(x)$ в (25) заменить на $F(y)$, а функционально избыточные ограничения (27) – на их аналоги в реберных переменных $y_{ij} = x_i x_j$. В результате, целевая функция из (25) перейдет в целевую функцию (15), а ограничения (27) примут вид

$$\begin{aligned} y_{ij} + y_{ik} + y_{jk} &\geq -1, \\ -y_{ij} - y_{ik} + y_{jk} &\geq -1, \\ -y_{ij} + y_{ik} - y_{jk} &\geq -1, \\ y_{ij} - y_{ik} - y_{jk} &\geq -1, \end{aligned} \quad \forall i, j, k : 1 \leq i < j < k \leq n, \quad (28)$$

который соответствует записи полного семейства функционально избыточных ограничений в форме (23) для всех «троек» бинарных переменных из n . Легко видеть, что ограничения (28) с точностью до их порядка (второе и четвертое ограничения следует поменять местами) соответствуют ограничениям (16) из линейной модели. В результате этих замен ограничения (26) становятся несущественными и, убрав их, приходим к модели (15)–(16).

Итак, из D-модели легко получается известная линейная модель из [6]. Однако, имеется огромный резерв для «расширения» D-модели за счет функционально избыточных квадратичных ограничений для оставшихся «нечетных» подмножеств бинарных переменных, т.е. всех возможных «пятерок», «семерок» вершинных переменных и т.д. Это приводит к тому, что в определенных случаях оценки ψ^* для «расширенных» D-моделей будут улучшать точность, но уже оценки $\psi_D^*(G, w)$. Оказывается, что на основе таких «расширенных» D-моделей легко строить «расширенные» (по количеству ограничений) аналоги линейной модели (15)–(16), которым будут соответствовать более сильные утверждения об точности линейных оценок в MaxCut-задачах.

Полное семейство функционально избыточных квадратичных ограничений для каждой «пятерки» (здесь $k = 2$) вершинных переменных x_i, x_j, x_k, x_l и x_m , где $1 \leq i < j < k < l < m \leq n$, состоит из $16 = 4^2$ квадратичных ограничений-неравенств. В форме (21) это такие ограничения:

$$\begin{aligned}
 (+x_i + x_j + x_k + x_l + x_m)^2 &\geq 1, & (-x_i + x_j + x_k - x_l + x_m)^2 &\geq 1, \\
 (-x_i + x_j + x_k + x_l - x_m)^2 &\geq 1, & (-x_i + x_j + x_k + x_l + x_m)^2 &\geq 1, \\
 (+x_i - x_j + x_k + x_l + x_m)^2 &\geq 1, & (+x_i - x_j - x_k + x_l + x_m)^2 &\geq 1, \\
 (+x_i + x_j - x_k + x_l + x_m)^2 &\geq 1, & (+x_i - x_j + x_k - x_l + x_m)^2 &\geq 1, \\
 (+x_i + x_j + x_k - x_l + x_m)^2 &\geq 1, & (+x_i - x_j + x_k + x_l - x_m)^2 &\geq 1, \\
 (+x_i + x_j + x_k + x_l - x_m)^2 &\geq 1, & (+x_i + x_j - x_k - x_l + x_m)^2 &\geq 1, \\
 (-x_i - x_j + x_k + x_l + x_m)^2 &\geq 1, & (+x_i + x_j - x_k + x_l - x_m)^2 &\geq 1, \\
 (-x_i + x_j - x_k + x_l + x_m)^2 &\geq 1, & (+x_i + x_j + x_k - x_l - x_m)^2 &\geq 1.
 \end{aligned} \tag{29}$$

Если ограничения (29) записать в форме (22) и прибавить их к D-модели для всех возможных «пятерок» вершинных переменных, то получим «расширенную» D-модель, где полные семейства функционально избыточных квадратичных ограничений используются для таких «нечетных» подмножеств бинарных переменных, как всевозможные «тройки» и всевозможные «пятерки» реберных переменных. «Расширенная» D-модель содержит $4C_n^3 + 16C_n^5$ функционально избыточных квадратичных ограничений, и соответствующая ей оценка ψ^* в определенных случаях будет лучше, чем оценка $\psi_D^*(G, w)$.

«Расширенной» D-модели будет соответствовать «расширенный» аналог линейной модели из [6]. Он получается в результате прибавления к линейной модели (15)–(16) линейных ограничений-неравенств, которые в «расширенной» D-модели соответствуют функционально избыточным ограничениям для «пятерок», записанных с помощью реберных переменных в форме (23). Таких линейных ограничений будет $16C_n^5$. Так, например, при заданных i, j, k, l, m , таких что $1 \leq i < j < k < l < m \leq n$, первому ограничению из (29) соответствует линейное неравенство:

$$y_{ij} + y_{ik} + y_{il} + y_{im} + y_{jk} + y_{jl} + y_{jm} + y_{kl} + y_{km} + y_{lm} \geq -2,$$

второму – линейное неравенство

$$-y_{ij} - y_{ik} - y_{il} - y_{im} + y_{jk} + y_{jl} + y_{jm} + y_{kl} + y_{km} + y_{lm} \geq -2,$$

и т.д., а последнему –

$$y_{ij} + y_{ik} - y_{il} - y_{im} + y_{jk} - y_{jl} - y_{jm} - y_{kl} - y_{km} + y_{lm} \geq -2.$$

«Расширенная» линейная модель дает точную верхнюю оценку для $mc(G, w)$ тогда, когда неполный граф G принадлежит семейству графов, нестягиваемых к K_7 (7-клика).

«Расширение» D-модели можно продолжить, дополняя ее еще и полным семейством функционально избыточных ограничений для всех возможных «семерок» ($k = 3$) реберных переменных. При этом прибавится $4^3 C_n^7 = 64C_n^7$ функционально избыточных квадратичных ограничений. «Расширенной» D-модели (за счет всех «пятерок» и «семерок») соответствует «расширенный» аналог линейной модели, который будет включать еще $64C_n^7$ линейных ограничений-неравенств, которые характеризуют все возможные «семерки» вершинных переменных. Такой линейной модели соответствует точная верхняя оценка для $mc(G, w)$ тогда, когда неполный граф G принадлежит семейству графов, нестягиваемых к K_9 (9-клика).

Эту процедуру можно повторять и далее, формируя полные семейства для «нечетного» количества вершин до тех пор, пока мы не исчерпаем все нечетные числа до n . В итоге будет получена квадратичная модель с огромным количеством (равным N_{sc}) функционально избыточных квадратичных ограничений. Этой модели будет соответствовать линейная модель с точно таким же (к сожалению, неполиномиальным от размеров задачи) количеством ограничений. Тем не менее, такая линейная модель дает точную верхнюю линейную оценку для $mc(G, w)$.

Итак, простая идея, связанная с построением функционально избыточных квадратичных ограничений для «нечетного» количества бинарных переменных, обусловила возможность построения «усиленных» аналогов известной линейной модели из [6] для MaxCut-задач.

Заключение.

В работе предложены три новых модели квадратичного типа, с помощью которых можно уточнять верхние оценки в MaxCut-задачах. Все модели тесно связаны с известной линейной моделью [6] для нахождения линейных верхних оценок в MaxCut-задачах. Более того, третья модель дает алгоритм для построения вариантов линейных моделей с более сильными утверждениями о точности линейных оценок в MaxCut-задачах.

Данные модели получены с помощью двух разных способов построения функционально избыточных квадратичных ограничений в моделях квадратичного типа. В связи с этим можно сделать вывод, подтверждающий один из главных тезисов монографии Н.З.Шора [2].

Функционально избыточные квадратичные ограничения играют важную роль при исследовании различного рода многоэкстремальных задач, в том числе и экстремальных задач на графах. Так, предметом исследования в статье была задача о максимальном взвешенном разрезе графа, но подобная схема подойдет и для любой экстремальной задачи на графах, которую можно сформулировать в форме модели квадратичного типа: например, для задач о максимальном взвешенном множестве вершин графа, об оптимальной бисекции графа, о максимальной клике графа и др. Техника функционально избыточных ограничений, очевидно, сможет многое прояснить о проблеме точности верхних оценок для этих задач.

Автор благодарен Науму Зуселевичу Шору за помощь, оказанную при написании работы.

Литература

1. ГЭРИ М., ДЖОНСОН Д. Вычислительные машины и труднорешаемые задачи. – М.: Мир, 1982. – 416 с.
2. SHOR N.Z. Nondifferentiable Optimization and Polynomial Problems. – Dordrecht, Kluwer. – 1998. – 394 p.

3. ШОР Н.З., БЕРЕЗОВСКИЙ О.А. Новые алгоритмы решения задачи о максимальном разрезе неориентированного графа // Кибернетика и систем. анализ. – 1995. – № 2. – С. 100-106.
4. DELORME C., POLJAK S. Laplacian eigenvalues and the maximum cut problem // Math. Programming. – 1993. – 62. – P. 557–574.
5. СТЕЦЮК П.И. О функционально избыточных ограничениях для булевых оптимизационных задач квадратичного типа // Кибернетика и систем. анализ. – 2005. – № 6. – С. 168–172.
6. BARAHONA F., MANJOUR A.R. On the cut polytope // Math. Programming. – 1986. – 36. – P. 157-173.
7. LAURENT M., RENDL F. Semidefinite Programming and Integer Programming. – Report PNA-R0210, CWI, Amsterdam. – 2002. – 91p.
8. GROTSCHEL M., LOVASZ L., SCHRIJVER A. Geometric Algorithms and Combinatorial Optimization. – Springer-Verlag, Berlin. – 1988. – 362 p.

Поступила 23.09.2005

Об уточнении лагранжевых двойственных оценок в бинарных и булевых квадратичных задачах

Стецюк П.И., Пардалос П.М.

Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2006. – С. 145–153.

Рассмотрен подход для улучшения лагранжевых двойственных оценок в квадратичных оптимизационных задачах с бинарными (± 1) и булевыми $(0, 1)$ переменными. Он базируется на использовании семейств функционально избыточных квадратичных ограничений-равенств, которые для этих задач можно построить при введении новых переменных в форме произведений уже существующих переменных. Показано, что введение этих ограничений улучшает точность лагранжевых двойственных оценок.

Многие из NP -трудных экстремальных задач на графах (максимальное устойчивое множество вершин графа, максимальный разрез графа, оптимальная бисекция графа, максим-мальная клика графа и ряд других) формулируются с помощью оптимизационных квадратичных задач на максимум, используя либо бинарные (± 1), либо булевы $(0, 1)$ переменные. Верхнюю оценку для глобального максимума целевой функции в этих задачах позволяет находить разработанная Н.З. Шором техника нахождения оптимальных лагранжевых двойственных оценок [1, 2], которая базируется на решении задачи минимизации выпуклой недифференцируемой функции, определенной на параметрическом семействе неотрицательно определенных симметричных матриц.

Для уточнения лагранжевых двойственных верхних оценок важную роль играют функционально избыточные ограничения [1, 2], которые являются нетривиальными следствиями из уже имеющихся ограничений, но в то же время не являются линейными комбинациями уже существующих ограничений. Такие ограничения не изменяют оптимального значения исходной квадратичной задачи. Но их добавление к исходной квадратичной задаче приводит к новой функции Лагранжа и может оказаться, что новая лагранжева двойственная оценка – более точная верхняя оценка, чем лагранжева двойственная оценка, которая соответствует исходной задаче.

Строить функционально избыточные ограничения можно по-разному. Так, например, в [1, 2, 3] для экстремальных задач на графах (ма-

ксимальное устойчивое множество вершин графа, максимальный разрез графа и др.) исследованы различные семейства функционально избыточных квадратичных ограничений. Большинство из них ориентированы на учет специфических особенностей каждой конкретной задачи для того или иного семейства графов. Оказывается, этого неудобства можно избежать, используя простой прием „расширения“ множества бинарных (булевых) переменных. Он дает возможность генерировать функционально избыточные квадратичные ограничения-равенства, не обращая внимания на специфику задачи, для которой была сформулирована оптимизационная квадратичная модель.

Такой „автомат“ для генерации функционально-избыточных ограничений в бинарных и булевых квадратичных задачах обсудим в данной работе. Сделаем это на примере простейших квадратичных задач, которые связаны с максимизацией квадратичной функции от n бинарных переменных y_1, \dots, y_n , каждая из которых может принимать значение либо $+1$, либо -1 , либо n булевых переменных x_1, \dots, x_n , каждая из которых может принимать значение либо 0 , либо 1 . Заметим, что ничего не изменится, если рассмотреть квадратичную задачу от бинарных (булевых) переменных в общем виде, т.е. при наличии квадратичных ограничений равенств или неравенств.

1. Простейшие бинарная и булева квадратичные задачи. Простейшую бинарную квадратичную задачу рассмотрим в следующей формулировке:

$$Q_y^* = Q(y^*) = \max \left\{ Q(y) = \sum_{i=1}^{n-1} \sum_{j>i}^n q_{ij} y_i y_j + \sum_{i=1}^n q_{i0} y_i + q_{00} \right\} \quad (1)$$

при ограничениях:

$$y_i^2 = 1, \quad i = 1, \dots, n, \quad (2)$$

а простейшую булеву квадратичную задачу – в формулировке:

$$Q_x^* = Q(x^*) = \max \left\{ Q(x) = \sum_{i=1}^{n-1} \sum_{j>i}^n q_{ij} x_i x_j + \sum_{i=1}^n q_{i0} x_i + q_{00} \right\} \quad (1')$$

при ограничениях:

$$x_i^2 - x_i = 0, \quad i = 1, \dots, n. \quad (2')$$

Пусть Ψ_y^* и Ψ_x^* – наилучшие лагранжевые двойственные оценки, соответствующие задачам (1)–(2) и (1')–(2'). Их можно найти с помощью решения задачи минимизации выпуклой недифференцируемой

функции, определенной на параметрическом семействе неотрицательно определенных симметричных матриц [2]. Здесь не будем обсуждать алгоритмы нахождения оценок Ψ_y^* и Ψ_x^* , а отметим лишь, что эти оценки служат оценками сверху для глобальных максимумов Q_y^* и Q_x^* , т.е. $\Psi_y^* \geq Q_y^*$, $\Psi_x^* \geq Q_x^*$.

Оценки Ψ_y^* и Ψ_x^* не всегда есть достаточно хорошими верхними оценками для Q_y^* и Q_x^* . Так, например, для функций от двух переменных имеем

$$\begin{aligned} Q(y) &= -y_1 y_2 - y_1 - y_2 & \implies & Q_y^* = 1, \quad \Psi_y^* = 1.5, \\ Q(y) &= -y_1 y_2 - 2y_1 - y_2 & \implies & Q_y^* = 2, \quad \Psi_y^* = 2.25, \\ Q(x) &= -2x_1 x_2 + 3x_1 + 2x_2 & \implies & Q_x^* = 3, \quad \Psi_x^* = 3.125, \\ Q(x) &= -x_1 x_2 + x_1 + x_2, & \implies & Q_x^* = 1, \quad \Psi_x^* = 1.125, \end{aligned}$$

и следовательно, для всех указанных примеров они есть неточными. Более того, в первом примере имеем наихудшую относительную точность оценки Ψ_y^* , которая составляет 50%.

Поэтому оценки Ψ_y^* и Ψ_x^* нуждаются в уточнении. Сделать это можно за счет использования функционально избыточных ограничений. Далее рассмотрим получение новых более точных верхних оценок, чем оценки Ψ_y^* и Ψ_x^* . Новые оценки получены благодаря использованию функционально избыточных ограничений, которые можно построить, если для задач (1)–(2) и (1')–(2') определенным образом расширить множество бинарных и булевых переменных.

2. Расширенные бинарная и булева квадратичные задачи.

Для задачи (1)–(2) множество бинарных переменных расширим введением новых бинарных переменных $y_{ij} = y_i y_j$ для всех $i, j : 1 \leq i < j \leq n$. В результате получим „расширенную“ формулировку квадратичной задачи (1)–(2):

$$Q_y^* = Q(y^*) = \max \left\{ Q(y) = \sum_{i=1}^{n-1} \sum_{j>i}^n q_{ij} y_i y_j + \sum_{i=1}^n q_{i0} y_i + q_{00} \right\} \quad (3)$$

при ограничениях:

$$y_i^2 = 1, \quad i = 1, \dots, n. \quad (4)$$

$$\begin{cases} y_{ij} = y_i y_j, & \forall i, j : 1 \leq i < j \leq n. \\ y_{ij}^2 = 1, \end{cases} \quad (5)$$

Множество булевых переменных для задачи (1')–(2') расширим за счет введения новых булевых переменных $x_{ij} = x_i x_j$ для всех $i, j : 1 \leq i < j \leq n$.

В результате получим „расширенную“ формулировку булевой квадратичной задачи (1')–(2'):

$$Q_x^* = Q(x^*) = \max \left\{ Q(x) = \sum_{i=1}^{n-1} \sum_{j>i}^n q_{ij} x_i x_j + \sum_{i=1}^n q_{i0} x_i + q_{00} \right\} \quad (3')$$

при ограничениях:

$$x_i^2 - x_i = 0, \quad i = 1, \dots, n. \quad (4')$$

$$\begin{cases} x_{ij} = x_i x_j, \\ x_{ij}^2 - x_{ij} = 0, \end{cases} \quad \forall i, j: 1 \leq i < j \leq n. \quad (5')$$

Справедлива следующая теорема.

Теорема 1. *Наилучшие лагранжевые двойственные оценки для задач (3)–(5) и (3')–(5') удовлетворяют условиям:*

- (i) для задачи (3)–(5) оценка не больше, чем Ψ_y^* ,
- (ii) для задачи (3')–(5') оценка не больше, чем Ψ_x^* .

Однако ожидать, что расширение множества переменных значительно улучшит точность лагранжевых двойственных оценок не приходится. Зато „расширенные“ формулировки квадратичных задач позволяют описывать семейства функционально избыточных ограничений в форме квадратичных ограничений-равенств [4], добавление которых к задачам (3)–(5) и (3')–(5') может способствовать уточнению лагранжевых двойственных оценок. Содержательный смысл этих семейств ограничений состоит в отражении различных видов „избыточных“ квадратичных зависимостей между новыми переменными y_{ij} и старыми переменными y_i для бинарной квадратичной задачи, и между новыми переменными x_{ij} и старыми переменными x_i для булевой квадратичной задачи.

3. Семейства функционально избыточных ограничений для расширенных бинарной и булевой задач. Для бинарных переменных имеем следующие семейства функционально избыточных ограничений.

Для каждой двойки бинарных переменных y_i, y_j имеем два ограничения

$$y_i - y_{ij} y_j = 0, \quad y_j - y_{ij} y_i = 0, \quad (6)$$

которые являются следствием очевидных равенств

$$y_i = y_i y_j^2 = (y_i y_j) y_j = y_{ij} y_j,$$

$$y_j = y_j y_i^2 = (y_i y_j) y_i = y_{ij} y_i$$

и справедливы при любых i и j , таких что $1 \leq i < j \leq n$.

Для каждой тройки бинарных переменных y_i , y_j и y_k имеем пять функционально избыточных ограничений. Первые три из них такие:

$$y_{ij} - y_{ik} y_{jk} = 0, \quad y_{ik} - y_{ij} y_{jk} = 0, \quad y_{jk} - y_{ij} y_{ik} = 0 \quad (7)$$

и являются следствием очевидных равенств

$$y_{ij} = y_i y_j = y_i y_j y_k^2 = (y_i y_k)(y_j y_k) = y_{ik} y_{jk},$$

$$y_{ik} = y_i y_k = y_i y_k y_j^2 = (y_i y_j)(y_j y_k) = y_{ij} y_{jk},$$

$$y_{jk} = y_j y_k = y_j y_k y_i^2 = (y_i y_j)(y_i y_k) = y_{ij} y_{ik},$$

которые, учитывая (2), справедливы при любых индексах i , j и k . Три ограничения (7) дополняются двумя ограничениями

$$y_{ij} y_k - y_{ik} y_j = 0, \quad y_{ij} y_k - y_{jk} y_i = 0, \quad (8)$$

которые следуют из неоднозначности представления произведения $y_i y_j y_k$, т.е.

$$y_i y_j y_k = (y_i y_j) y_k = y_{ij} y_k = (y_i y_k) y_j = y_{ik} y_j = (y_j y_k) y_i = y_{jk} y_i.$$

Заметим, что из неоднозначного представления $y_i y_j y_k$ следует три ограничения. Но одно из них – линейно зависимо от двух оставшихся ограничений, и не будет влиять на точность лагранжевой двойственной оценки.

Для каждой четверки бинарных переменных y_i , y_j , y_k и y_l имеем два ограничения:

$$y_{ij} y_{kl} - y_{ik} y_{jl} = 0, \quad y_{ij} y_{kl} - y_{il} y_{jk} = 0, \quad (9)$$

которые линейно независимы и следуют из неоднозначности представления произведения всех четырех переменных

$$y_i y_j y_k y_l = y_{ij} y_{kl} = y_{ik} y_{jl} = y_{il} y_{jk}.$$

По аналогичной схеме для булевых переменных имеем следующие семейства функционально избыточных ограничений.

Для каждой двойки булевых переменных x_i, x_j имеем два ограничения

$$x_{ij}x_i - x_{ij} = 0, \quad x_{ij}x_j - x_{ij} = 0, \quad (6')$$

которые являются следствием равенств:

$$0 = x_i^2 - x_i = (x_i^2 - x_i)x_j = x_i(x_ix_j) - x_ix_j = x_ix_{ij} - x_{ij} = 0,$$

$$0 = x_j^2 - x_j = (x_j^2 - x_j)x_i = x_j(x_jx_i) - x_jx_i = x_jx_{ij} - x_{ij} = 0,$$

и справедливы при любых i и j , таких что $1 \leq i < j \leq n$.

Для каждой тройки булевых переменных x_i, x_j и x_k имеем пять ограничений. Первые три ограничения

$$x_{ij}x_{ik} - x_{ij}x_{jk} = 0, \quad x_{ij}x_{jk} - x_{jx_{ik}} = 0, \quad x_{ik}x_{jk} - x_kx_{ij} = 0 \quad (7')$$

являются следствием таких очевидных равенств

$$0 = (x_i^2 - x_i)x_jx_k = (x_ix_j)(x_ix_k) - x_i(x_jx_k) = x_{ij}x_{ik} - x_ix_{jk} = 0,$$

$$0 = (x_j^2 - x_j)x_ix_k = (x_ix_j)(x_jx_k) - x_j(x_ix_k) = x_{ij}x_{jk} - x_jx_{ik} = 0,$$

$$0 = (x_k^2 - x_k)x_ix_j = (x_ix_k)(x_jx_k) - x_k(x_ix_j) = x_{ik}x_{jk} - x_kx_{ij} = 0,$$

которые, учитывая (2'), справедливы при любых индексах i, j и k . Они точно также, как и в случае бинарных переменных, дополняются еще двумя такими ограничениями

$$x_{ij}x_k - x_{ik}x_j = 0, \quad x_{ij}x_k - x_{jk}x_i = 0, \quad (8')$$

которые следуют из неоднозначного представления

$$x_ix_jx_k = (x_ix_j)x_k = x_{ij}x_k = (x_ix_k)x_j = x_{ik}x_j = (x_jx_k)x_i = x_{jk}x_i.$$

Для каждой четверки булевых переменных x_i, x_j, x_k и x_l имеем два ограничения

$$x_{ij}x_{kl} - x_{ik}x_{jl} = 0, \quad x_{ij}x_{kl} - x_{il}x_{jk} = 0, \quad (9')$$

которые линейно независимы и следуют из неоднозначности представления произведения четырех переменных

$$x_ix_jx_kx_l = x_{ij}x_{kl} = x_{ik}x_{jl} = x_{il}x_{jk}.$$

Семейства функционально избыточных ограничений (6)–(9) и (6')–(9') дают возможность построить новые улучшенные бинарные и булевы квадратичные задачи, которым в ряде случаев будут соответствовать более точные оценки, чем оценки Ψ_y^* и Ψ_x^* .

4. Улучшенные бинарная и булева квадратичные задачи.

Если квадратичную задачу (3)–(5) дополнить группами функционально избыточных ограничений (6), (7), (8) и (9), то получим новую квадратичную задачу в бинарных переменных:

$$Q_y^* = \max \left\{ Q(y) = \sum_{i=1}^{n-1} \sum_{j>i}^n q_{ij} y_i y_j + \sum_{i=1}^n q_{i0} y_i + q_{00} \right\} \quad (10)$$

при ограничениях

$$y_i^2 - 1 = 0, \quad i = 1, \dots, n, \quad (11)$$

$$\begin{cases} y_{ij}^2 - 1 = 0, \\ y_i y_j - y_{ij} = 0, \\ y_{ij} y_i - y_j = 0, \\ y_{ij} y_j - y_i = 0, \end{cases} \quad \forall i, j: 1 \leq i < j \leq n, \quad (12)$$

$$\begin{cases} y_{ij} y_k - y_{ik} y_j = 0, \\ y_{ij} y_k - y_{jk} y_i = 0, \\ y_{ik} y_j - y_{ij} = 0, \\ y_{ij} y_j - y_{ik} = 0, \\ y_{ij} y_{jk} - y_{ik} = 0, \end{cases} \quad \forall i, j, k: 1 \leq i < j < k \leq n, \quad (13)$$

$$\begin{cases} y_{ij} y_{kl} - y_{ik} y_{jl} = 0, \\ y_{ij} y_{kl} - y_{il} y_{jk} = 0, \end{cases} \quad \forall i, j, k, l: 1 \leq i < j < k < l \leq n. \quad (14)$$

Пусть ψ_y^* – наилучшая лагранжева двойственная оценка для задачи (10)–(14). Она всегда будет не хуже, чем оценка Ψ_y^* , а в ряде случаев будет более точной верхней оценкой для Q_y^* .

Аналогично, если квадратичную задачу (3')–(5') дополнить группами функционально избыточных ограничений (6'), (7'), (8') и (9'), то получим новую квадратичную задачу в булевых переменных:

$$Q_x^* = \max \left\{ Q(x) = \sum_{i=1}^{n-1} \sum_{j>i}^n q_{ij} x_i x_j + \sum_{i=1}^n q_{i0} x_i + q_{00} \right\} \quad (10')$$

при ограничениях

$$x_i^2 - x_i = 0, \quad i = 1, \dots, n, \quad (11')$$

$$\begin{cases} x_{ij}^2 - x_{ij} = 0, \\ x_i x_j - x_{ij} = 0, \\ x_{ij} x_i - x_{ij} = 0, \\ x_{ij} x_j - x_{ij} = 0, \end{cases} \quad \forall i, j: 1 \leq i < j \leq n, \quad (12')$$

$$\begin{cases} x_{ij} x_k - x_{ik} y_j = 0, \\ x_{ij} x_k - x_{jk} y_i = 0, \\ x_{ik} x_{jk} - x_{ij} x_k = 0, \\ x_{ij} x_{jk} - x_{ik} x_j = 0, \\ x_{ij} x_{ik} - x_{jk} x_i = 0, \end{cases} \quad \forall i, j, k: 1 \leq i < j < k \leq n, \quad (13')$$

$$\begin{cases} x_{ij} x_{kl} - x_{ik} x_{jl} = 0, \\ x_{ij} x_{kl} - x_{il} x_{jk} = 0, \end{cases} \quad \forall i, j, k, l: 1 \leq i < j < k < l \leq n. \quad (14')$$

Пусть ψ_x^* – наилучшая лагранжевая двойственная оценка для задачи (10')–(14'). Она всегда будет не хуже, чем оценка Ψ_x^* , а в ряде случаев будет более точной верхней оценкой для Q_x^* .

Итак, новые квадратичные задачи (10)–(14) и (10')–(14') получены дополнением простейших квадратичных задач (1)–(2) и (1')–(2') полными семействами функционально избыточных ограничений, которые рассмотрены выше. Ограничения (12) и (12') описывают связи внутри каждой двойки булевых и бинарных переменных. Их количество равно $2n(n-1)$. Ограничения (13) и (13') описывают дополнительные связи внутри каждой тройки булевых и бинарных переменных и их количество равно $5n(n-1)(n-2)/6$. Количество ограничений (14) и (14') равно $n(n-1)(n-2)(n-3)/12$ и они описывают дополнительные связи для переменных внутри каждой четверки булевых и бинарных переменных.

Справедлива следующая теорема.

Теорема 2. *Оценки ψ_y^* и ψ_x^* (при произвольном $n \geq 4$) удовлетворяют следующим неравенствам*

$$Q_y^* \leq \psi_y^* \leq \Psi_y^*, \quad Q_x^* \leq \psi_x^* \leq \Psi_x^*.$$

В заключение отметим, что точность оценок ψ_y^* и ψ_x^* требует детального исследования. Так, неравенства в теореме 2 – очень грубые и рассчитаны на наихудшие случаи бинарной и булевой квадратичных задач. Конечно в ряде случаев оценки ψ_y^* и ψ_x^* будут более точными, чем оценки Ψ_y^* и Ψ_x^* . Так, например, для бинарной квадратичной задачи с $n = 4$, где $Q(y) = -y_1 y_2 - y_1 - y_2 - y_3 y_4 - y_3 - y_4$, имеем $Q_y^* = 2$, $\Psi_y^* = 3$, $\psi_y^* = 2$. Для булевой квадратичной задачи с $n = 4$, где $Q(x) = -x_1 x_2 + x_1 + x_2 - x_3 x_4 + x_3 + x_4$, имеем $Q_x^* = 2$, $\Psi_x^* = 2.25$, $\psi_x^* = 2$. В обоих случаях оценки ψ_y^* и ψ_x^* являются точными.

Литература

1. SHOR N.Z. Nondifferentiable Optimization and Polynomial Problems. – Dordrecht: Kluwer, 1998. – 394 p.
2. ШОР Н.З. Роль избыточных ограничений в улучшении двойственных оценок для полиномиальных оптимизационных задач // Кибернетика и системный анализ. – 1998. – № 4. – С. 106–121.
3. SHOR N.Z., STETSYUK P.I. Lagrangian bounds in multiextremal polynomial and discrete optimization problems // J. of Global Optimization. – 2002 – 23. – P. 1–41.
4. СТЕЦЮК П.И. О функционально избыточных ограничениях для булевых оптимизационных задач квадратичного типа // Кибернетика и системный анализ. – 2005. – №6. – С. 168–172.

Получено 17.07.2006

Об уточнении верхних оценок Н.З. Шора для числа независимости графа

П.И. Стецюк

*Информационный бюллетень Ассоциации математического
программирования. № 11. Научное издание.
– Екатеринбург: УрО РАН, 2007. – С. 80–81.*

Пусть $G=(V, E)$ – неориентированный граф с множеством вершин V и множеством ребер E . Подмножество вершин $S \subseteq V$ называется независимым множеством графа G , если для любых двух вершин $i, j \in S$ выполняется $ij \notin E$. Максимальный размер независимого множества называют числом независимости графа G и обозначают $\alpha(G)$. Задача нахождения $\alpha(G)$ – *NP*-трудная.

В [1] Н.З. Шор предложил три оптимальных лагранжевых двойственных оценки $\psi_1^*(G)$, $\psi_2^*(G)$ и $\psi_3^*(G)$, каждая из которых есть верхней оценкой (оценкой сверху) для $\alpha(G)$ и может быть найдена с помощью методов недифференцируемой оптимизации. Оценка $\psi_1^*(G)$ связана с квадратичной булевой задачей

$$\alpha(G) = \max \sum_{i \in V} x_i : x_i x_j = 0 \ (\forall ij \in E), \ x_i^2 - x_i = 0 \ (\forall i \in V), \quad (1)$$

с помощью которой формулируется проблема нахождения $\alpha(G)$. Оценки $\psi_2^*(G)$ и $\psi_3^*(G)$ связаны с квадратичными булевыми задачами, полученными последовательным добавлением к задаче (1) двух семейств функционально избыточных квадратичных ограничений $x_i x_j \geq 0 \ (\forall ij \notin E)$ и $x_i x_k + x_j x_k \leq x_k \ (\forall ij \in E, k \neq i, j)$, которые следуют из квадратичных ограничений задачи (1).

В [1] показано, что $\psi_1^*(G) = \theta(G)$, а $\psi_2^*(G) = \theta'(G)$, где $\theta(G)$ и $\theta'(G)$ – числа Ловаса для графа G . Для произвольного графа G справедливо следующее соотношение

$$\psi_1^*(G) = \theta(G) \geq \psi_2^*(G) = \theta'(G) \geq \psi_3^*(G) \geq \alpha(G). \quad (2)$$

Оценка $\psi_3^*(G)$ есть уточнением $\theta'(G)$, которое ориентировано на специальные классы графов. Так, например, в [1] показано, что $\psi_3^*(G)$ есть точной верхней оценкой для $\alpha(G)$, когда граф G есть t -перфектным, либо h -перфектным.

Новые оптимальные лагранжевые двойственные оценки $\psi_a^*(G)$, $\psi_b^*(G)$ и $\psi_c^*(G)$ построим аналогично тому, как построены оценки Н.З. Шора. В качестве базовой возьмем задачу (1) и будем последовательно дополнять ее такими семействами функционально избыточных ограничений из [2] для бинарных (± 1) переменных

$$y_i y_j = y_{ij}, y_{ij}^2 = 1, y_{ij} y_i = y_j, y_{ij} y_j = y_i \quad (\forall i, j \in V), \quad (3)$$

$$\begin{cases} y_{ij} y_k = y_{ik} y_j, y_{ij} y_k = y_{jk} y_i, \\ y_{ik} y_{jk} = y_{ij}, y_{ij} y_{jk} = y_{ik}, y_{ij} y_{ik} = y_{jk} \end{cases} \quad (\forall i, j, k \in V), \quad (4)$$

$$y_{ij} y_{kl} = y_{ik} y_{jl}, y_{ij} y_{kl} = y_{il} y_{jk} \quad (\forall i, j, k, l \in V), \quad (5)$$

переводя бинарные переменные y_i и y_{ij} в булевы переменные x_i и x_{ij} с помощью замены $y_i = 1 - 2x_i$ и $y_{ij} = 1 - 2x_{ij}$.

Теорема 1. *Для произвольного графа G справедливы соотношения*

$$\psi_2^*(G) = \theta'(G) \geq \psi_a^*(G) \geq \psi_b^*(G) \geq \psi_c^*(G) \geq \alpha(G), \quad (6)$$

$$\psi_3^*(G) \geq \psi_b^*(G) \geq \psi_c^*(G) \geq \alpha(G). \quad (7)$$

Из (7) следует, что если граф G — t -перфектный, либо h -перфектный, то $\psi_3^*(G) = \psi_b^*(G) = \psi_c^*(G) = \alpha(G)$. Следовательно, верхние оценки $\psi_b^*(G)$ и $\psi_c^*(G)$ есть точными для t - и h -перфектных графов. Из (6) следует, что для произвольного графа G все три новые оценки есть не менее точными, чем оценка $\psi_2^*(G)$, а значит, и число Ловаса $\theta'(G)$. Оценки $\psi_b^*(G)$ и $\psi_c^*(G)$ часто оказываются более точными, чем $\psi_3^*(G)$. В качестве примеров приведем граф в форме икосаэдра I_{12} (12 вершин и 30 ребер) и дополнительный к нему граф \bar{I}_{12} (12 вершин и 36 ребер), для которых $\theta'(I_{12}) = \theta'(\bar{I}_{12}) = 1 + \sqrt{5} \approx 3.236068$ и оценка $\psi_3^*(G)$ не улучшает ни $\theta'(I_{12})$, ни $\theta'(\bar{I}_{12})$. Для графов I_{12} и \bar{I}_{12} имеем $\psi_b^*(I_{12}) = \psi_b^*(\bar{I}_{12}) \approx 3.185116$ и верхняя оценка $\psi_b^*(G)$ есть более точной, чем $\psi_3^*(G)$. Оценка $\psi_c^*(G)$ оказывается точной верхней оценкой для $\alpha(G)$ в обоих графах, т.е. $\psi_c^*(\bar{I}_{12}) = \alpha(\bar{I}_{12}) = 3$, $\psi_c^*(I_{12}) = \alpha(I_{12}) = 3$.

Литература

1. SHOR N.Z. Nondifferentiable Optimization and Polynomial Problems. — Dordrecht, Kluwer. — 1998. — 394 p.
2. СТЕЦЮК П.И. О функционально избыточных ограничениях для булевых оптимизационных задач квадратичного типа // Кибернетика и системный анализ. 2005, № 6, С. 168–172.

Бинарный квадратичный многогранник и его аппроксимации

П. И. Стецюк, Н. Ю. Золотых

Журнал вычислительной та прикладной математики. — 2010. — № 2(101). — С. 139–149.

В работе введен бинарный квадратичный многогранник $BiQP_n$ для описания множества допустимых решений в задаче нахождения экстремума квадратичной функции от n бинарных ± 1 -переменных. Построено семейство многогранников, заданных с помощью семейств линейных неравенств, и показано, что они аппроксимируют сверху многогранник $BiQP_n$.

ВВЕДЕНИЕ

В работе рассматриваются вопросы, связанные с задачами минимизации (максимизации) квадратичных функций от булевых $(0 - 1)$ или бинарных $(-1, 1)$ переменных. В общем случае эти задачи являются NP -трудными. Так, например, к ним сводятся такие NP -трудные задачи из теории графов, как задача о максимальном независимом множестве вершин графа и задача об максимальном взвешенном разрезе графа [1].

Для описания множества допустимых решений в задаче нахождения экстремума квадратичной функции от n бинарных переменных введен бинарный квадратичный многогранник $BiQP_n$ (раздел 5) и сделано это по аналогии с булевым квадратичным многогранником Падберга [2]. Для аппроксимации сверху многогранника $BiQP_n$ построено семейство многогранников с конечным количеством линейных неравенств (раздел 6). В основу положено свойство нечетного количества ± 1 (раздел 1) и вытекающие из него квадратичные неравенства для четного и нечетного количества бинарных переменных (разделы 2, 3). В разделе 4 описано применение этих неравенств в бинарных и булевых квадратичных задачах, для которых они и были предложены в [3, 4].

1. Нечетное количество ± 1

Пусть последовательность $l(k) = (l_1, \dots, l_k)$ состоит из $k \geq 2$ бинарных (± 1) чисел. Условимся ее называть бинарной последовательностью и будем использовать для нее еще и такое обозначение $l \in \{-1, 1\}^k$. Будем различать нечетной длины бинарную последовательность $l(2r + 1) = (l_1, \dots, l_{2r+1})$ и четной длины бинарную последовательность $l(2r) = (l_1, \dots, l_{2r})$. Здесь $r = 1, 2, 3, \dots$ — натуральные числа.

Лемма 1. Для произвольной нечетной длины бинарной последовательности $l(2r+1) = (l_1, \dots, l_{2r+1})$ всегда справедливо квадратичное неравенство:

$$S_{2r+1}^2 = \left(\sum_{i=1}^{2r+1} l_i \right)^2 = (l_1 + l_2 + \dots + l_{2r+1})^2 \geq 1. \quad (1)$$

Для заданного r количество всех возможных квадратичных неравенств (1) равно 4^r .

Доказательство. Легко видеть, что для нечетной длины бинарной последовательности $l(2r+1)$ сумма всех ее членов $S_{2r+1} = l_1 + l_2 + \dots + l_{2r+1}$ не может быть равна нулю, так как S_{2r+1} получается в результате суммирования нечетного количества ± 1 . Самые близкие к нулю целые значения суммы S_{2r+1} такие: $S_{2r+1} = 1$, когда ровно $(r+1)$ произвольных членов последовательности $l(2r+1)$ равны $+1$, а оставшиеся r ее членов равны -1 ; $S_{2r+1} = -1$, когда ровно r членов последовательности $l(2r+1)$ равны $+1$ и оставшиеся $(r+1)$ ее членов равны -1 .

Для бинарной последовательности $l(2r+1) = (l_1, \dots, l_{2r+1})$ всегда справедливо следующее квадратичное неравенство:

$$S_{2r+1}^2 = (l_1 + l_2 + \dots + l_{2r+1})^2 \geq (\pm 1)^2 = 1,$$

что доказывает неравенство (1).

Количество последовательностей $l(2r+1) = (l_1, \dots, l_{2r+1})$ для заданного r равно $2^{2r+1} = 2 \cdot 4^r$. Однако, последовательность $l(2r+1) = (l_1, \dots, l_{2r+1})$ и последовательность $l'(2r+1) = (-l_1, \dots, -l_{2r+1})$ дают одно и то же неравенство (1). Следовательно, для заданного r количество всех возможных неравенств вида (1) равно 4^r . \square

2. КВАДРАТИЧНЫЕ НЕРАВЕНСТВА ДЛЯ БИНАРНЫХ ПЕРЕМЕННЫХ

Пусть имеется вектор-строка $y = (y_1, \dots, y_k)$, которая состоит из k бинарных (± 1) переменных. Будем различать векторы $y \in \{-1, 1\}^{2r+1}$ с нечетным количеством бинарных переменных и векторы $y \in \{-1, 1\}^{2r}$ с четным количеством бинарных переменных, где $r = 1, 2, \dots$ — натуральные числа.

Лемма 2. Для нечетного количества переменных $y \in \{-1, 1\}^{2r+1}$, $r = 1, 2, \dots$ и произвольной нечетной длины бинарной последовательности

$l(2r+1) = (l_1, \dots, l_{2r+1})$ справедливо квадратичное неравенство

$$\left(\sum_{i=1}^{2r+1} l_i y_i \right)^2 \geq 1 \quad \text{или его аналог} \quad \sum_{i=1}^{2r} \sum_{j>i}^{2r+1} l_i l_j y_i y_j \geq -r. \quad (2)$$

Для четного количества бинарных переменных $y \in \{-1, 1\}^{2r}$, $r = 1, 2, \dots$, $l_0 = \pm 1$ и произвольной четной длины бинарной последовательности $l(2r) = (l_1, \dots, l_{2r})$ справедливо квадратичное неравенство

$$\left(l_0 + \sum_{i=1}^{2r} l_i y_i \right)^2 \geq 1 \quad \text{или его аналог} \quad \sum_{i=1}^{2r} l_0 l_i y_i + \sum_{i=1}^{2r-1} \sum_{j>i}^{2r} l_i l_j y_i y_j \geq -r. \quad (3)$$

Для заданного r количество всех возможных квадратичных неравенств как вида (2), так и вида (3), равно 4^r .

Доказательство. Первое неравенство в (2) следует из леммы 1, если ее применить к нечетной длины бинарной последовательности $l'(2r+1) = (l'_1, \dots, l'_{2r+1})$, члены которой получены по правилу: $l'_i = l_i y_i$, $i = 1, \dots, 2r+1$. Учитывая, что $l_i^2 = 1$ и $y_i^2 = 1$ для всех $i = 1, \dots, 2r+1$, правую часть первого неравенства из (2) можно записать так:

$$\left(\sum_{i=1}^{2r+1} l_i y_i \right)^2 = \sum_{i=1}^{2r+1} l_i^2 y_i^2 + 2 \sum_{i=1}^{2r} \sum_{j>i}^{2r+1} l_i l_j y_i y_j = 2r+1 + 2 \sum_{i=1}^{2r} \sum_{j>i}^{2r+1} l_i l_j y_i y_j.$$

Тогда первое неравенство в (2) можно переписать следующим образом

$$2r+1 + 2 \sum_{i=1}^{2r} \sum_{j>i}^{2r+1} l_i l_j y_i y_j \geq 1 \quad \text{или} \quad 2 \sum_{i=1}^{2r} \sum_{j>i}^{2r+1} l_i l_j y_i y_j \geq -2r,$$

откуда, после деления на 2, получаем второе квадратичное неравенство в (2).

По аналогичной схеме доказывается и (3). Так, первое неравенство в (3) следует из леммы 1, если ее применить к нечетной длины бинарной последовательности $l'(2r+1) = (l'_1, \dots, l'_{2r+1})$, члены которой получены по правилу: $l'_1 = l_0$ и $l'_{i+1} = l_i y_i$, $i = 1, \dots, 2r$. Учитывая, что $l_0^2 = 1$, $l_i^2 = 1$ и $y_i^2 = 1$ для всех $i = 1, \dots, 2r$, правую часть первого неравенства из (3) можно записать так:

$$\begin{aligned} \left(l_0 + \sum_{i=1}^{2r} l_i y_i \right)^2 &= l_0^2 + \sum_{i=1}^{2r} l_i^2 y_i^2 + 2 \sum_{i=1}^{2r} l_0 l_i y_i + 2 \sum_{i=1}^{2r-1} \sum_{j>i}^{2r} l_i l_j y_i y_j \\ &= 2r+1 + 2 \sum_{i=1}^{2r} l_0 l_i y_i + 2 \sum_{i=1}^{2r-1} \sum_{j>i}^{2r} l_i l_j y_i y_j. \end{aligned}$$

Тогда первое неравенство в (3) можно переписать следующим образом

$$2r + 1 + 2 \sum_{i=1}^{2r} l_0 l_i y_i + 2 \sum_{i=1}^{2r-1} \sum_{j>i}^{2r} l_i l_j y_i y_j \geq 1,$$

$$2 \sum_{i=1}^{2r} l_0 l_i y_i + 2 \sum_{i=1}^{2r-1} \sum_{j>i}^{2r} l_i l_j y_i y_j \geq -2r,$$

откуда, после деления на 2, получаем второе квадратичное неравенство из (3).

Для одного и того же r ($r = 1, 2, 3, \dots$) квадратичным неравенствам (2) и (3) соответствуют пусть и различные, но одинаковой длины нечетные последовательности $l'(2r+1) = (l'_1, \dots, l'_{2r+1})$. Поэтому, для заданного r количество всевозможных квадратичных неравенств как вида (2), так и вида (3) равно 4^r (следует из леммы 1). \square

3. ПРИМЕРЫ ПОЛНЫХ СЕМЕЙСТВ КВАДРАТИЧНЫХ НЕРАВЕНСТВ

Приведем полные семейства квадратичных неравенств, которые следуют из леммы 2, для двоек и троек бинарных переменных (им соответствует $r = 1$), четверок и пятерок бинарных переменных (им соответствует $r = 2$). Для удобства будем рассматривать бинарные переменные y_i, y_j, y_k, y_l, y_m , т.е. каждой бинарной переменной от первой до пятой припишем свой отдельный индекс.

Полные семейства квадратичных неравенств для двоек и троек бинарных переменных содержат по 4 различных неравенства. Для двоек бинарных переменных полное семейство квадратичных неравенств следует из (3) при $r = 1$ и включает неравенства

$$\left\{ \begin{array}{l} (+1 + y_i + y_j)^2 \geq 1, \\ (-1 + y_i + y_j)^2 \geq 1, \\ (+1 - y_i + y_j)^2 \geq 1, \\ (+1 + y_i - y_j)^2 \geq 1, \end{array} \right. \quad \text{или их аналоги} \quad \left\{ \begin{array}{l} +y_i + y_j + y_i y_j \geq -1, \\ -y_i - y_j + y_i y_j \geq -1, \\ -y_i + y_j - y_i y_j \geq -1, \\ +y_i - y_j - y_i y_j \geq -1. \end{array} \right. \quad (4)$$

Для троек бинарных переменных полное семейство квадратичных неравенств следует из (2) при $r = 1$ и включает неравенства

$$\left\{ \begin{array}{l} (+y_i + y_j + y_k)^2 \geq 1, \\ (-y_i + y_j + y_k)^2 \geq 1, \\ (+y_i - y_j + y_k)^2 \geq 1, \\ (+y_i + y_j - y_k)^2 \geq 1, \end{array} \right. \quad \text{или их аналоги} \quad \left\{ \begin{array}{l} +y_i y_j + y_i y_k + y_j y_k \geq -1, \\ -y_i y_j - y_i y_k + y_j y_k \geq -1, \\ -y_i y_j + y_i y_k - y_j y_k \geq -1, \\ +y_i y_j - y_i y_k - y_j y_k \geq -1. \end{array} \right. \quad (5)$$

Полные семейства квадратичных неравенств для четверок и пятерок бинарных переменных содержат по $16 = 4^2$ различных неравенств. Для четверок они следуют из (3) при $r = 2$, а для пятерок – из (2) при $r = 2$. Полные семейства квадратичных неравенств для четверок и пятерок бинарных переменных в форме первых неравенств из (3) и (2) включают следующие 16 неравенств:

$$\left\{ \begin{array}{l} (+1 + y_i + y_j + y_k + y_l)^2 \geq 1, \\ (-1 + y_i + y_j + y_k + y_l)^2 \geq 1, \\ (+1 - y_i + y_j + y_k + y_l)^2 \geq 1, \\ (+1 + y_i - y_j + y_k + y_l)^2 \geq 1, \\ (+1 + y_i + y_j - y_k + y_l)^2 \geq 1, \\ (+1 + y_i + y_j + y_k - y_l)^2 \geq 1, \\ (-1 - y_i + y_j + y_k + y_l)^2 \geq 1, \\ (-1 + y_i - y_j + y_k + y_l)^2 \geq 1, \\ (-1 + y_i + y_j - y_k + y_l)^2 \geq 1, \\ (-1 + y_i + y_j + y_k - y_l)^2 \geq 1, \\ (+1 - y_i - y_j + y_k + y_l)^2 \geq 1, \\ (+1 - y_i + y_j - y_k + y_l)^2 \geq 1, \\ (+1 - y_i + y_j + y_k - y_l)^2 \geq 1, \\ (+1 + y_i - y_j - y_k + y_l)^2 \geq 1, \\ (+1 + y_i - y_j + y_k - y_l)^2 \geq 1, \\ (+1 + y_i + y_j - y_k - y_l)^2 \geq 1, \\ (+1 + y_i + y_j + y_k - y_l)^2 \geq 1, \end{array} \right. \text{ и } \left\{ \begin{array}{l} (+y_i + y_j + y_k + y_l + y_m)^2 \geq 1, \\ (-y_i + y_j + y_k + y_l + y_m)^2 \geq 1, \\ (+y_i - y_j + y_k + y_l + y_m)^2 \geq 1, \\ (+y_i + y_j - y_k + y_l + y_m)^2 \geq 1, \\ (+y_i + y_j + y_k - y_l + y_m)^2 \geq 1, \\ (+y_i + y_j + y_k + y_l - y_m)^2 \geq 1, \\ (-y_i - y_j + y_k + y_l + y_m)^2 \geq 1, \\ (-y_i + y_j - y_k + y_l + y_m)^2 \geq 1, \\ (-y_i + y_j + y_k - y_l + y_m)^2 \geq 1, \\ (-y_i + y_j + y_k + y_l - y_m)^2 \geq 1, \\ (+y_i - y_j - y_k + y_l + y_m)^2 \geq 1, \\ (+y_i - y_j + y_k - y_l + y_m)^2 \geq 1, \\ (+y_i - y_j + y_k + y_l - y_m)^2 \geq 1, \\ (+y_i + y_j - y_k - y_l + y_m)^2 \geq 1, \\ (+y_i + y_j - y_k + y_l - y_m)^2 \geq 1, \\ (+y_i + y_j + y_k - y_l - y_m)^2 \geq 1. \end{array} \right. \quad (6)$$

Аналоги квадратичных неравенств (6) приводить не будем в силу их громоздкости. Отметим лишь, что по мере увеличения r ($r = 3, 4, \dots$) полные семейства квадратичных неравенств приобретают очень громоздкую форму даже при их описании с помощью первых квадратичных неравенств из (3) и (2) в форме (6), которая использована для четверок и пятерок бинарных переменных. Так, например, при $r = 3$ (соответствует четным „шестеркам“ и нечетным „семеркам“ бинарных переменных) имеется $64 = 4^3$ неравенства. Форма (6) для них требует слишком много места, чтобы отразить все возможные комбинации ± 1 для шести и семи-индексных бинарных переменных.

Описать полные семейства неравенств (2) и (3) для больших r можно в более экономной форме, но это не будет предметом обсуждения в работе.

4. ОСНОВНЫЕ ВЫВОДЫ ИЗ ЛЕММЫ 2

Отметим два центральных момента, которые следуют из леммы 2 для

квадратичных экстремальных задач от n булевых $x \in \{0, 1\}^n$ или бинарных $y \in \{-1, 1\}^n$ переменных. Булевая переменная $x_i \in \{0, 1\}$ описывается квадратичным равенством $x_i^2 = x_i$, а бинарная переменная $y_i \in \{-1, +1\}$ описывается квадратичным равенством $y_i^2 = 1$. Принципиальной разницы между этими задачами нет. С помощью замены переменных $y_i = 1 - 2x_i$, $i = 1, \dots, n$ любую квадратичную экстремальную задачу с булевыми переменными можно свести к квадратичной экстремальной задаче с бинарными переменными (и наоборот).

Первый вывод из леммы 2: *использование идеи нечетного количества ± 1 позволяет в квадратичных экстремальных задачах с бинарными (а значит и с булевыми) переменными автоматизировать построение дополнительных квадратичных неравенств, справедливых для всевозможных значений бинарных переменных $y \in \{-1, 1\}^n$ (либо булевых переменных $x \in \{0, 1\}^n$).* В своем методе получения двойственных оценок Н.З.Шор такие неравенства называет функционально избыточными квадратичными ограничениями, так как они не изменяют множества допустимых решений в квадратичных экстремальных задачах с бинарными (булевыми) переменными [1, 5]. Двойственные алгоритмы для квадратичной экстремальной задачи, дополненной такими неравенствами, позволяют более точно аппроксимировать оптимальное значение целевой функции: сверху в задачах на максимум и снизу в задачах на минимум.

Для квадратичных экстремальных задач с бинарными переменными более удобной является первая форма неравенств (2) и (3) (для задач с булевыми переменными аналоги этой первой формы переписываются с учетом замены переменных $y_i = 1 - 2x_i$, $i = 1, \dots, n$). Первая форма определяется как полный квадрат от линейной комбинации бинарных переменных и для нее легко вычислять значение, которое принимает эта линейная комбинация при некотором заданном векторе $\bar{y} = (\bar{y}_1, \dots, \bar{y}_n)$. В двойственных алгоритмах это является существенным преимуществом, так как невязки ограничений в квадратичной задаче являются компонентами суб(супер)градиента негладких функций от неизвестных множителей Лагранжа.

Второй вывод из леммы 2 связан со второй формой неравенств в (2) и (3) и утверждает: *использование второй формы квадратичных неравенств для четного и нечетного количества бинарных переменных позволяет автоматизировать переход от квадратичных экстремальных задач с бинарными переменными к их линейным аналогам, которые получены релаксацией (ослаблением) бинарных переменных.* На самом деле, вторые формы квадратичных неравенств в (2) и (3) фа-

ктически подготовлены для использования переменных $y_{ij} = y_i y_j$ и их можно записать в виде следующих линейных неравенств:

$$\sum_{i=1}^{2r} \sum_{j>i}^{2r+1} l_i l_j y_{ij} \geq -r \quad \text{и} \quad \sum_{i=1}^{2r} l_0 l_i y_i + \sum_{i=1}^{2r-1} \sum_{j>i}^{2r} l_i l_j y_{ij} \geq -r.$$

При $r = 1$ (соответствует тройкам и двойкам бинарных переменных) линейные неравенства приобретают следующий вид

$$\left\{ \begin{array}{l} +y_{ij} + y_{ik} + y_{jk} \geq -1, \\ -y_{ij} - y_{ik} + y_{jk} \geq -1, \\ -y_{ij} + y_{ik} - y_{jk} \geq -1, \\ +y_{ij} - y_{ik} - y_{jk} \geq -1, \end{array} \right. \quad \text{и} \quad \left\{ \begin{array}{l} +y_i + y_j + y_{ij} \geq -1, \\ -y_i - y_j + y_{ij} \geq -1, \\ -y_i + y_j - y_{ij} \geq -1, \\ +y_i - y_j - y_{ij} \geq -1. \end{array} \right.$$

В теории математического программирования первая четверка линейных неравенств известна как правило „треугольника“, а вторая четверка линейных неравенств известна как самое простое правило для линеаризации квадратичных булевых задач. Это правило следует из релаксации бинарных переменных y_i и y_j :

$$y_i \in \{-1, 1\} \quad \Longrightarrow \quad -1 \leq y_i \leq 1 \quad \equiv \quad \left\{ \begin{array}{l} 1 + y_i \geq 0, \\ 1 - y_i \geq 0, \end{array} \right.$$

$$y_j \in \{-1, 1\} \quad \Longrightarrow \quad -1 \leq y_j \leq 1 \quad \equiv \quad \left\{ \begin{array}{l} 1 + y_j \geq 0, \\ 1 - y_j \geq 0, \end{array} \right.$$

и использовании для релаксированных переменных очевидных неравенств:

$$\left\{ \begin{array}{l} (1 + y_i)(1 + y_j) \geq 0, \\ (1 - y_i)(1 - y_j) \geq 0, \\ (1 - y_i)(1 + y_j) \geq 0, \\ (1 + y_i)(1 - y_j) \geq 0, \end{array} \right. \quad \Longrightarrow \quad \left\{ \begin{array}{l} +y_i + y_j + y_i y_j \geq -1, \\ -y_i - y_j + y_i y_j \geq -1, \\ -y_i + y_j - y_i y_j \geq -1, \\ +y_i - y_j - y_i y_j \geq -1. \end{array} \right.$$

При $r = 2$ (соответствует пятеркам и четверкам бинарных переменных), $r = 3$ (соответствует семеркам и шестеркам бинарных переменных) и т. д. линейные неравенства имеют аналогичный вид и их можно считать правилами „пятиугольника“, „семиугольника“ и т. д.

Ниже рассмотрим применение леммы 2 в рамках второго вывода, а именно будем использовать вытекающие из нее семейства линейных неравенств для аппроксимации сверху некоторого многогранника, который построен по типу булевого квадратичного многогранника введенного М. Падбергом (Padberg M.W.) в [2].

5. Многогранники $BiQP_n$ и $BoQP_n$

Рассмотрим выпуклый многогранник

$$BiQP_n = \text{conv}\{(y, x) \in \{-1, 1\}^{n+\frac{n(n-1)}{2}} : x_{ij} = y_i y_j \forall (i, j) : 1 \leq i < j \leq n\}$$

и условимся его называть бинарным квадратичным многогранником (binary quadratic polytope). Он есть выпуклая оболочка всех возможных пар бинарных векторов $y \in \{-1, 1\}^n$ и $x \in \{-1, 1\}^{\frac{n(n-1)}{2}}$, где на компоненты вектора x наложены следующие ограничения: бинарная переменная x_{ij} равна 1 только тогда, когда обе бинарные переменные y_i и y_j одинаковы, и равна -1 только тогда, когда бинарные переменные y_i и y_j различные. Эти ограничения описываются с помощью функции $x_{ij} = y_i y_j$ (логическая функция два к двум), т.е. в двух случаях x_{ij} будет равно 1 (когда $y_i = y_j = 1$ и когда $y_i = y_j = -1$), а в двух случаях равно -1 (когда $y_i = 1$ и $y_j = -1$, и когда $y_i = -1$ и $y_j = 1$).

Другая логическая функция положена в основу булевого квадратичного многогранника (boolean quadratic polytope) Падберга (условимся его обозначать $BoQP_n$). Он определяется так:

$$BoQP_n = \text{conv}\{(x, y) \in \{0, 1\}^{n+\frac{n(n-1)}{2}} : y_{ij} = x_i x_j \forall (i, j) : 1 \leq i < j \leq n\}$$

и есть выпуклая оболочка всех возможных пар булевых векторов $x \in \{0, 1\}^n$ и $y \in \{0, 1\}^{\frac{n(n-1)}{2}}$, где ограничения на компоненты булевого вектора y определяются некоторой логической функцией один к трем. На самом деле, булева компонента y_{ij} равна единице только в одном случае, а именно когда обе булевы компоненты x_i и x_j равны единице. В трех оставшихся случаях она принимает значение равное нулю.

Эти две различные по своей сути логические функции играют важную роль в теории графов. Так, например, с помощью многогранника $BiQP_n$ удобно описывать разрезы графа (ребрами связаны только вершины из двух различных подмножеств вершин графа), а с помощью многогранника $BoQP_n$ – свойства отдельных подмножеств графа (ребрами связаны только вершины этого подмножества вершин графа).

Для многогранников $BiQP_n$ и $BoQP_n$ при значениях $n = 2 \div 6$ с помощью программы Skeleton (<http://www.uic.mnov.ru/~zny/skeleton>) были построены неприводимые системы линейных неравенств [6]. Для описания многогранников $BiQP_5$ и $BoQP_5$ системы линейных неравенств включают по 368 неравенства, а для описания многогранников $BiQP_6$ и $BoQP_6$ – по 116764 неравенства и чтобы их найти потребовалось 30 минут счетного времени на процессоре Pentium M 1.60 GHz.

6. АППРОКСИМАЦИИ МНОГОГРАННИКА $BiQP_n$

Легко видеть, что квадратичные неравенства леммы 2 очень подходят для верхней аппроксимации многогранника $BiQP_n$ с помощью многогранников, заданных семействами линейных неравенств для четных и нечетных подмножеств бинарных переменных $y \in \{-1, 1\}^n$. Действительно, с помощью переменных $x_{ij} = y_i y_j$ вторые формы квадратичных неравенств из (2) и (3) записываются в форме линейных неравенств. Все возможные линейные неравенства для нечетной длины последовательностей и всех $r \in 1, \dots, r_1$, где r_1 такое, что $2r_1 + 1 \leq n$, можно записать в следующем виде:

$$\sum_{j=1}^{2r} \sum_{k>j}^{2r+1} l_j l_k x_{ij} x_{ik} \geq -r, \quad r=1, \dots, r_1, \quad \forall \left\{ \begin{array}{l} 1 \leq i_1 < \dots < i_{2r+1} \leq n, \\ l \in \{-1, 1\}^{2r+1} : l_1 = 1. \end{array} \right. \quad (7)$$

Неравенства (7) следуют из второй формы квадратичных неравенств (2). Все возможные линейные неравенства для четной длины последовательностей и всех $r \in 1, \dots, r_2$, где r_2 такое, что $2r_2 \leq n$, можно записать таким образом:

$$\sum_{j=1}^{2r} l_j y_{i_j} + \sum_{j=1}^{2r-1} \sum_{k>j}^{2r} l_j l_k x_{i_j i_k} \geq -r, \quad r=1, \dots, r_2, \quad \forall \left\{ \begin{array}{l} 1 \leq i_1 < \dots < i_{2r} \leq n, \\ l \in \{-1, 1\}^{2r}. \end{array} \right. \quad (8)$$

Рассмотрим следующий многогранник

$$LBiQP_n(r_1, r_2) = \{(y, x) \in R^{n+\frac{n(n-1)}{2}} : (y, x) \text{ удовлетворяют (7) и (8)}\},$$

где r_1 и r_2 такие, что $2r_1 + 1 \leq n$ и $2r_2 \leq n$. Самый полный многогранник $LBiQP_n(r_1, r_2)$, где r_1 и r_2 принимают все возможные значения, условимся обозначать $LBiQP_n$. Этот многогранник реализуется для четных n при $r_1 = \frac{n}{2} - 1$ и $r_2 = \frac{n}{2}$, а для нечетных n при $r_1 = r_2 = \frac{n-1}{2}$.

Теорема 1. Многогранник $LBiQP_n(r_1, r_2)$ аппроксимирует сверху многогранник $BiQP_n$, его описание содержит

$$M(r_1, r_2) = \sum_{r=1}^{r_2} 4^r C_n^{2r} + \sum_{r=1}^{r_1} 4^r C_n^{2r+1}$$

линейных неравенств. Описание многогранника $LBiQP_n$ содержит $M(n) = \frac{3^{n+1} + (-1)^n}{4} - (n+1)$ линейных неравенств: количество неравенств для четных n равно $\frac{3(3^n - 1)}{4} - n$, а для нечетных n равно $\frac{3^{n+1} - 5}{4} - n$.

Доказательство. Многогранник $BiQP_n$ может быть задан с помощью следующей системы квадратичных неравенств:

$$y_i^2 = 1, \quad i = 1, \dots, n, \quad x_{ij} = y_i y_j, \quad x_{ij}^2 = 1 \quad \forall (i, j) : 1 \leq i < j \leq n. \quad (9)$$

Учитывая бинарность переменных $y \in \{-1, 1\}^n$ из леммы 2 следует, что множество решений системы (9) не изменится, если к ней прибавить семейство квадратичных неравенств

$$\sum_{j=1}^{2r} \sum_{k>j}^{2r+1} l_j l_k y_{i_j} y_{i_k} \geq -r, \quad r=1, \dots, r_1, \quad \forall \left\{ \begin{array}{l} 1 \leq i_1 < \dots < i_{2r+1} \leq n, \\ l \in \{-1, 1\}^{2r+1} : l_1 = 1, \end{array} \right. \quad (10)$$

где $2r_1 + 1 \leq n$, и семейство квадратичных неравенств

$$\sum_{j=1}^{2r} l_j y_{i_j} + \sum_{j=1}^{2r-1} \sum_{k>j}^{2r} l_j l_k y_{i_j} y_{i_k} \geq -r, \quad r=1, \dots, r_2, \quad \forall \left\{ \begin{array}{l} 1 \leq i_1 < \dots < i_{2r} \leq n, \\ l \in \{-1, 1\}^{2r}. \end{array} \right. \quad (11)$$

где $2r_2 \leq n$. Здесь квадратичные неравенства (10) есть квадратичными неравенствами вида (2) для всех возможных нечетных подмножеств бинарных переменных от троек (соответствует $r = 1$) до $2r_1 + 1$. Их количество $M(r_1) = \sum_{r=1}^{r_1} 4^r C_n^{2r+1}$. Квадратичные неравенства (11) есть

квадратичными неравенствами вида (3) для всех возможных четных подмножеств бинарных переменных от двоек ($r = 1$) до $2r_2$. Их количество $M(r_2) = \sum_{r=1}^{r_2} 4^r C_n^{2r}$.

Множество решений системы (9)–(11) не изменится, если в неравенствах (10) и (11) квадратичные термы $y_{i_j} y_{i_k}$ заменить на линейные $x_{i_j i_k}$ для всех $1 \leq j < k \leq n$. В результате получаем систему (7)–(9), которая эквивалентна системе (9)–(11). Релаксация (ослабление) бинарных переменных $y \in \{-1, 1\}^n$ и $x \in \{-1, 1\}^{\frac{n(n-1)}{2}}$ для системы (7)–(9) дает многогранник, определяемый следующими линейными неравенствами

$$-1 \leq y_i \leq 1, \quad i = 1, \dots, n, \quad -1 \leq x_{ij} \leq 1, \quad \forall (i, j) : 1 \leq i < j \leq n, \quad (12)$$

$$(y, x) \in LBiQP_n(r_1, r_2). \quad (13)$$

В силу своего построения многогранник (12), (13) аппроксимирует сверху бинарный квадратичный многогранник $BiQP_n$. Однако, неравенства (12) и (13) являются избыточными (несущественными). Действительно, в многогранник $LBiQP_n(r_1, r_2)$ как обязательные включены линейные

неравенства для всех возможных двоек (пар) переменных y_i и y_j . Они следуют из квадратичных неравенств (4) и имеют такую форму:

$$\begin{cases} +y_i + y_j + x_{ij} \geq -1, \\ -y_i - y_j + x_{ij} \geq -1, \\ -y_i + y_j - x_{ij} \geq -1, \\ +y_i - y_j - x_{ij} \geq -1, \end{cases} \quad \forall (i, j) : 1 \leq i < j \leq n. \quad (14)$$

Легко видеть, что из неравенств (14) следуют линейные неравенства (12). Так, для произвольной пары (i, j) , сложив первое и второе неравенства из (14), имеем: $2x_{ij} \geq -2$ или $x_{ij} \geq -1$, а сложив третье и четвертое неравенства из (14) имеем: $-2x_{ij} \geq -2$ или $x_{ij} \leq 1$. По аналогичной схеме следуют и двусторонние границы на переменные y_i , $i = 1, \dots, n$. Так, сложив первое и четвертое неравенство из (14), получаем неравенство $y_i \geq -1$, а сложив второе и третье неравенства, получаем $y_i \leq 1$. Учитывая то, что индекс i принимает все значения $1, \dots, n-1$, получаем ограничения на все компоненты переменной y , за исключением последней. Границы на последнюю компоненту получаем путем сложения первого и третьего неравенств ($y_j \geq -1$) и второго и четвертого неравенств ($y_j \leq 1$), а индекс переменной y_j может принимать значение равное n .

Многогранник (12), (13) тождественен многограннику $LBiQP_n(r_1, r_2)$ и, следовательно, последний аппроксимирует сверху бинарный квадратичный многогранник $BiQP_n$. Многогранник $LBiQP_n(r_1, r_2)$ включает все неравенства для нечетных подмножеств $M(r_1)$ и все неравенства для четных подмножеств $M(r_2)$. Отсюда получаем полное количество неравенств

$$M(r_1, r_2) = M(r_1) + M(r_2) = \sum_{r=1}^{r_1} 4^r C_n^{2r+1} + \sum_{r=1}^{r_2} 4^r C_n^{2r}.$$

В заключение подсчитаем максимально возможное количество линейных неравенств в многограннике $LBiQP_n(r_1, r_2)$. Для этого воспользуемся двумя равенствами:

$$(1+2)^n = 1 + 2n + \sum_{i=2}^n 2^i C_n^i = 1 + 2n + \sum_{r=1}^{2r \leq n} 2^{2r} C_n^{2r} + \sum_{r=1}^{2r+1 \leq n} 2^{2r+1} C_n^{2r+1}$$

и

$$(1-2)^n = 1 - 2n + \sum_{i=2}^n (-2)^i C_n^i = 1 - 2n + \sum_{r=1}^{2r \leq n} 2^{2r} C_n^{2r} - \sum_{r=1}^{2r+1 \leq n} 2^{2r+1} C_n^{2r+1},$$

которые следуют из бинома Ньютона. Обозначив $\sum_{r=1}^{2r \leq n} 2^{2r} C_n^{2r}$ как S_{2r} , а $\sum_{r=1}^{2r+1 \leq n} 2^{2r+1} C_n^{2r+1}$ как S_{2r+1} , получаем систему из двух линейных уравнений

$$3^n = 1 + 2n + S_{2r} + S_{2r+1}$$

и

$$(-1)^n = 1 - 2n + S_{2r} - S_{2r+1},$$

из которой находим $S_{2r} = \frac{3^n + (-1)^n}{2} - 1$ и $S_{2r+1} = \frac{3^n - (-1)^n}{2} - 2n$. Полное количество линейных неравенств в многограннике $LBiQP_n$ равно

$$\begin{aligned} M(n) &= \sum_{r=1}^{2r+1 \leq n} 2^{2r} C_n^{2r+1} + \sum_{r=1}^{2r \leq n} 2^{2r} C_n^{2r} = \\ &= \frac{1}{2} S_{2r+1} + S_{2r} = \frac{3^{n+1} + (-1)^n}{4} - (n+1). \end{aligned}$$

Из последнего вытекает, что когда n – четное, то число неравенств равно $\frac{3(3^n - 1)}{4} - n$ (реализуется при $r_1 = \frac{n}{2} - 1$ и $r_2 = \frac{n}{2}$), а когда n – нечетное, то число неравенств равно $\frac{3^{n+1} - 5}{4} - n$ (реализуется при $r_1 = r_2 = \frac{n-1}{2}$). Теорема доказана. \square

При $n = 2, 3, 4$ многогранники $LBiQP_n$ полностью совпали с описываемыми многогранники $BiQP_n$ неприводимыми системами линейных неравенств, которые получены с помощью программы *Skeleton*. Многогранник $BiQP_2$ описывается 4 линейными неравенствами для пары бинарных переменных. Система неравенств для многогранника $BiQP_3$ включает 16 неравенств: из них 4 для тройки и 3 по 4 для трех возможных пар. В описании многогранника $BiQP_4$ используется 56 неравенств (из них 16 для четверки, 4 по 4 для четырех троек и 6 по 4 для шести пар переменных).

ЗАКЛЮЧЕНИЕ

В работе описано свойство нечетного количества ± 1 и показано, что оно позволяет в квадратичных экстремальных задачах с бинарными и булевыми переменными автоматизировать построение дополнительных квадратичных неравенств, справедливых независимо от того, какие значения принимают бинарные переменные $y \in \{-1, 1\}^n$, либо булевы переменные $x \in \{0, 1\}^n$. На его основе построены семейства квадратичных неравенств для четного и нечетного количества бинарных переменных.

Их линейризованные аналоги использованы для внешней аппроксимации бинарного квадратичного многогранника $BiQP_n$ (задает множество допустимых решений в задаче нахождения экстремума квадратичной функции от n бинарных переменных) с помощью многогранников с конечным количеством линейных неравенств от $n(n+1)/2$ переменных. Подсчитано максимальное количество линейных неравенств в аппроксимирующем многограннике.

Литература

1. Shor N.Z. Nondifferentiable optimization and polynomial problems. – Dordrecht: Kluwer, 1998. – 394 p.
2. Padberg M.W. The boolean quadratic polytope: some characteristics, facets and relatives // Mathematical Programming. – 1989. – Vol. 45. – P. 139-172.
3. Стецюк П.И. Новые модели квадратичного типа для задачи о максимальном взвешенном разрезе графа // Кибернетика и системный анализ. – 2006. – №1. – С. 63-75.
4. Stetsyuk P.I., Bardadym T.O. The Use of Odd Sets of Variables to Improve Bounds in Boolean Quadratic Problems // Abstracts of International Conference in Honor of the 70th Birthday of Academician I. V. Sergienko „Applied Optimization and Metaheuristic Innovations“ (July 19-21, 2006, Yalta, Ukraine). – P. 51.
5. Шор Н.З., Стеценко С.И. Квадратичные экстремальные задачи и недифференцируемая оптимизация. – Киев: Наукова думка, 1989. – 200 с.
6. Золотых Н.Ю. Новая модификация метода двойного описания для построения остова многогранного конуса // III Всероссийская конференция „Проблемы оптимизации и экономические приложения“: Материалы конференции. Омск: изд-во ОмГТУ, 2006. С. 140

Поступила 02.08.2010

О внешней аппроксимации бинарного и булевого квадратичных многогранников

Золотых Н.Ю., Стецюк П.И

V-а Міжнародна школа-семінар "Теорія прийняття рішень",

Ужгород, 27 вересня – 1 жовтня 2010.

Праці школи-семінару. – С. 103–104.

Многогранники $BiQP_n$ и $BoQP_n$. Бинарный квадратичный многогранник (binary quadratic polytope) определим следующим образом:

$$BiQP_n = \text{conv}\{(y, x) \in \{-1, 1\}^{n+\frac{n(n-1)}{2}} : x_{ij} = y_i y_j \forall (i, j) : 1 \leq i < j \leq n\}. \quad (1)$$

Многогранник $BiQP_n$ определен по типу булевого квадратичного многогранника (boolean quadratic polytope), введенного М. Падбергом (M.W. Padberg) в [1]:

$$BoQP_n = \text{conv}\{(x, y) \in \{0, 1\}^{n+\frac{n(n-1)}{2}} : y_{ij} = x_i x_j \forall (i, j) : 1 \leq i < j \leq n\}. \quad (1')$$

Правильные неравенства для $BiQP_n$ и $BoQP_n$. Для всех возможных пар $(i, j) : 1 \leq i < j \leq n$ бинарных переменных $y \in \{-1, 1\}^n$ справедливы неравенства:

$$\left\{ \begin{array}{l} (+1 + y_i + y_j)^2 \geq 1, \\ (-1 + y_i + y_j)^2 \geq 1, \\ (+1 - y_i + y_j)^2 \geq 1, \\ (+1 + y_i - y_j)^2 \geq 1, \end{array} \right. \Rightarrow \left\{ \begin{array}{l} +y_i + y_j + y_i y_j \geq -1, \\ -y_i - y_j + y_i y_j \geq -1, \\ -y_i + y_j - y_i y_j \geq -1, \\ +y_i - y_j - y_i y_j \geq -1, \end{array} \right. \Rightarrow \left\{ \begin{array}{l} +y_i + y_j + x_{ij} \geq -1, \\ -y_i - y_j + x_{ij} \geq -1, \\ -y_i + y_j - x_{ij} \geq -1, \\ +y_i - y_j - x_{ij} \geq -1, \end{array} \right. \quad (2)$$

а булевых переменных $x \in \{0, 1\}^n$ – неравенства:

$$\begin{aligned}
 \left\{ \begin{array}{l} (+3 - 2x_i - 2x_j)^2 \geq 1, \\ (+1 - 2x_i - 2x_j)^2 \geq 1, \\ (+1 + 2x_i - 2x_j)^2 \geq 1, \\ (+1 - 2x_i + 2x_j)^2 \geq 1, \end{array} \right. & \Rightarrow \left\{ \begin{array}{l} -x_i - x_j + x_i x_j \geq -1, \\ +x_i x_j \geq 0, \\ +x_i - x_i x_j \geq 0, \\ +x_j - x_i x_j \geq 0, \end{array} \right. \Rightarrow \\
 & \Rightarrow \left\{ \begin{array}{l} -x_i - x_j + y_{ij} \geq -1, \\ +y_{ij} \geq 0, \\ +x_i - y_{ij} \geq 0, \\ +x_j - y_{ij} \geq 0. \end{array} \right. \quad (2')
 \end{aligned}$$

Для всех возможных троек $(i, j, k) : 1 \leq i < j < k \leq n$ бинарных переменных $y \in \{-1, 1\}^n$ справедливы неравенства:

$$\begin{aligned}
 \left\{ \begin{array}{l} (+y_i + y_j + y_k)^2 \geq 1, \\ (-y_i + y_j + y_k)^2 \geq 1, \\ (+y_i - y_j + y_k)^2 \geq 1, \\ (+y_i + y_j - y_k)^2 \geq 1, \end{array} \right. & \Rightarrow \left\{ \begin{array}{l} +y_i y_j + y_i y_k + y_j y_k \geq -1, \\ -y_i y_j - y_i y_k + y_j y_k \geq -1, \\ -y_i y_j + y_i y_k - y_j y_k \geq -1, \\ +y_i y_j - y_i y_k - y_j y_k \geq -1, \end{array} \right. \Rightarrow \\
 & \Rightarrow \left\{ \begin{array}{l} +x_{ij} + x_{ik} + x_{jk} \geq -1, \\ -x_{ij} - x_{ik} + x_{jk} \geq -1, \\ -x_{ij} + x_{ik} - x_{jk} \geq -1, \\ +x_{ij} - x_{ik} - x_{jk} \geq -1, \end{array} \right. \quad (3)
 \end{aligned}$$

а булевых переменных $x \in \{0, 1\}^n$ – неравенства:

$$\begin{aligned}
 \left\{ \begin{array}{l} (+3 - 2x_i - 2x_j - 2x_k)^2 \geq 1, \\ (+1 + 2x_i - 2x_j - 2x_k)^2 \geq 1, \\ (+1 - 2x_i + 2x_j - 2x_k)^2 \geq 1, \\ (+1 - 2x_i - 2x_j + 2x_k)^2 \geq 1, \end{array} \right. & \Rightarrow \dots \\
 \dots & \Rightarrow \left\{ \begin{array}{l} +y_{ij} + y_{ik} + y_{jk} - x_i - x_j - x_k \geq -1, \\ -y_{ij} - y_{ik} + y_{jk} + x_i \geq 0, \\ -y_{ij} + y_{ik} - y_{jk} + x_j \geq 0, \\ +y_{ij} - y_{ik} - y_{jk} + x_k \geq 0. \end{array} \right. \quad (3')
 \end{aligned}$$

Для всех возможных четверок $(i, j, k, l) : 1 \leq i < j < k < l \leq n$ би-

нарных переменных $y \in \{-1, 1\}^n$ справедливы следующие неравенства:

$$\left\{ \begin{array}{l} +y_i + y_j + y_k + y_l + x_{ij} + x_{ik} + x_{il} + x_{jk} + x_{jl} + x_{kl} \geq -2, \\ -y_i - y_j - y_k - y_l + x_{ij} + x_{ik} + x_{il} + x_{jk} + x_{jl} + x_{kl} \geq -2, \\ -y_i + y_j + y_k + y_l - x_{ij} - x_{ik} - x_{il} + x_{jk} + x_{jl} + x_{kl} \geq -2, \\ +y_i - y_j + y_k + y_l - x_{ij} + x_{ik} + x_{il} - x_{jk} - x_{jl} + x_{kl} \geq -2, \\ +y_i + y_j - y_k + y_l + x_{ij} - x_{ik} + x_{il} - x_{jk} + x_{jl} - x_{kl} \geq -2, \\ +y_i + y_j + y_k - y_l + x_{ij} + x_{ik} - x_{il} + x_{jk} - x_{jl} - x_{kl} \geq -2, \\ +y_i - y_j - y_k - y_l - x_{ij} - x_{ik} - x_{il} + x_{jk} + x_{jl} + x_{kl} \geq -2, \\ -y_i + y_j - y_k - y_l - x_{ij} + x_{ik} + x_{il} - x_{jk} - x_{jl} + x_{kl} \geq -2, \\ -y_i - y_j + y_k - y_l + x_{ij} - x_{ik} + x_{il} - x_{jk} + x_{jl} - x_{kl} \geq -2, \\ -y_i - y_j - y_k + y_l + x_{ij} + x_{ik} - x_{il} + x_{jk} - x_{jl} - x_{kl} \geq -2, \\ -y_i - y_j + y_k + y_l + x_{ij} - x_{ik} - x_{il} - x_{jk} - x_{jl} + x_{kl} \geq -2, \\ -y_i + y_j - y_k + y_l - x_{ij} + x_{ik} - x_{il} - x_{jk} + x_{jl} - x_{kl} \geq -2, \\ -y_i + y_j + y_k - y_l - x_{ij} - x_{ik} + x_{il} + x_{jk} - x_{jl} - x_{kl} \geq -2, \\ +y_i - y_j - y_k + y_l - x_{ij} - x_{ik} + x_{il} + x_{jk} - x_{jl} - x_{kl} \geq -2, \\ +y_i - y_j + y_k - y_l - x_{ij} + x_{ik} - x_{il} - x_{jk} + x_{jl} - x_{kl} \geq -2, \\ +y_i + y_j - y_k - y_l + x_{ij} - x_{ik} - x_{il} - x_{jk} - x_{jl} + x_{kl} \geq -2, \end{array} \right. \quad (4)$$

которые следуют из неравенств

$$\left\{ \begin{array}{l} (+1 + y_i + y_j + y_k + y_l)^2 \geq 1, \\ (-1 + y_i + y_j + y_k + y_l)^2 \geq 1, \\ (+1 - y_i + y_j + y_k + y_l)^2 \geq 1, \\ (+1 + y_i - y_j + y_k + y_l)^2 \geq 1, \\ (+1 + y_i + y_j - y_k + y_l)^2 \geq 1, \\ (+1 + y_i + y_j + y_k - y_l)^2 \geq 1, \\ (-1 - y_i + y_j + y_k + y_l)^2 \geq 1, \\ (-1 + y_i - y_j + y_k + y_l)^2 \geq 1, \\ (-1 + y_i + y_j - y_k + y_l)^2 \geq 1, \\ (-1 + y_i + y_j + y_k - y_l)^2 \geq 1, \\ (+1 - y_i - y_j + y_k + y_l)^2 \geq 1, \\ (+1 - y_i + y_j - y_k + y_l)^2 \geq 1, \\ (+1 - y_i + y_j + y_k - y_l)^2 \geq 1, \\ (+1 + y_i - y_j - y_k + y_l)^2 \geq 1, \\ (+1 + y_i - y_j + y_k - y_l)^2 \geq 1, \\ (+1 + y_i + y_j - y_k - y_l)^2 \geq 1. \end{array} \right.$$

Неравенства (4') для булевых переменных $x \in \{0, 1\}^n$ приводить не будем в силу их громоздкости. Отметим лишь общую схему построения всех этих семейств правильных неравенств. Квадратичные неравенства

для многогранника $BiQP_n$ получаются из правила, что сумма нечетного количества бинарных чисел не может равняться нулю, и использования для бинарных переменных квадратичного равенства $y_i^2 = 1$, $i = 1, \dots, n$. Неравенства для многогранника $VoQP_n$ получаются из квадратичных неравенств для многогранника $BiQP_n$, если бинарные переменные $y \in \{-1, 1\}^n$ заменить на булевы переменные с помощью замены переменных $y_i = 1 - 2x_i$, $i = 1, \dots, n$, а уже для булевых переменных использовать квадратичное равенство $x_i^2 = x_i$, $i = 1, \dots, n$.

Выводы. С помощью программы Skeleton (<http://www.uic.mnov.ru/~zny/skeleton>) были построены неприводимые системы линейных неравенств, описывающие многогранники $BiQP_n$ и $VoQP_n$ для $n = 2, 3, 4, 5, 6$. Оказалось, что приведенные в настоящей заметке неравенства для многогранников $BiQP_n$ и $VoQP_n$ при $n = 2, 3, 4$ полностью совпали с полученными с помощью программы. Так, например, многогранник $BiQP_2$ описывается 4 линейными неравенствами (2). Система неравенств для многогранника $BiQP_3$ включает 16 неравенств – из них 4 для тройки и 3 по 4 для трех возможных пар. В описании многогранника $BiQP_4$ используется 56 неравенств (из них 16 для четверки, 4 по 4 для четырех троек и 6 по 4 для шести пар переменных)

Литература

1. PADBERG M.W. The boolean quadratic polytope: some characteristics, facets and relatives // *Mathematical Programming*. – 1989. – v. 45. – P. 139–172.

Р А З Д Е Л IV

Задача о максимальном независимом множестве вершин графа

В раздел включены статьи:

- 4.1. О новых лагранжевых двойственных оценках для числа устойчивости графа (Стецюк, Пардалос, Крошко, 2006) 250
- 4.2. Об одной верхней оценке для взвешенного числа устойчивости графа (Стецюк, Бутенко, Березовский, 2007) 261
- 4.3. Точная ЛП-оценка для взвешенного числа устойчивости t -совершенных графов (Стецюк, Ляшко, Нурминский, 2009) 271
- 4.4. О новых свойствах оценок Шора для взвешенного числа устойчивости графа (Стецюк, 2008) 284
- 4.5. ЛП-ориентированная верхняя оценка для числа устойчивости графа на основе p -колес (Стецюк, Бутенко, Лиховид, 2008) 298
- 4.6. Об ЛП-ориентированных верхних оценках для взвешенного числа устойчивости графа (Стецюк, Лиховид, 2009) 310
- 4.7. Оценки Н.З. Шора в квадратичных экстремальных задачах и их применение в комбинаторной оптимизации (Стецюк, 2010) 333

О новых лагранжевых двойственных оценках для числа устойчивости графа

Стецюк П.И., Пардалос П.М., Крошко Д.Л.

*Компьютерная математика. – Киев: Ин-т кибернетики
им. В.М. Глушкова НАН Украины, 2006. – Выпуск 3.
– С. 149–158.*

Исследуются две новые верхние лагранжевые двойственные оценки для числа устойчивости графа. Они базируются на использовании специального вида функционально избыточных квадратичных ограничений. Доказано, что обе оценки являются не менее точными, чем предложенная Н.З.Шором самая точная верхняя оценка.

Введение. Пусть $G = (V, E)$ – неориентированный граф с множеством вершин $V = \{1, \dots, n\}$ и множеством ребер E . Подмножество вершин $S \subseteq V$ называется устойчивым множеством (stable set) графа G , если для любых $(i, j) \in S$ выполняется $(i, j) \notin E$. Другими словами, S – устойчивое множество, если порожденный (индуцированный) им подграф из G не содержит ни одного ребра из E . Число устойчивости графа G определяется $\alpha(G) = \max\{|S| : S \subseteq V\}$, где S – устойчивое множество.

Задача о максимальном устойчивом множестве для графа G состоит в нахождении числа устойчивости $\alpha(G)$. В общем случае эта задача является NP-трудной [1], учитывая, что для общего вида графа многогранник устойчивых множеств (принято обозначать $\text{STAB}(G)$) может иметь очень сложную структуру [2]. В силу этого актуальны вопросы нахождения эффективных по точности верхних оценок для $\alpha(G)$.

Один из механизмов для этого – разработанная Н.З. Шором техника оптимальных лагранжевых двойственных оценок ψ^* в сочетании с использованием функционально избыточных ограничений для улучшения точности таких оценок [3–5]. Суть этой техники такова. Пусть проблема сформулирована в форме невыпуклой квадратичной оптимизационной задачи (исходная квадратичная задача), которой соответствует оптимальная лагранжевая двойственная оценка ψ_0^* . Она может быть найдена с помощью современных методов недифференцируемой оптимизации [4]. Оценка ψ_0^* – оценка сверху в задаче на максимум (оценка

снизу в задаче на минимум) для оптимального значения целевой функции. Пусть имеются функционально избыточные квадратичные ограничения, которые построены как нетривиальные следствия из имеющихся ограничений исходной квадратичной задачи, т.е. они не являются линейными комбинациями уже существующих ограничений. Добавление таких ограничений к исходной квадратичной задаче не изменяет множества ее оптимальных решений, но изменяет функцию Лагранжа для новой квадратичной задачи. В результате соответствующая новой квадратичной задаче оптимальная лагранжевая двойственная оценка ψ_1^* всегда не менее точная, чем оценка ψ_0^* , соответствующая исходной квадратичной задаче. Часто оказывается, что оценка ψ_1^* является более точной верхней (нижней) оценкой для оптимального значения целевой функции, чем оценка ψ_0^* .

В работе рассмотрим ряд новых лагранжевых двойственных оценок для задачи о максимальном устойчивом множестве и докажем, что две из них являются не менее точными верхними оценками для $\alpha(G)$, чем наилучшая из известных оценок Н.З. Шора.

1. О трех оценках Н.З.Шора. В работах [3,4] Н.З.Шором предложено три оптимальные лагранжевые двойственные оценки для задачи о максимальном устойчивом множестве вершин графа. Обозначим их $\psi_1^*(G)$, $\psi_2^*(G)$ и $\psi_3^*(G)$. Каждая оценка является верхней оценкой (оценкой сверху) для $\alpha(G)$, и самая точная из них – оценка $\psi_3^*(G)$.

Оценка $\psi_1^*(G)$ – оптимальная лагранжевая двойственная оценка для квадратичной задачи в следующей формулировке:

$$\alpha(G) = \max \sum_{i=1}^n x_i \quad (1)$$

при ограничениях

$$x_i x_j = 0, \quad \forall (i, j) \in E, \quad (2)$$

$$x_i^2 - x_i = 0, \quad i = 1, \dots, n. \quad (3)$$

Здесь булева переменная x_i равна единице, если вершина i включается в устойчивое множество, и равна нулю, иначе. Задача (1)–(3) – простейшая квадратичная задача в булевых переменных, которая для графа G полностью описывает проблему о максимальном устойчивом множестве. Здесь булевые переменные $x_i \in \{0, 1\}$ описаны квадратичными равенствами (3). Квадратичные ограничения (2) означают, что если в графе G две вершины связаны ребром, то они обе одновременно не могут принадлежать устойчивому множеству.

Оценка $\psi_2^*(G)$ – оптимальная лагранжевая двойственная оценка для квадратичной задачи, которая получена в результате прибавления к простейшей задаче семейства функционально избыточных ограничений

$$x_i x_j \geq 0, \quad \forall (i, j) \notin E. \quad (4)$$

Ограничения (4) следуют из того, что положительность переменных $x_i = x_i^2 \geq 0$ для всех $i = 1, \dots, n$ позволяет сформулировать квадратичные неравенства в форме $x_i x_j \geq 0$ для всех ребер (i, j) дополнительно к G графа. Ограничения (4) дополняют простейшую квадратичную задачу отсутствующими связями между переменными для тех вершин, которые не связаны ребрами. Тем самым, они благоприятствуют тому, что для некоторых графов оценка $\psi_2^*(G)$, которая соответствует задаче (1)–(4), оказывается более точной, чем оценка $\psi_1^*(G)$.

Оценки $\psi_1^*(G)$ и $\psi_2^*(G)$ связаны с числами Ловаса $\vartheta(G)$ и $\vartheta'(G)$ [2]. В работе [4] показано, что $\psi_1^*(G) = \vartheta(G)$, а $\psi_2^*(G) = \vartheta'(G)$. Очередным уточнением оценки $\psi_2^*(G)$, а значит и числа Ловаса $\vartheta'(G)$, является оценка $\psi_3^*(G)$. Она является оптимальной лагранжевой двойственной оценкой для квадратичной задачи, полученной из задачи (1)–(4) прибавлением такого семейства функционально избыточных ограничений

$$x_i x_k + x_j x_k \leq x_k, \quad \forall (i, j) \in E, \quad k \neq i, j. \quad (5)$$

Ограничения (5) построены по следующему правилу. Для булевых переменных x_i и x_j , когда $(i, j) \in E$, всегда справедливо линейное неравенство $x_i + x_j \leq 1$. Если его умножить на переменную x_k , такую, что $k \neq i, j$, то знак неравенства не изменится в силу того, что $x_k = x_k^2 \geq 0$. В результате получаем квадратичное неравенство $x_i x_k + x_j x_k \leq x_k$, справедливое для $(i, j) \in E$ и $k \neq i, j$. Семейство ограничений (5) содержит все возможные квадратичные неравенства, построенные по этому правилу.

Оценка $\psi_3^*(G)$ будет всегда не менее точной, чем оценка $\psi_2^*(G)$. Функционально избыточные ограничения (5) наполняют квадратичную задачу (1)–(5) всеми возможными связями для троек (i, j, k) вершин графа G , таких, что $(i, j) \in E$ и $k \neq i, j$. Наличие этих связей придает оценке $\psi_3^*(G)$ ряд замечательных свойств, которые связаны со специальными классами графов. Так, например, в [4] показано, что оценка $\psi_3^*(G)$ – точная верхняя оценка для $\alpha(G)$, когда граф G является t -перфектным, либо h -перфектным.

2. Построение новых верхних оценок. Новые лагранжевые двойственные оценки для $\alpha(G)$ построим аналогично тому, как построены

оценки Н.З. Шора, т.е. в качестве базовой возьмем простейшую квадратичную задачу и будем последовательно начинать ее семействами функционально избыточных ограничений по методике из [6], которая базируется на свойствах бинарных (± 1) переменных. Поэтому, нашей базовой квадратичной задачей будет не простейшая задача (1)–(3) в булевых переменных, а ее аналог, сформулированный в бинарных переменных y_i , где

$$y_i = 1 - 2x_i, \quad i = 1, \dots, n. \quad (6)$$

Здесь вершине i соответствует бинарная переменная y_i , которая равна -1 , если вершина i включается в устойчивое множество, и равна 1 в противном случае.

Опишем последовательность действий (шагов) для построения квадратичной оптимизационной задачи, которая стоит за новыми верхними оценками. Проанализируем, что будет происходить с оптимальными лагранжевыми двойственными оценками на каждом шаге.

Шаг 1. Переформулируем задачу (1)–(3) с помощью замены переменных (6). В результате для максимального устойчивого множества получим квадратичную задачу в следующей формулировке:

$$\alpha(G) = \max \left(\frac{n}{2} - \sum_{i=1}^n \frac{y_i}{2} \right) \quad (7)$$

при ограничениях

$$y_i y_j - y_i - y_j + 1 = 0, \quad \forall (i, j) \in E, \quad (8)$$

$$y_i^2 = 1, \quad i = 1, \dots, n, \quad (9)$$

которая является простейшей квадратичной задачей в бинарных переменных. Задаче (7)–(9) соответствует в точности такая же оценка $\psi_1^*(G)$, как и задаче (1)–(3).

Шаг 2. Добавим к задаче (7)–(9) следующие квадратичные ограничения:

$$\begin{cases} y_i y_j - y_{ij} = 0, \\ y_{ij}^2 - 1 = 0, \end{cases} \quad \forall i, j : 1 \leq i < j \leq n, \quad (10)$$

чем “расширим” множество бинарных переменных в задаче (7)–(10) по отношению к простейшей квадратичной задаче. Здесь ограничения (10) описывают новые переменные, которые для каждой пары (i, j) , $i < j$ вводятся по правилу $y_{ij} = y_i y_j$, и то, что новые переменные являются бинарными, следует из правила их введения. Количество бинарных

переменных в "расширенной" задаче (7)–(10) равно $\frac{n(n+1)}{2}$, из них n старых переменных и $\frac{n(n-1)}{2}$ новых переменных.

Оптимальная лагранжевая двойственная оценка для "расширенной" задачи (7)–(10) будет всегда не менее точной, чем $\psi_1^*(G)$. Более того, ее можно улучшить, дополняя задачу (7)–(10) функционально избыточными ограничениями, которые связывают старые и новые переменные. Собственно в этом и состоят последующие шаги, каждый из которых последовательно включает в задачу семейства функционально избыточных ограничений для двоек, троек и четверок переменных. Естественно, что каждый очередной шаг способствует уточнению лагранжевых двойственных оценок, так как он включает больше дополнительных связей между новыми и старыми переменными, чем предыдущий. В результате можно утверждать, что всегда найдется такой граф G , что оценка после каждого очередного шага будет более точной, чем предыдущая.

Шаг 3. С целью учета связей для всех возможных двоек переменных добавим такое семейство функционально избыточных квадратичных ограничений:

$$\begin{cases} y_{ij}y_i - y_j = 0, \\ y_{ij}y_j - y_i = 0. \end{cases} \quad \forall i, j : 1 \leq i < j \leq n. \quad (11)$$

Они следуют из того, что для произвольной пары бинарных переменных y_i, y_j справедливы равенства $y_i = y_i y_j^2 = (y_i y_j) y_j = y_{ij} y_j$, $y_j = y_j y_i^2 = (y_i y_j) y_i = y_{ij} y_i$.

Оценку для задачи (7)–(11) обозначим $\psi_a^*(G)$. Она не менее точна, чем оценка $\psi_2^*(G)$, что докажем далее при доказательстве свойств более сильных оценок.

Шаг 4. С целью учета связей для всех возможных троек переменных добавим семейство функционально избыточных квадратичных ограничений

$$\begin{cases} y_{ij}y_k - y_{ik}y_j = 0, \\ y_{ij}y_k - y_{jk}y_i = 0, \\ y_{ik}y_{jk} - y_{ij} = 0, \\ y_{ij}y_{jk} - y_{ik} = 0, \\ y_{ij}y_{ik} - y_{jk} = 0, \end{cases} \quad \forall i, j, k : 1 \leq i < j < k \leq n, \quad (12)$$

которое является следствием двух факторов. Во-первых, для каждой тройки бинарных переменных y_i, y_j и y_k справедливы такие равенства:

$$\begin{aligned} y_{ij} &= y_i y_j = y_i y_j y_k^2 = (y_i y_k)(y_j y_k) = y_{ik} y_{jk}, \\ y_{ik} &= y_i y_k = y_i y_k y_j^2 = (y_i y_j)(y_j y_k) = y_{ij} y_{jk}, \\ y_{jk} &= y_j y_k = y_j y_k y_i^2 = (y_i y_j)(y_i y_k) = y_{ij} y_{ik}. \end{aligned}$$

Во-вторых, из неоднозначности представления произведения

$$y_i y_j y_k = (y_i y_j) y_k = y_{ij} y_k = (y_i y_k) y_j = y_{ik} y_j = (y_j y_k) y_i = y_{jk} y_i$$

имеем еще два квадратичных равенства. Заметим, что из неоднозначного представления $y_i y_j y_k$ следует три равенства. Но третье – линейная комбинация двух первых равенств и не влияет на точность лагранжевой двойственной оценки. Оценку для задачи (7)–(12) условимся обозначать $\psi_b^*(G)$. Она будет не менее точной, чем оценка $\psi_3^*(G)$ (доказательство опишем далее).

Шаг 5. Для учета дополнительных связей для всех возможных четверок переменных добавим семейство функционально избыточных ограничений

$$\begin{cases} y_{ij} y_{kl} - y_{ik} y_{jl} = 0, \\ y_{ij} y_{kl} - y_{il} y_{jk} = 0, \end{cases} \quad \forall i, j, k, l : 1 \leq i < j < k < l \leq n. \quad (13)$$

Ограничения (13) содержат два линейно независимых равенства для каждой четверки вершин i, j, k и l , которые следуют из неоднозначного представления произведения $y_i y_j y_k y_l = y_{ij} y_{kl} = y_{ik} y_{jl} = y_{il} y_{jk}$. Здесь по той же причине, что и для троек, не включено третье равенство, которое не будет влиять на точность оптимальной лагранжевой двойственной оценки.

Оценку для задачи (7)–(13) условимся обозначать $\psi_c^*(G)$. Естественно, что она будет не менее точной верхней оценкой для $\alpha(G)$, чем оценка $\psi_b^*(G)$.

Итак, для максимального устойчивого множества имеем новую квадратичную задачу (7)–(13) в бинарных переменных. Она получена дополнением простейшей квадратичной задачи полным семейством функционально избыточных ограничений из [6]. Если от бинарных переменных перейти к булевым переменным (по правилу $y_i = 1 - 2x_i$ и $y_{ij} = 1 - 2x_{ij}$), то тогда квадратичной задаче (7)–(13) соответствует квадратичная задача в булевых переменных, которую рассмотрим далее и изучим свойства двух самых сильных оценок.

3. Оценки $\psi_b^*(G)$ и $\psi_c^*(G)$ и их свойства. Обе новые оценки связаны с формулировкой проблемы о максимальном устойчивом множестве

в форме следующей квадратичной задачи:

$$\alpha(G) = \max \sum_{i=1}^n x_i \quad (14)$$

при ограничениях

$$x_i x_j = 0, \forall (i, j) \in E, \quad (15)$$

$$x_i^2 - x_i = 0, \quad i = 1, \dots, n, \quad (16)$$

$$\begin{cases} 2x_i x_j - x_i - x_j + x_{ij} = 0, \\ x_{ij} - x_{ji} = 0, \\ x_{ij}^2 - x_{ij} = 0, \end{cases} \quad \forall i, j, \quad (17)$$

$$\begin{cases} 2x_i x_{ij} - x_i + x_j - x_{ij} = 0, \\ 2x_j x_{ij} + x_i - x_j - x_{ij} = 0, \end{cases} \quad \forall i, j \quad (18)$$

$$\begin{cases} 2x_k x_{ij} - 2x_j x_{ik} + x_j - x_k - x_{ij} + x_{ik} = 0, \\ 2x_k x_{ij} - 2x_i x_{jk} + x_i - x_k - x_{ij} + x_{jk} = 0, \\ 2x_{ik} x_{jk} + x_{ij} - x_{ik} - x_{jk} = 0, \\ 2x_{ij} x_{jk} - x_{ij} + x_{ik} - x_{jk} = 0, \\ 2x_{ij} x_{ik} - x_{ij} - x_{ik} + x_{jk} = 0, \end{cases} \quad \forall i, j, k \quad (19)$$

$$\begin{cases} 2x_{ij} x_{kl} - 2x_{ik} x_{jl} - x_{ij} + x_{ik} + x_{jl} - x_{kl} = 0, \\ 2x_{ij} x_{kl} - 2x_{il} x_{jk} - x_{ij} + x_{il} + x_{jk} - x_{kl} = 0, \end{cases} \quad \forall i, j, k, l. \quad (20)$$

Оценки $\psi_b^*(G)$ и $\psi_c^*(G)$ соответствуют оптимальным лагранжевым двойственным оценкам для задач (14)–(19), (14)–(20).

Формулировка задачи (14)–(20) отличается от формулировки (7)–(13) тем, что в ней отсутствуют упорядоченность индексов тех переменных, которые входят в соответствующие группы ограничений. Для этого в группу ограничений (15), которые связаны с новыми переменными, добавлено ограничение $x_{ij} = x_{ji}$ в явной форме. Это делает формулировку задачи (7)–(13) в терминах булевых переменных более удобной для доказательства свойств оценок (хотя заметим, что многие функционально избыточные ограничения при этом дублируются). Чтобы формулировка задачи (14)–(20) была более краткой, в ней опущена и запись того, что в ограничениях, которые справедливы для многих индексов, среди этих индексов нет двух совпадающих. Это будем считать само собой разумеющимся.

Для оценок $\psi_b^*(G)$ и $\psi_c^*(G)$ справедлива следующая теорема.

Теорема. Для произвольного графа G имеет место неравенство

$$\psi_3^*(G) \geq \psi_b^*(G) \geq \psi_c^*(G),$$

и, следовательно, оценка $\psi_b^*(G)$ (а тем более $\psi_c^*(G)$) – не менее точная верхняя оценка для $\alpha(G)$, чем оценка $\psi_3^*(G)$.

Доказательство. Тот факт, что $\psi_b^*(G) \geq \psi_c^*(G)$ следует из способа построения этих оценок: $\psi_c^*(G)$ соответствует задаче (14)–(20), которая содержит в себе задачу (14)–(19), которой соответствует оценка $\psi_b^*(G)$.

Для доказательства того, что $\psi_3^*(G) \geq \psi_b^*(G)$ покажем, что из квадратичной задачи (14)–(19) можно получить задачу (1)–(5), которой соответствует оценка $\psi_3^*(G)$. Учитывая, что задача (14)–(16) совпадает с задачей (1)–(3), достаточно показать, что из ограничений задачи (14)–(19) следуют семейства функционально избыточных ограничений (4) и (5).

Вначале покажем, как из ограничений (15)–(18) получаются функционально избыточные ограничения (4). Заметим, что для произвольной пары (i, j) справедливо неравенство

$$x_i + x_j - x_{ij} \geq 0. \quad (21)$$

На самом деле, из ограничений (17) и (18) следуют три равенства:

$$\begin{cases} 2x_i x_j - x_i - x_j + x_{ij} = 0, \\ 2x_i x_{ij} - x_i + x_j - x_{ij} = 0, \\ 2x_j x_{ij} + x_i - x_j - x_{ij} = 0, \end{cases} \quad (22)$$

которые будут справедливы для произвольной пары (i, j) . Если из первого равенства из (22) вычтеть второе и третье, то получаем равенство

$$2x_i x_j - 2x_i x_{ij} - 2x_j x_{ij} - x_i - x_j + 3x_{ij} = 0,$$

которое, учитывая что $x_{ij}^2 = x_{ij}$, $x_i^2 = x_i$, $x_j^2 = x_j$ перепишем так

$$x_i^2 + x_j^2 + x_{ij}^2 + 2x_i x_j - 2x_i x_{ij} - 2x_j x_{ij} - 2x_i - 2x_j + 2x_{ij} = 0.$$

Последнее можно записать в такой форме:

$$(x_i + x_j - x_{ij})^2 - 2(x_i + x_j - x_{ij}) = 0,$$

откуда имеем

$$(x_i + x_j - x_{ij}) = 0, \text{ либо } (x_i + x_j - x_{ij}) = 2,$$

что обеспечивает справедливость (21), так как

$$x_i + x_j - x_{ij} \geq \min\{0, 2\} = 0.$$

Пусть $(i, j) \notin E$. Тогда из первого уравнения из (22) имеем

$$2x_i x_j = x_i + x_j - x_{ij}, \quad (i, j) \notin E,$$

откуда с учетом (21), получаем неравенство

$$x_i x_j \geq 0, \quad (i, j) \notin E,$$

которое совпадает с функционально избыточным ограничением из (4). Это означает доказательство того, что $\psi_a^*(G) \leq \psi_2^*(G)$ поскольку из задачи (14)–(18), соответствующей оценке $\psi_a^*(G)$, получена задача (1)–(4), которая соответствует оценке $\psi_2^*(G)$.

Далее покажем, как из ограничений (15)–(19) следуют функционально избыточные ограничения (5). Пусть $(i, j) \in E$ и $k \neq (i, j)$. Из ограничений (17) следует, что для пар (i, k) и (j, k) имеют место равенства

$$2x_i x_k - x_i - x_k + x_{ik} = 0,$$

$$2x_j x_k - x_j - x_k + x_{jk} = 0,$$

а для пары $(i, j) \in E$ – равенство

$$-x_i - x_j + x_{ij} = 0,$$

учитывая, что согласно (16) $x_i x_j = 0$. Сложив два первых равенства и вычитая третье, получаем равенство

$$2x_i x_k + 2x_j x_k - 2x_k - x_{ij} + x_{ik} + x_{jk} = 0,$$

которое запишем в виде

$$x_i x_k + x_j x_k = x_k + \frac{1}{2}(x_{ij} - x_{ik} - x_{jk}). \quad (23)$$

Для оценки правой части в (23) воспользуемся тем, что из ограничений (19) имеем три равенства

$$\begin{cases} 2x_{ik} x_{jk} + x_{ij} - x_{ik} - x_{jk} = 0, \\ 2x_{ij} x_{jk} - x_{ij} + x_{ik} - x_{jk} = 0, \\ 2x_{ij} x_{ik} - x_{ij} - x_{ik} + x_{jk} = 0, \end{cases}$$

которые связывают между собой переменные x_{ij}, x_{ik}, x_{jk} . Вычтем из первого равенства второе и третье. В результате получаем

$$-2x_{ij}x_{ik} - 2x_{ij}x_{jk} + 2x_{ik}x_{jk} + 3x_{ij} - x_{ik} - x_{jk} = 0,$$

которое, с учетом того, что $x_{ij}^2 = x_{ij}$, $x_{ik}^2 = x_{ik}$ и $x_{jk}^2 = x_{jk}$, можно переписать следующим образом:

$$x_{ij}^2 + x_{ik}^2 + x_{jk}^2 - 2x_{ij}x_{jk} - 2x_{ij}x_{ik} + 2x_{ij}x_{jk} + 2(x_{ij} - x_{ik} - x_{jk}) = 0.$$

Последнее равенство можно записать

$$(x_{ij} - x_{ik} - x_{jk})^2 + 2(x_{ij} - x_{ik} - x_{jk}) = 0,$$

откуда имеем

$$x_{ij} - x_{ik} - x_{jk} \leq \max\{-2, 0\} = 0.$$

Учитывая последнее неравенство, из (23) следует, что для произвольной пары $(i, j) \in E$ и произвольного $k \neq i, j$ справедливо неравенство

$$x_i x_k + x_j x_k = x_k + \frac{1}{2}(x_{ij} - x_{ik} - x_{jk}) \leq x_k,$$

которое дает функционально избыточные ограничения (5). Теорема доказана.

Заключение. Из теоремы следует, что если граф G t -перфектный, либо h -перфектный, то $\psi_b^*(G) = \psi_c^*(G) = \alpha(G)$. Следовательно, обе оценки – точные верхние оценки для $\alpha(G)$ для t - и h -перфектных графов. На самом деле свойства оценок $\psi_b^*(G)$ и $\psi_c^*(G)$ более сильные, чем сформулированные в теореме. Так, например, оценка $\psi_b^*(G)$ улучшает число Ловаса $\vartheta'(G)$ для графа в форме икосаэдра I_{12} (содержит 12 вершин и 30 ребер). Для него $\psi_b^*(I_{12}) \approx 3.185116$, а $\vartheta'(I_{12}) = 1 + \sqrt{5} \approx 3.236068$. Заметим, что оценка $\psi_3^*(G)$ не улучшает $\vartheta'(I_{12})$. Оценка $\psi_c^*(G)$ является точной для I_{12} , т.е. $\psi_c^*(I_{12}) = \alpha(I_{12}) = 3$.

Литература

1. ГЭРИ М., ДЖОНСОН Д. Вычислительные машины и труднорешаемые задачи. – М.: Мир, 1982. – 416 с.
2. GROTSCHEL M., LOVASZ L., SCHRIJVER A. Geometric algorithms and combinatorial optimization. – Berlin: Springer-Verlag, 1988. – 362 p.

3. ШОР Н.З., СТЕЦЕНКО С.И. Квадратичные экстремальные задачи и недифференцируемая оптимизация. – Киев: Наук. думка, 1989. – 208 с.
4. SHOR N.Z. Nondifferentiable optimization and polynomial problems. – Dordrecht: Kluwer, 1998. – 394 p.
5. ШОР Н.З. Роль избыточных ограничений в улучшении двойственных оценок для полиномиальных оптимизационных задач // Кибернетика и системный анализ. 1998. – № 4. – С. 106–121.
6. СТЕЦЮК П.И. О функционально избыточных ограничениях для булевых оптимизационных задач квадратичного типа // Кибернетика и системный анализ. – 2005. – № 6. – С. 168–172.

Получено 24.07.2006

Об одной верхней оценке для взвешенного числа устойчивости графа

Стецюк П.И., Бутенко С.И., Березовский О.А.

*Теория оптимальных решений. – Киев: Ин-т кибернетики
им. В.М.Глушкова НАН Украины, 2007. – № 6. – С. 80–89.*

Для взвешенного числа устойчивости неориентированного графа G выведена верхняя оценка, которая является решением задачи линейного программирования с числом ограничений $O(|V|^3)$, где V – количество вершин в графе. Доказано, что полученная верхняя оценка не хуже, чем известная оценка, связанная с многогранником $CSTAB(G)$, а также является точной оценкой сверху для взвешенного числа устойчивости t -совершенного графа.

Введение. Пусть $G = (V, E)$ – взвешенный неориентированный граф с множеством вершин V и множеством ребер E ; вес каждой вершины $i \in V$ задан положительным целым числом w_i . Подмножество вершин $S \subseteq V$ называется устойчивым (или независимым) множеством графа G , если для любых $i, j \in S$ ребро (i, j) не принадлежит E . Взвешенное число устойчивости графа G определяется как $\alpha(G, w) = \max \sum_{i \in S} w_i$, где $S \subseteq V$ – устойчивое множество. Множество S^* , на котором достигается $\alpha(G, w)$, называется максимальным взвешенным устойчивым (или независимым) множеством.

В общем случае задача нахождения взвешенного числа устойчивости графа G NP-трудная. Доказано [1], что она является NP-трудной даже в частном случае, когда все веса вершин равны единице. Поэтому актуален поиск оценок, приближающих как взвешенное число устойчивости графа $\alpha(G, w)$, так и обычное число устойчивости графа $\alpha(G)$, когда все веса вершин графа равны единице. Величина $\alpha(G)$ характеризует мощность максимального по числу вершин устойчивого множества графа G .

Пусть χ^S – "инцидентный" вектор для подмножества $S \subseteq V$: $\chi^S \in R^{|V|}$, $\chi_i^S = 1$, если $i \in S$, $\chi_i^S = 0$, если $i \in V \setminus S$. Выпуклая оболочка "инцидентных" векторов χ^S для всех устойчивых множеств S графа G

$$STAB(G) := \text{conv}\{\chi^S : S \subseteq V \text{ – устойчивое множество}\}$$

называется многогранником устойчивых множеств (stable set polytope) [2]. Тогда задача нахождения взвешенного числа устойчивости графа G формулируется следующим образом:

$$\alpha(G, w) = \max \sum_{i \in V} w_i x_i, \quad x \in STAB(G), \quad (1)$$

В общем случае многогранник устойчивых множеств $STAB(G)$ может иметь очень сложную структуру [2], чем обусловлено то, что задача нахождения $\alpha(G, w)$ является NP -трудной. Вычисление "тесных" верхних оценок для $\alpha(G, w)$, которые достаточно хорошо аппроксимируют сверху значение целевой функции в задаче (1), имеет как практический, так и теоретический интерес.

Рассмотрим верхнюю оценку, связанную с многогранником

$$CSTAB(G) = \begin{cases} 0 \leq x_i \leq 1 & \text{для каждой вершины } i \in V, \\ x_i + x_j \leq 1 & \text{для каждого ребра } (i, j) \in E, \\ \sum_{i \in V(C_{2k+1})} x_i \leq k & \text{для каждого нечетного цикла } C_{2k+1} \in G, \end{cases}$$

который аппроксимирует (сверху) многогранник $STAB(G)$. Условимся эту оценку обозначать $\alpha_C^*(G, w)$. Для ее нахождения требуется решить задачу линейного программирования (ЛП-задачу):

$$\alpha_C^*(G, w) = \max \sum_{i \in V} w_i x_i, \quad x \in CSTAB(G), \quad (2)$$

которая в общем случае содержит неполиномиальное количество ограничений. Несмотря на это, задача (2) разрешима за полиномиальное время [2]. Полиномиальный алгоритм для нахождения $\alpha_C^*(G, w)$ базируется на использовании метода эллипсоидов и на том, что если точка \underline{x} удовлетворяет первым двум семействам линейных неравенств для многогранника $CSTAB(G)$, то за полиномиальное время можно либо убедиться, что точка \underline{x} удовлетворяет ограничениям для всех нечетных циклов, либо найти нечетный цикл, для которого ограничение будет максимально нарушаться. Для этого требуется найти $|V|$ кратчайших путей в специально построенном неориентированном графе G' , содержащем $2|V|$ вершин (вершины V и их копии V') и $2|E|$ ребер.

t -совершенным (t -perfect) графами называют семейство графов, для которых многогранник устойчивых множеств $STAB(G)$ совпадает с многогранником $CSTAB(G)$ ($STAB(G) = CSTAB(G)$). Для t -совершенных

графов оценка $\alpha_C^*(G, w)$ является точной оценкой для $\alpha(G, w)$, т.е. справедливо следующее равенство

$$\alpha_C^*(G, w) = \alpha(G, w). \quad (3)$$

Естественно, что когда граф G принадлежит семейству t -совершенных графов, то задача (1) полиномиально разрешима в том смысле, что $\alpha(G, w)$ может быть найдено с любой заданной точностью за полиномиальное время.

Многогранник $CSTAB(G)$ можно описать с помощью полиномиального количества линейных ограничений [3, с. 1187], используя переменные y_{ij} для каждого ребра $(i, j) \in E$ и переменные $z_{u,v}$ для каждой пары вершин $u, v \in V$. Область действия ограничений, связанных с нечетным циклом, эквивалентна следующей системе линейных ограничений:

$$\left\{ \begin{array}{ll} y_{ij} = 1 - x_i - x_j & \text{для каждого ребра } (i, j) \in E, \\ z_{v,v} \geq 1 & \text{для каждой вершины } v \in V, \\ z_{u,v} \leq y_{uv} & \text{для каждого ребра } (u, v) \in E, \\ z_{t,w} \leq z_{t,v} + y_{uv} + y_{vw} & \text{для всех } t, u, v, w \in V : (u, v), (v, w) \in E. \end{array} \right. \quad (4)$$

Таким образом, для t -совершенных графов $\alpha(G, w)$ может быть найдено за полиномиальное время с помощью любого из полиномиальных алгоритмов для задачи линейного программирования. Заметим, что результирующее количество ограничений для описания многогранника $CSTAB(G)$ с помощью (4), равно $O(|V|^2|E|)$.

Рассмотрим оценку $\alpha_\Delta^*(G, w)$, нахождению которой соответствует ЛП-задача с полиномиальным количеством ограничений порядка $O(|V|^3)$. Покажем, что оценка $\alpha_\Delta^*(G, w)$ не менее точная оценка сверху для $\alpha(G, w)$, чем оценка $\alpha_C^*(G, w)$. Естественно, что оценка $\alpha_\Delta^*(G, w)$ будет точной верхней оценкой для $\alpha(G, w)$, если граф G является t -совершенным.

Вывод оценки $\alpha_\Delta^*(G, w)$. ЛП-задача для оценки $\alpha_\Delta^*(G, w)$ следует из релаксации (ослабления) определенной формулировки квадратичной булевой задачи, с помощью которой формулируется проблема нахождения $\alpha(G, w)$. Построение этой ЛП-задачи опишем последовательностью этапов, которые будем комментировать необходимыми сведениями.

Этап 1. Рассмотрим задачу о максимальном взвешенном устойчивом множестве графа G в форме следующей квадратичной булевой задачи [4]

$$\alpha(G, w) = \max \sum_{i \in V} w_i x_i \quad (5)$$

при ограничениях

$$x_i x_j = 0 \quad \forall (i, j) \in E, \quad (6)$$

$$x_i^2 - x_i = 0 \quad \forall i \in V, \quad (7)$$

где булева переменная $x_i \in \{0, 1\}$ равна единице, если вершина i включается в устойчивое множество, и равна нулю – иначе. Булевы переменные описаны квадратичными ограничениями-равенствами (7). Квадратичные ограничения (6) означают, что если две вершины связаны ребром в графе G , то они обе не могут одновременно принадлежать независимому множеству.

Этап 2. Переформулируем задачу (5)–(7) с помощью бинарных (± 1)-переменных

$$y_i = 1 - 2x_i \quad \forall i \in V, \quad (8)$$

где вершине i соответствует бинарная переменная y_i , которая равна -1 , если вершина i включается в устойчивое множество, и равна 1 в противном случае. В результате получаем квадратичную задачу для бинарных переменных в следующей формулировке:

$$\alpha(G, w) = \max \sum_{i \in V} \left(\frac{w_i}{2} - \frac{w_i y_i}{2} \right) \quad (9)$$

при ограничениях

$$y_i y_j - y_i - y_j + 1 = 0 \quad \forall (i, j) \in E, \quad (10)$$

$$y_i^2 = 1 \quad \forall i \in V. \quad (11)$$

Этап 3. Добавим к задаче (9)–(11) семейство квадратичных ограничений:

$$\left\{ \begin{array}{l} +y_i y_j + y_i y_k + y_j y_k \geq -1, \\ -y_i y_j - y_i y_k + y_j y_k \geq -1, \\ -y_i y_j + y_i y_k - y_j y_k \geq -1, \\ +y_i y_j - y_i y_k - y_j y_k \geq -1, \end{array} \right. \quad \forall i, j, k \in V : i < j < k. \quad (12)$$

Ограничения (12) не изменяют множество оптимальных решений задачи (9)–(11) [5]. Они следуют из того, что для произвольной тройки бинарных (± 1)-переменных y_i, y_j и y_k , такой что $i \neq j \neq k$ всегда правильными являются квадратичные ограничения в виде неравенств

$$\left\{ \begin{array}{l} (+y_i + y_j + y_k)^2 \geq 1, \\ (-y_i + y_j + y_k)^2 \geq 1, \\ (+y_i - y_j + y_k)^2 \geq 1, \\ (+y_i + y_j - y_k)^2 \geq 1, \end{array} \right.$$

Последние легко преобразовываются к форме (12), учитывая, что $y_i^2=1$, $y_j^2=1$ и $y_k^2=1$.

Этап 4. Линеаризуем задачу (9)–(12), положив $y_{ij}=y_i y_j$ для каждой пары $i, j \in V$, таких что $i < j$ и релаксируя ограничения (11) линейными неравенствами $-1 \leq y_i \leq 1 \quad \forall i \in V$. В результате получаем ЛП-задачу:

$$\alpha_{\Delta}^*(G, w) = \max \sum_{i \in V} \left(\frac{w_i}{2} - \frac{w_i y_i}{2} \right) \quad (13)$$

при ограничениях

$$y_{ij} - y_i - y_j + 1 = 0 \quad \forall (i, j) \in E, \quad (14)$$

$$-1 \leq y_i \leq 1 \quad i \in V, \quad (15)$$

$$\begin{cases} +y_{ij} + y_{ik} + y_{jk} \geq -1, \\ -y_{ij} - y_{ik} + y_{jk} \geq -1, \\ -y_{ij} + y_{ik} - y_{jk} \geq -1, \\ +y_{ij} - y_{ik} - y_{jk} \geq -1, \end{cases} \quad \forall i, j, k \in V : i < j < k. \quad (16)$$

Учитывая, что ЛП-задача (13)–(16) получена „ослаблением“ квадратичной задачи (9)–(12), имеем

$$\alpha_{\Delta}^*(G, w) \geq \alpha(G, w). \quad (17)$$

Следовательно, оценка $\alpha_{\Delta}^*(G, w)$ есть оценкой сверху для взвешенного числа устойчивости $\alpha(G, w)$.

Этап 5. Переформулируем задачу (9)–(12) с помощью „релаксированных“ булевых переменных. Для этого „релаксированные“ бинарные переменные y_i , $i \in V$ переведем в „релаксированные“ булевые переменные с помощью замены (8), а „релаксированные“ бинарные переменные y_{ij} – в „релаксированные“ булевые переменные x_{ij} с помощью замены

$$y_{ij} = 1 - 2x_{ij} \quad \forall i, j \in V : i < j. \quad (18)$$

В результате получаем, что оценке $\alpha_{\Delta}^*(G)$ соответствует ЛП-задача

$$\alpha_{\Delta}^*(G, w) = \max \sum_{i \in V} w_i x_i \quad (19)$$

при ограничениях

$$x_i + x_j - x_{ij} = 0 \quad \forall (i, j) \in E, \quad (20)$$

$$\left\{ \begin{array}{l} +x_{ij} + x_{ik} + x_{jk} \leq 2, \\ -x_{ij} - x_{ik} + x_{jk} \leq 0, \\ -x_{ij} + x_{ik} - x_{jk} \leq 0, \\ +x_{ij} - x_{ik} - x_{jk} \leq 0, \end{array} \right. \quad \forall i, j, k \in V : i < j < k \quad (21)$$

$$0 \leq x_i \leq 1 \quad \forall i \in V. \quad (22)$$

ЛП-задача (19)–(22) содержит $O(|V|^3)$ ограничений. Точное количество ограничений равно $|E| + 2|V| + 2|V|(|V| - 1)(|V| - 2)/3$. Из них самую значительную часть, равную $2|V|(|V| - 1)(|V| - 2)/3$, определяют ограничения (21), которые (а также их форма в бинарных переменных), известны как ограничения треугольника (triangle inequalities). Их наличие придает ЛП-задаче (19)–(22) интересные геометрические свойства, связанные с нечетным циклом C_{2k+1} в графе G .

Свойства оценки $\alpha_{\Delta}^*(G, w)$. Справедлива следующая теорема.

Теорема 1. Из ограничений (20) и (21) следует справедливость линейных неравенств

$$\sum_{i \in V(C_{2k+1})} x_i \leq k \quad \text{для каждого нечетного цикла } C_{2k+1} \in G. \quad (23)$$

Доказательство. Рассмотрим произвольный нечетный цикл C_{2k+1} с вершинами $\{i_1, \dots, i_{2k+1}\}$ и ребрами (i_r, i_{r+1}) , $r=1, \dots, 2k$, (i_1, i_{2k+1}) . Не ограничивая общности, будем считать, что $i_1 < i_2 < \dots < i_{2k+1}$. Для цикла C_{2k+1} ограничения (20) имеют следующий вид

$$x_{i_r} + x_{i_{r+1}} = x_{i_r i_{r+1}}, \quad r = 1, 2, \dots, 2k, \quad x_{i_1} + x_{i_{2k+1}} = x_{i_1 i_{2k+1}},$$

и, сложив их, получим

$$2 \sum_{r=1}^{2k+1} x_{i_r} = x_{i_1 i_{2k+1}} + \sum_{r=1}^{2k} x_{i_r i_{r+1}}. \quad (24)$$

Оценим правую часть в равенстве (24).

Если $k = 1$ (соответствует нечетному циклу C_3), то для тройки вершин (i_1, i_2, i_3) из первого неравенства в (21) следует справедливость неравенства:

$$x_{i_1 i_2} + x_{i_1 i_3} + x_{i_2 i_3} \leq 2.$$

Тогда из равенства (24) имеем

$$2 \sum_{r=1}^3 x_{i_r} = x_{i_1 i_3} + \sum_{r=1}^2 x_{i_r i_{r+1}} = x_{i_1 i_2} + x_{i_1 i_3} + x_{i_2 i_3} \leq 2,$$

откуда следует

$$\sum_{r=1}^3 x_{i_r} \leq 1 \quad \text{или} \quad \sum_{i \in V(C_3)} x_i \leq 1,$$

что дает доказательство теоремы для нечетного цикла C_3 .

Пусть k – произвольное натуральное число, такое что $k \geq 2$. Рассмотрим „покрытие“ вершин нечетного цикла C_{2k+1} двумя типами треугольников (тройками вершин), которое для нечетного цикла C_9 показано на рис. 1. Для первого типа треугольников („незаштрихованные“)

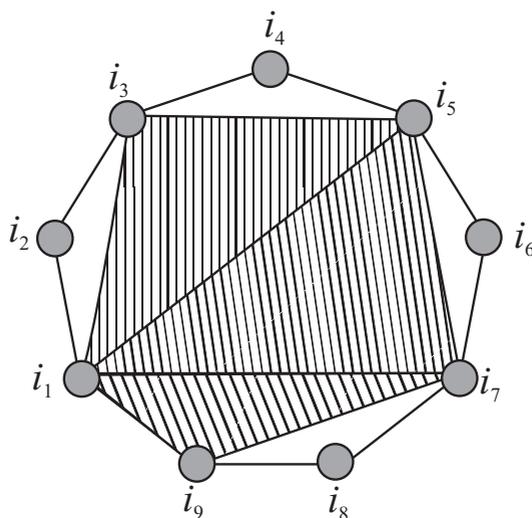


Рис. 1. Покрытие треугольниками вершин нечетного цикла C_9

будем использовать первое неравенство из (21), а для второго типа треугольников („заштрихованные“) – третье неравенство из (21).

Первый тип треугольников содержит следующие тройки вершин: $(i_1, i_2, i_3), \dots, (i_{2k-1}, i_{2k}, i_{2k+1})$. Количество таких троек равно k , и из первого неравенства в (21) для них следует справедливость неравенств

$$x_{i_{2t-1}i_{2t}} + x_{i_{2t-1}i_{2t+1}} + x_{i_{2t}i_{2t+1}} \leq 2, \quad t = 1, \dots, k.$$

Сложив их, получаем

$$\sum_{t=1}^k x_{i_{2t-1}i_{2t}} + \sum_{t=1}^k x_{i_{2t-1}i_{2t+1}} + \sum_{t=1}^k x_{i_{2t}i_{2t+1}} \leq 2k.$$

Затем запишем в виде:

$$x_{i_1 i_3} + \sum_{t=2}^k x_{i_{2t-1} i_{2t+1}} + \sum_{t=1}^k x_{i_{2t-1} i_{2t}} + \sum_{t=1}^k x_{i_{2t} i_{2t+1}} \leq 2k. \quad (25)$$

Второй тип треугольников содержит следующие тройки вершин $(i_1, i_3, i_5), \dots, (i_1, i_{2k-1}, i_{2k+1})$. Количество таких троек равно $(k-1)$, и из третьего неравенства в (21) для них следует справедливость неравенств

$$-x_{i_1 i_{2t-1}} + x_{i_1 i_{2t+1}} - x_{i_{2t-1} i_{2t+1}} \leq 0, \quad t = 2, \dots, k,$$

сложив которые, получаем неравенство

$$-x_{i_1 i_3} + x_{i_1 i_{2k+1}} - \sum_{t=2}^k x_{i_{2t-1} i_{2t+1}} \leq 0.$$

Сложив последнее неравенство с неравенством (25) получаем неравенство

$$x_{i_1 i_{2k+1}} + \sum_{t=1}^k x_{i_{2t-1} i_{2t}} + \sum_{t=1}^k x_{i_{2t} i_{2t+1}} \leq 2k. \quad (26)$$

Учитывая, что

$$\sum_{t=1}^k x_{i_{2t} i_{2t+1}} + \sum_{t=1}^k x_{i_{2t-1} i_{2t}} = \sum_{t=1}^k (x_{i_{2t-1} i_{2t}} + x_{i_{2t} i_{2t+1}}) = \sum_{r=1}^{2k} x_{i_r i_{r+1}},$$

из неравенства (26) имеем

$$x_{i_1 i_{2k+1}} + \sum_{r=1}^{2k} x_{i_r i_{r+1}} \leq 2k. \quad (27)$$

Далее, подставляя (27) в (24), имеем

$$2 \sum_{r=1}^{2k+1} x_{i_r} = x_{i_1 i_{2k+1}} + \sum_{r=1}^{2k} x_{i_r i_{r+1}} \leq 2k,$$

откуда следует неравенство

$$\sum_{r=1}^{2k+1} x_{i_r} \leq k \quad \text{или} \quad \sum_{i \in V(C_{2k+1})} x_i \leq k,$$

что завершает доказательство теоремы. \square

Перейдем к исследованию свойств оценки $\alpha_{\Delta}^*(G, w)$. Справедлива следующая теорема.

Теорема 2. Для произвольного графа G оценка $\alpha_{\Delta}^*(G, w)$ удовлетворяет соотношению:

$$\alpha_C^*(G, w) \geq \alpha_{\Delta}^*(G, w) \geq \alpha(G, w).$$

Доказательство. Из определения оценки $\alpha_{\Delta}^*(G, w)$ следует что, что $\alpha_{\Delta}^*(G, w) \geq \alpha(G, w)$ (см. формулу (17)). Докажем, что $\alpha_C^*(G, w) \geq \alpha_{\Delta}^*(G, w)$. От задачи (19)–(22) легко перейти к ЛП задаче:

$$\alpha_{\Delta}^*(G, w) = \max \sum_{i \in V} w_i x_i \quad (28)$$

при ограничениях

$$x_i + x_j - x_{ij} = 0 \quad \forall (i, j) \in E, \quad (29)$$

$$\begin{cases} +x_{ij} + x_{ik} + x_{jk} \leq 2, \\ -x_{ij} - x_{ik} + x_{jk} \leq 0, \\ -x_{ij} + x_{ik} - x_{jk} \leq 0, \\ +x_{ij} - x_{ik} - x_{jk} \leq 0, \end{cases} \quad \forall i, j, k \in V : i < j < k. \quad (30)$$

$$0 \leq x_i \leq 1 \quad \forall i \in V, \quad (31)$$

$$x_{ij} \leq 1 \quad \forall (i, j) \in E, \quad (32)$$

$$x_i + x_j \leq 1 \quad \forall (i, j) \in E. \quad (33)$$

$$\sum_{i \in V(C_{2k+1})} x_i \leq k \quad \forall C_{2k+1} \in G, \quad (34)$$

Здесь дополнительные ограничения (32) получены из ограничений (30). Так, например, сложив первое и последнее из неравенств (30), имеем $2x_{ij} \leq 2$, откуда следует справедливость неравенств (32). Ограничения (33) получены в результате сложения ограничений (29) с ограничениями (32). Ограничения (34) связаны с нечетными циклами в графе G и следуют из теоремы 1.

Задача (28)–(34) включает избыточные линейные ограничения (32)–(34), которые следуют из ограничений (29)–(31). От задачи (28)–(34) перейдем к „ослабленной“ ЛП-задаче, убрав из нее все ограничения за исключением ограничений (31), (33) и (34). В результате получаем ЛП-задачу

$$\alpha^*(G, w) = \max \sum_{i \in V} w_i x_i \quad (35)$$

при ограничениях

$$0 \leq x_i \leq 1 \quad \forall i \in V, \quad (36)$$

$$x_i + x_j \leq 1 \quad \forall (i, j) \in E. \quad (37)$$

$$\sum_{i \in V(C_{2k+1})} x_i \leq k \quad \forall C_{2k+1} \in G, \quad (38)$$

для которой $\alpha^*(G, w) \geq \alpha_{\Delta}^*(G, w)$. Задача (35)–(38) является ни чем иным, как формулировкой задачи (2) для нахождения $\alpha_C^*(G, w)$, так как ограничения (36)–(38) – определение многогранника $CSTAB(G)$, т.е. $\alpha^*(G, w) = \alpha_C^*(G, w)$. Отсюда имеем $\alpha_C^*(G, w) \geq \alpha_{\Delta}^*(G, w)$. Теорема доказана. \square

Из теоремы 2 и равенства (3), которое означает что для t -совершенных графов оценка $\alpha_C^*(G, w)$ является точной оценкой сверху для $\alpha(G, w)$, следует справедливость теоремы.

Теорема 3. *Если граф G – t -совершенный, то $\alpha_{\Delta}^*(G, w) = \alpha(G, w)$.*

Следовательно, для t -совершенных графов оценка $\alpha_{\Delta}^*(G, w)$ – точная верхняя оценка для взвешенного числа устойчивости $\alpha(G, w)$.

Литература

1. ГЭРИ М., ДЖОНСОН Д. Вычислительные машины и труднорешаемые задачи. – М.: Мир, 1982. – 416 с.
2. GRÖTSCHEL M., LOVASZ L., SCHRIJVER A. Geometric algorithms and combinatorial optimization. – Berlin: Springer-Verlag, 1988. – 362 p.
3. SCHRIJVER A. Combinatorial Optimization: Polyhedra and Efficiency. – Berlin: Springer-Verlag, 2003. – 1881 p.
4. SHOR N.Z. Nondifferentiable optimization and polynomial problems. – Boston; Dordrecht; London: Kluwer Academic Publishers, 1998. – 412 p.
5. Стецюк П.И. Новые модели квадратичного типа для задачи о максимальном взвешенном разрезе графа // Кибернетика и системный анализ. – 2006. – № 1. – С. 63-75.

Получено 17.04.2007

Точная ЛП-оценка для взвешенного числа устойчивости t -совершенных графов

П. И. Стецюк, В. И. Ляшко, Е. А. Нурминский

Журнал вычислительной и прикладной математики. — 2009. — № (99). — С. 106–115.

Для взвешенного числа устойчивости неориентированного графа $G = (V, E)$ построена верхняя ЛП-оценка, которая равна оптимальному значению целевой функции в задаче линейного программирования с числом переменных $|V| + |\bar{E}|$ и числом ограничений $O(|V|^3)$, где V – количество вершин в графе, а $|\bar{E}|$ – количество ребер в графе, дополнительном к графу G . Доказано, что полученная ЛП-оценка является не менее точной оценкой сверху для взвешенного числа устойчивости произвольного графа, чем известная верхняя оценка, связанная с многогранником нечетных циклов. Следствием этого факта есть то, что ЛП-оценка является точной для взвешенного числа устойчивости t -совершенных графов.

Введение

Задача о вычислении взвешенного числа устойчивости графа и задача об определении взвешенного наибольшего устойчивого множества вершин графа относятся к классу NP-трудных задач в области теории графов. Обе они связаны с неориентированным графом, вершинам которого приписаны целые веса, однако это две принципиально разные задачи и решение первой не слишком помогает при попытках решить вторую.

Для случая, когда все веса вершин графа равны единице, сами задачи и различие между ними можно продемонстрировать на примере известной проблемы восьми ферзей, которую связывают с именем К. Гаусса. Эта проблема состоит в нахождении наибольшего числа ферзей, которые можно расставить на шахматной доске так, чтобы они не атаковали друг друга. Очевидно, что таких ферзей не может быть более восьми, т.к. никакие два из них не должны находиться на одной вертикали или горизонтали.

С этой задачей можно сопоставить неориентированный граф, вершинам которого соответствуют клетки шахматной доски, а ребрами соединены пары клеток, лежащие на одной вертикали, горизонтали или диагонали. Расстановке ферзей, не бьющих друг друга, будет соответствовать выбор такого подмножества вершин-клеток, в котором ни одна из вершин не соединена с другой. Именно такое подмножество и называется устойчивым.

В графе для задачи о восьми ферзях можно довольно легко найти одну из возможных расстановок (наибольших устойчивых множеств) и поскольку большее число ферзей не может быть расставлено, то число устойчивости этого графа равно 8. Вместе с тем существуют 92 такие расстановки [1], с. 43-44 и понятно, что указать все соответствующие им наибольшие устойчивые множества существенно труднее.

Данная работа связана с исследованием верхней границы для взвешенного числа устойчивости неориентированного графа, которое является обобщением числа устойчивости. Эта граница названа ЛП-оценкой и она равна оптимальному значению целевой функции в специально сконструированной задаче линейного программирования с полиномиальным количеством линейных ограничений. Она может быть перенесена на случай взвешенного кликового числа, не менее известной NP-трудной задаче в теории графов. Кликой в неориентированном графе называется подмножество вершин, каждые две из которых соединены ребром. Обе задачи по сути эквивалентны: каждая из них получается из другой, путем построения дополнения графа – такого графа, в котором есть все вершины исходного графа, причем в дополнении графа вершины соединены ребром тогда и только тогда, если они не были соединены в исходном графе. Устойчивое (независимое) множество вершин в исходном графе является кликой в дополнительном графе к исходному, и наоборот.

1. Взвешенное число устойчивости $\alpha(G, w)$

Пусть $G=(V, E)$ – взвешенный неориентированный граф (не содержащий петель) с множеством вершин V и множеством ребер E , вес каждой вершины $i \in V$ задан положительным целым числом w_i . Основным объектом исследований в данной работе является так называемое устойчивое (или независимое) множество графа G .

Определение 1. *Подмножество вершин $S \subseteq V$ называется устойчивым (или независимым) множеством графа G , если для любых $i, j \in S$*

ребро $e = (i, j)$ не принадлежит E .

Взвешенное число устойчивости графа G определяется как

$$\alpha(G, w) = \max \sum_{i \in S} w_i,$$

где $S \subseteq V$ – устойчивое множество. Подмножество S^* , на котором достигается $\alpha(G, w)$, называется максимальным взвешенным устойчивым (или независимым) множеством графа G . В частном случае, когда все веса вершин в графе равны единице, оно совпадает с обычным числом устойчивости графа G , которое принято обозначать $\alpha(G)$. Число устойчивости $\alpha(G)$ характеризует мощность максимального по числу входящих в него вершин устойчивого множества в графе G . В общем случае задача нахождения $\alpha(G, w)$ принадлежит к NP -трудным задачам [2].

Пусть $STAB(G)$ – многогранник устойчивых множеств (stable polytope), определяемый как выпуклая оболочка булевых индикаторных векторов устойчивых множеств S в графе G :

$$STAB(G) = \text{conv}\{x^S, S \text{ — устойчивое множество в графе } G.\} \quad (1.1)$$

Индикаторный вектор множества S определяется как $x^S = (x_i^S, i \in V) \in \{0, 1\}^{|V|}$, где

$$x_i^S = \begin{cases} 1 & \text{если } i \in S; \\ 0 & \text{в противном случае.} \end{cases}$$

Условие устойчивости множества S может быть представлено в виде системы квадратичных равенств для компонент вектора x^S :

$$x_i^S x_j^S = 0 \text{ для всех } (i, j) = e \in E.$$

Квадратичные равенства означают, что если две вершины связаны ребром в графе G , то они обе не могут одновременно принадлежать устойчивому множеству S .

Нахождение $\alpha(G, w)$ связано с задачей максимизации линейной функции $\sum_{i \in V} w_i x_i = wx$ на выпуклом многограннике $STAB(G)$:

$$\alpha(G, w) = \max_{x \in STAB(G)} wx = \max_S wx^S, \quad (1.2)$$

где последний максимум берется по всем устойчивым подмножествам V . Максимум линейной функции в задаче (1.2) достигается в одной или нескольких из вершин многогранника $STAB(G)$. В общем случае

многогранник $STAB(G)$ имеет сложную структуру, из-за чего задача нахождения $\alpha(G, w)$ принадлежит к NP -трудным задачам.

Однако, существует много семейств графов, для которых $\alpha(G, w)$ может быть найдено за полиномиальное время [2, 3]. Одним из таких семейств графов являются t -совершенные графы. Их название происходит от французского слова "trou", которое в переводе на русский язык означает "дыра". Само это название выделяет нечетный цикл в графе G (содержит нечетное количество вершин), который является центральным при определении t -совершенных графов. Нечетный цикл в графе G далее будем обозначать C_{2k+1} , $k = 1, 2, \dots$, а соответствующее ему подмножество вершин графа через $V(C_{2k+1})$.

Полиномиальный алгоритм нахождения $\alpha(G, w)$ для t -совершенных графов в значительной мере обязан такой задаче линейного программирования (ЛП-задаче)

$$\alpha_C^*(G, w) = \max \sum_{i \in V} w_i x_i, \quad (1.3)$$

при ограничениях

$$\left\{ \begin{array}{ll} 0 \leq x_i \leq 1 & \text{для каждой вершины } i \in V, \\ x_i + x_j \leq 1 & \text{для каждого ребра } (i, j) = e \in E, \\ \sum_{i \in V(C_{2k+1})} x_i \leq k & \text{для каждого нечетного цикла } C_{2k+1} \in G. \end{array} \right. \quad (1.4)$$

Многогранник, заданный ограничениями (1.4), называется многогранником нечетных циклов (odd-cycle polytope), и его принято обозначать $CSTAB(G)$. Он определяется справедливыми для многогранника $STAB(G)$ семействами линейных неравенств для вершин, линейных неравенств для ребер и линейных неравенств для нечетных циклов в графе G .

Для произвольного графа G величина $\alpha_C^*(G, w)$ удовлетворяет условию

$$\alpha_C^*(G, w) \geq \alpha(G, w) \quad (1.5)$$

и является оценкой сверху (верхней оценкой) для $\alpha(G, w)$. Неравенство (1.5) следует из того, что многогранник $CSTAB(G)$ аппроксимирует (сверху) многогранник $STAB(G)$. Семейство графов, для которых многогранник устойчивых множеств $STAB(G)$ совпадает с многогранником $CSTAB(G)$, называют t -совершенными графами. Учитывая, что $STAB(G) = CSTAB(G)$, для t -совершенных графов справедливо равенство

$$\alpha_C^*(G, w) = \alpha(G, w) \quad (1.6)$$

и величина $\alpha_C^*(G, w)$ есть точной верхней оценкой для $\alpha(G, w)$.

В общем случае ЛП-задача (1.3)–(1.4) содержит неполиномиальное количество ограничений, что связано с ограничениями в форме неравенств для всех нечетных циклов C_{2k+1} ($k = 1, 2, \dots$) в графе G . Несмотря на это ЛП-задача (1.3)–(1.4) полиномиально разрешима в том смысле, что для произвольного графа G оптимальное значение целевой функции $\alpha_C^*(G, w)$ может быть найдено с любой заданной точностью за полиномиальное время. Алгоритм на основе метода эллипсоидов дан в [2], с. 275–276. Если граф G принадлежит семейству t -совершенных графов, то равенство (1.6) обеспечивает нахождение $\alpha(G, w)$ за полиномиальное время.

Полиномиальный алгоритм для нахождения верхней оценки $\alpha_C^*(G, w)$ [2] базируется на использовании метода эллипсоидов. Он имеет скорее теоретическую ценность, чем практическую, и неприменим для нахождения оценки $\alpha_C^*(G, w)$ в том случае, когда граф содержит около сотни вершин. В то же время практически эффективные алгоритмы нахождения $\alpha_C^*(G, w)$ для графов такого размера и больше могут быть реализованы на основе современных ЛП-солверов, для которых решение ЛП-задач с сотнями тысяч переменных и миллионами ограничений не представляет особых проблем для современных компьютеров с оперативной памятью в несколько десятков гигабайт. Чтобы обеспечить полиномиальность таких алгоритмов, достаточно ограничить размеры ЛП-задачи вида (1.3)–(1.4) по числу ограничений, и в качестве ЛП-солвера использовать такой, который позволяет решить ЛП-задачу за полиномиальное время.

Для верхней оценки $\alpha_C^*(G, w)$ существует аналог ЛП-задачи с полиномиальным количеством линейных ограничений [3], стр. 1187. Такая ЛП-задача содержит $O(|V|^2)$ переменных и $O(|V|^2|E|)$ ограничений. В [4] построена верхняя оценка $\alpha_\Delta^*(G, w)$, которая по точности не хуже чем $\alpha_C^*(G, w)$, при этом ей соответствует ЛП-задача с меньшим количеством переменных и ограничений. Она включает $O(|V|^3)$ ограничений и $|V| + |\bar{E}|$ переменных, где $|\bar{E}|$ – количество ребер в дополнительном графе к графу G . Предметом обсуждения ниже будет еще одна верхняя оценка подобного типа, назовем ее ЛП-оценкой $\alpha_\nabla^*(G, w)$, которая получена линейной релаксацией квадратичных ограничений, связывающих между собой тройки булевых переменных.

2. ЛП-оценка $\alpha_{\nabla}^*(G, w)$

Приведем ЛП-оценку $\alpha_{\nabla}^*(G, w)$ и покажем, что она является решением ЛП-задачи с числом переменных $(|V| + |\bar{E}|)$ и числом ограничений $O(|V|^3)$. Нахождению ЛП-оценки $\alpha_{\nabla}^*(G, w)$ соответствует такая ЛП-задача: найти

$$\alpha_{\nabla}^*(G, w) = \max \sum_{i \in V} w_i x_i \quad (2.1)$$

при ограничениях

$$x_{ij} = 0 \quad \forall (i, j) = e \in E, \quad (2.2)$$

$$\begin{cases} -x_{ij} - x_{ik} - x_{jk} + x_i + x_j + x_k \leq 1, \\ +x_{ij} + x_{ik} - x_{jk} - x_i \leq 0, \\ +x_{ij} - x_{ik} + x_{jk} - x_j \leq 0, \\ -x_{ij} + x_{ik} + x_{jk} - x_k \leq 0, \end{cases} \quad \forall i, j, k \in V : i < j < k \quad (2.3)$$

$$0 \leq x_i \leq 1 \quad i \in V. \quad (2.4)$$

Справедлива следующая лемма.

Лемма 1. Для произвольного графа G ЛП-оценка $\alpha_{\nabla}^*(G, w)$ удовлетворяет соотношению:

$$\alpha_{\nabla}^*(G, w) \geq \alpha(G, w).$$

Доказательство. За основу возьмем квадратичную булеву формулировку задачи о максимальном взвешенном устойчивом множестве [5]:

$$\alpha(G, w) = \max \sum_{i \in V} w_i x_i \quad (2.5)$$

при ограничениях

$$x_i x_j = 0 \quad \forall (i, j) = e \in E, \quad (2.6)$$

$$x_i^2 - x_i = 0 \quad \forall i \in V. \quad (2.7)$$

Квадратичная задача (2.5)–(2.7) фактически есть другой формой записи задачи (1.2), для которой булевы переменные $x_i \in \{0, 1\}$ для всех вершин из V описаны квадратичными ограничениями-равенствами (2.7).

Добавим к задаче (2.5)–(2.7) следующее семейство квадратичных ограничений

$$\begin{cases} +x_i x_j + x_i x_k + x_j x_k - x_i - x_j - x_k \geq -1, \\ -x_i x_j - x_i x_k + x_j x_k + x_i \geq 0, \\ -x_i x_j + x_i x_k - x_j x_k + x_j \geq 0, \\ +x_i x_j - x_i x_k - x_j x_k + x_k \geq 0, \end{cases} \quad \forall i, j, k \in V : i < j < k. \quad (2.8)$$

Квадратичные ограничения (2.8) не изменяют множества допустимых решений задачи (2.5)–(2.7). На самом деле, они являются следствием квадратичных неравенств

$$\begin{cases} +y_i y_j + y_i y_k + y_j y_k \geq -1, \\ -y_i y_j - y_i y_k + y_j y_k \geq -1, \\ -y_i y_j + y_i y_k - y_j y_k \geq -1, \\ +y_i y_j - y_i y_k - y_j y_k \geq -1, \end{cases} \quad \forall i, j, k \in V : i < j < k. \quad (2.9)$$

для бинарных (± 1) -переменных y_i , $i \in V$, если бинарные переменные перевести в булевы $(0-1)$ -переменные x_i , $i \in V$ с помощью следующей замены

$$y_i = 1 - 2x_i \quad \forall i \in V.$$

Квадратичные ограничения (2.9) следуют из того, что для произвольной тройки бинарных (± 1) -переменных y_i , y_j и y_k , такой что $i \neq j \neq k$, всегда справедливы квадратичные ограничения в форме неравенств [6]

$$\begin{cases} (+y_i + y_j + y_k)^2 \geq 1, \\ (-y_i + y_j + y_k)^2 \geq 1, \\ (+y_i - y_j + y_k)^2 \geq 1, \\ (+y_i + y_j - y_k)^2 \geq 1, \end{cases}$$

Последние легко преобразуются к неравенствам (2.9), учитывая, что $y_i^2=1$, $y_j^2=1$ и $y_k^2=1$.

Линеаризуем задачу (2.5)–(2.8), положив $x_{ij}=x_i x_j \quad \forall i, j \in V : i < j$ и релаксируя ограничения (2.7) линейными неравенствами $0 \leq x_i \leq 1 \quad \forall i \in V$. Оптимальное значение целевой функции в релаксированной задаче обозначим $\alpha_{\nabla}^*(G, w)$. В результате получаем ЛП-задачу (2.1)–(2.4) для величины $\alpha_{\nabla}^*(G, w)$, которая и названа ЛП-оценкой $\alpha_{\nabla}^*(G, w)$. Учитывая, что ЛП-задача (2.1)–(2.4) получена в результате "ослабления" квадратичной задачи (2.5)–(2.8), то ЛП-оценка $\alpha_{\nabla}^*(G, w)$ для произвольного графа G удовлетворяет соотношению:

$$\alpha_{\nabla}^*(G, w) \geq \alpha(G, w). \quad (2.10)$$

Следовательно, ЛП-оценка $\alpha_{\nabla}^*(G, w)$ будет оценкой сверху для $\alpha(G, w)$. Соотношение (2.10) доказывает лемму 1.

ЛП-задача (2.1)–(2.4) содержит $|V|(|V|+1)/2$ переменных и $O(|V|^3)$ ограничений. Однако, число переменных в ней можно уменьшить. На самом деле $|E|$ линейных переменных, которые соответствуют ребрам графа G и заданы ограничениями (2.2), фиксированы и равны нулю.

Это означает, что в ЛП-задаче (2.1)–(2.4) можно избавиться от ограничений (2.2), подставляя нулевые значения для тех переменных, которые они определяют, непосредственно в ограничения (2.3). В результате в новой ЛП-задаче останутся только переменные x_{ij} для тех пар (i, j) , которые не принадлежат множеству ребер E . Все эти пары определяют множество ребер \bar{E} в дополнительном графе к графу G . Число переменных в новой ЛП-задаче будет равным $|V| + |\bar{E}|$, где $|\bar{E}|$ – количество ребер в графе, дополнительном к графу G . При этом порядок ограничений в новой задаче останется равным $O(|V|^3)$, но реальное число ограничений может уменьшиться. Так, например, если структура графа G такова, что существуют такие тройки вершин i, j и k , для которых все три пары (i, j) , (i, k) и (j, k) принадлежат множеству ребер E , то линейные переменные x_{ij} , x_{ik} и x_{jk} равны нулю из ограничений (2.2). Тогда, четыре ограничения в (2.3) вырождаются в единственное ограничение

$$x_i + x_j + x_k \leq 1,$$

так как остальные три ограничения связаны с неотрицательностью переменных x_i , x_j и x_k и дублируются ограничениями (2.4).

ЛП-задача, где отсутствуют ограничения в форме (2.2), будет более экономной для вычисления ЛП-оценки $\alpha_{\nabla}^*(G, w)$ с помощью стандартных ЛП-программ. Однако, мы далее будем придерживаться ее формы записи как ЛП-задачи (2.1)–(2.4), в основном, из-за удобства всех дальнейших доказательств, где будем в явном виде использовать факт равенства нулю тех или иных переменных вида x_{ij} , которые следуют из ограничений (2.2).

3. ЛП-оценка $\alpha_{\nabla}^*(G, w)$ и нечетный цикл

Наличие ограничений (2.3) придает ЛП-задаче (2.1)–(2.4) интересные "геометрические" свойства, связанные с нечетным циклом C_{2k+1} в графе G . Справедлива следующая лемма.

Лемма 2. *Из ограничений (2.2) и (2.3) следует справедливость линейных неравенств*

$$\sum_{i \in V(C_{2k+1})} x_i \leq k \quad \text{для каждого нечетного цикла } C_{2k+1} \in G.$$

Доказательство. Рассмотрим произвольный нечетный цикл C_{2k+1} с вершинами $\{i_1, \dots, i_{2k+1}\}$ и ребрами (i_r, i_{r+1}) , $r=1, \dots, 2k$, (i_1, i_{2k+1}) . Не ограничивая общности, будем считать что $i_1 < i_2 < \dots < i_{2k+1}$.

Если $k = 1$ (соответствует нечетному циклу C_3), то для тройки вершин (i_1, i_2, i_3) из первого неравенства из (2.3) имеем:

$$x_{i_1} + x_{i_2} + x_{i_3} \leq 1 + x_{i_1 i_2} + x_{i_1 i_3} + x_{i_2 i_3}.$$

Учитывая, что для ребер (i_1, i_2) , (i_1, i_3) и (i_2, i_3) из равенств (2.2) имеем $x_{i_1 i_2} = 0$, $x_{i_1 i_3} = 0$ и $x_{i_2 i_3} = 0$, то из приведенного выше равенства следует

$$\sum_{r=1}^3 x_{i_r} \leq 1 \quad \text{или} \quad \sum_{i \in V(C_3)} x_i \leq 1,$$

что дает доказательство леммы для нечетного цикла C_3 .

Пусть k – произвольное натуральное число, такое что $k \geq 2$. Рассмотрим "покрытие" нечетного цикла C_{2k+1} двумя типами треугольников (тройками вершин), которое для нечетного цикла C_9 проиллюстрировано на рис. 1. Для первого типа треугольников ("незаштрихованные") бу-

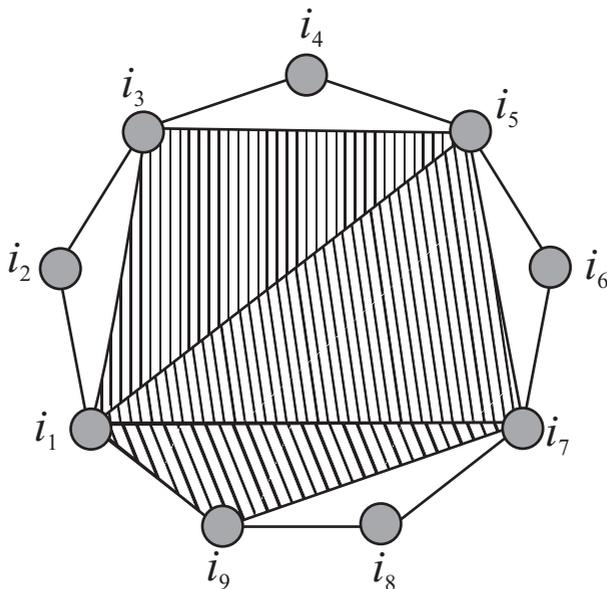


Рис. 1. Покрытие нечетного цикла C_9 треугольниками

дем использовать первое неравенство из (2.3), а для второго типа треугольников ("заштрихованные") будем использовать третье неравенство из (2.3).

Первый тип треугольников содержит тройки вершин $(i_{2t-1}, i_{2t}, i_{2t+1})$ для всех $t = 1, \dots, k$ и для каждой пары (i_{2t-1}, i_{2t}) и (i_{2t}, i_{2t+1}) из равенств (2.2) имеем $x_{i_{2t-1}i_{2t}} = 0$ и $x_{i_{2t}i_{2t+1}} = 0$. Для треугольников этого типа из первого неравенства в (2.3) следует справедливость таких неравенств

$$x_{i_{2t-1}} + x_{i_{2t}} + x_{i_{2t+1}} - x_{i_{2t-1}i_{2t+1}} \leq 1, \quad t = 1, \dots, k,$$

сложив которые получаем неравенство

$$\sum_{t=1}^k x_{i_{2t-1}} + \sum_{t=1}^k x_{i_{2t}} + \sum_{t=1}^k x_{i_{2t+1}} - \sum_{t=1}^k x_{i_{2t-1}i_{2t+1}} \leq k.$$

Учитывая, что

$$\sum_{t=1}^k x_{i_{2t-1}} + \sum_{t=1}^k x_{i_{2t}} = \sum_{t=1}^k (x_{i_{2t-1}} + x_{i_{2t}}) = \sum_{r=1}^{2k} x_{i_r},$$

последнее неравенство запишем в такой форме

$$\sum_{r=1}^{2k} x_{i_r} + \sum_{t=1}^k x_{i_{2t+1}} - \sum_{t=1}^k x_{i_{2t-1}i_{2t+1}} \leq k. \quad (3.1)$$

Второй тип треугольников ("заштрихованные") составляют такие тройки вершин $(i_1, i_3, i_5), \dots, (i_1, i_{2k-1}, i_{2k+1})$. Количество таких троек равно $(k-1)$. Из третьего неравенства в (2.3) для них следует справедливость таких неравенств

$$x_{i_1i_{2t-1}} - x_{i_1i_{2t+1}} + x_{i_{2t-1}i_{2t+1}} - x_{i_{2t-1}} \leq 0, \quad t = 2, \dots, k,$$

сложив которые получаем неравенство

$$x_{i_1i_3} - x_{i_1i_{2k+1}} + \sum_{t=2}^k x_{i_{2t-1}i_{2t+1}} - \sum_{t=2}^k x_{i_{2t-1}} \leq 0.$$

Учитывая, что для ребра (i_1, i_{2k+1}) из (2.2) следует $x_{i_1i_{2k+1}} = 0$, и то, что

$$x_{i_1i_3} + \sum_{t=2}^k x_{i_{2t-1}i_{2t+1}} = \sum_{t=1}^k x_{i_{2t-1}i_{2t+1}} \quad \text{и} \quad \sum_{t=2}^k x_{i_{2t-1}} = \sum_{t=1}^{k-1} x_{i_{2t+1}},$$

последнее неравенство запишем в следующем виде

$$\sum_{t=1}^k x_{i_{2t-1}i_{2t+1}} - \sum_{t=1}^{k-1} x_{i_{2t+1}} \leq 0. \quad (3.2)$$

Сложив неравенство (3.1) с неравенством (3.2) получаем

$$\sum_{r=1}^{2k} x_{i_r} + \sum_{t=1}^k x_{i_{2t+1}} - \sum_{t=1}^{k-1} x_{i_{2t+1}} = \sum_{r=1}^{2k} x_{i_r} + x_{2k+1} \leq k,$$

откуда следует неравенство

$$\sum_{r=1}^{2k+1} x_{i_r} \leq k \quad \text{или} \quad \sum_{i \in V(C_{2k+1})} x_i \leq k, \quad (3.3)$$

что завершает доказательство леммы.

4. Свойства ЛП-оценки $\alpha_{\nabla}^*(G, w)$

Связь верхней оценки $\alpha_C^*(G, w)$ с ЛП-оценкой $\alpha_{\nabla}^*(G, w)$ характеризует следующая теорема.

Теорема 1. *Для произвольного графа G ЛП-оценка $\alpha_{\nabla}^*(G, w)$ удовлетворяет соотношению:*

$$\alpha_C^*(G, w) \geq \alpha_{\nabla}^*(G, w) \geq \alpha(G, w).$$

Доказательство. Из леммы 1 следует, что $\alpha_{\nabla}^*(G, w) \geq \alpha(G, w)$. Докажем, что $\alpha_C^*(G, w) \geq \alpha_{\nabla}^*(G, w)$. Вначале покажем, что из ограничений (2.2) и (2.3) следует справедливость линейных неравенств

$$x_i + x_j \leq 1 \quad \forall (i, j) = e \in E. \quad (4.1)$$

Сложим первое неравенство из (2.3) соответственно со вторым, третьим и четвертым неравенствами из (2.3). В результате получаем такие неравенства

$$\begin{cases} -x_{jk} + x_j + x_k \leq 1, \\ -x_{ik} + x_i + x_k \leq 1, \\ -x_{ij} + x_i + x_j \leq 1, \end{cases} \quad \forall i, j, k \in V : i < j < k,$$

из которых следует справедливость линейных неравенств

$$-x_{ij} + x_i + x_j \leq 1 \quad \forall (i, j) = e \in E. \quad (4.2)$$

Учитывая, что из ограничений (2.2) имеем $x_{ij} = 0$ для любого $(i, j) = e \in E$, неравенства (4.2) гарантируют справедливость линейных неравенств (4.1).

ЛП-задача (2.1)–(2.4) эквивалентна такой ЛП задаче:

$$\alpha_{\nabla}^*(G, w) = \max \sum_{i \in V} w_i x_i \quad (4.3)$$

при ограничениях

$$x_{ij} = 0 \quad \forall (i, j) = e \in E, \quad (4.4)$$

$$\left\{ \begin{array}{l} -x_{ij} - x_{ik} - x_{jk} + x_i + x_j + x_k \leq 1, \\ +x_{ij} + x_{ik} - x_{jk} - x_i \leq 0, \\ +x_{ij} - x_{ik} + x_{jk} - x_j \leq 0, \\ -x_{ij} + x_{ik} + x_{jk} - x_k \leq 0, \end{array} \right. \quad \forall i, j, k \in V : i < j < k, \quad (4.5)$$

$$0 \leq x_i \leq 1 \quad i \in V, \quad (4.6)$$

$$x_i + x_j \leq 1 \quad \forall (i, j) = e \in E, \quad (4.7)$$

$$\sum_{i \in V(C_{2k+1})} x_i \leq k \quad \forall C_{2k+1} \in V, \quad (4.8)$$

Здесь ограничения (4.7) тождественно равны неравенствам (4.1), а ограничения (4.8) следуют из леммы 2.

От ЛП-задачи (4.3)–(4.8) легко перейти к "ослабленной" ЛП-задаче, убрав из нее ограничения (4.4), (4.5) и оставив ограничения (4.6), (4.7) и (4.8). В результате получаем такую ЛП-задачу

$$\alpha^*(G, w) = \max \sum_{i \in V} w_i x_i \quad (4.9)$$

при ограничениях

$$0 \leq x_i \leq 1 \quad \forall i \in V, \quad (4.10)$$

$$x_i + x_j \leq 1 \quad \forall (i, j) = e \in E. \quad (4.11)$$

$$\sum_{i \in V(C_{2k+1})} x_i \leq k \quad \forall C_{2k+1} \in V, \quad (4.12)$$

для которой $\alpha^*(G, w) \geq \alpha_{\Delta}^*(G, w)$. ЛП-задача (4.9)–(4.12) является ничем иным, как формулировкой ЛП-задачи (1.3)–(1.4) для величины

$\alpha_C^*(G, w)$, и, следовательно, $\alpha^*(G, w) = \alpha_C^*(G, w)$. Отсюда имеем $\alpha_C^*(G, w) \geq \alpha_{\nabla}^*(G, w)$. Теорема доказана.

Из теоремы 1 и соотношения (1.6) для t -совершенных графов следует справедливость такой теоремы.

Теорема 2. *Если граф G – t -совершенный, то*

$$\alpha_{\nabla}^*(G, w) = \alpha(G, w).$$

Следовательно, ЛП-оценка $\alpha_{\nabla}^*(G, w)$ является точной оценкой сверху для взвешенного числа устойчивости t -совершенных графов.

Литература

1. БЕРЖ К. Теория графов и ее применение. – М.: Издательство иностранной литературы, 1962. – 319 с.
2. GROTSCHEL M., LOVASZ L., SCHRIJVER A. Geometric Algorithms and Combinatorial Optimization. – Springer-Verlag, Berlin. – 1988. – 362 p.
3. SCHRIJVER A. Combinatorial Optimization: Polyhedra and Efficiency. – Springer, 2003. – 1881 p.
4. СТЕЦЮК П.И., С.И.БУТЕНКО, О.А.БЕРЕЗОВСКИЙ Об одной верхней оценке для взвешенного числа устойчивости графа // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 2007, №6. – С. 80–89.
5. SHOR N.Z. Nondifferentiable Optimization and Polynomial Problems. – Dordrecht, Kluwer. – 1998. – 394 p.
6. СТЕЦЮК П.И. Новые модели квадратичного типа для задачи о максимальном взвешенном разрезе графа // Кибернетика и системный анализ. – 2006. – N.1. – С. 63-75.

О новых свойствах оценок Шора для взвешенного числа устойчивости графа¹

Стецюк П.И.

*Праці міжнародної конференції "50 років Інституту кібернетики ім. В.М.Глушкова НАН України", К.: Ін-т кібернетики ім. В.М.Глушкова НАН України, 2008.
– С. 164–173.*

Большой вклад в развитие теории сложности алгоритмов для задач непрерывной и дискретной оптимизации внесла предложенная Н.З. Шором техника лагранжевых (двойственных) оценок для значения глобального экстремума целевой функции в невыпуклых квадратичных задачах. В задачах на минимум это будут нижние оценки (оценки снизу), а в задачах на максимум – верхние оценки (оценки сверху). Эта техника включает в себя алгоритмы нахождения лагранжевых оценок на основе методов недифференцируемой оптимизации и использование функционально избыточных ограничений для улучшения точности лагранжевых оценок. Техника лагранжевых оценок может быть использована при нахождении эффективных оценок для значения глобального экстремума целевой функции во многоэкстремальных задачах оптимизации, которые можно сформулировать в форме невыпуклых квадратичных задач. С ее помощью можно выделить такие подклассы NP-трудных задач, которые разрешимы за полиномиальное время.

Техника лагранжевых (двойственных) оценок для невыпуклых квадратичных задач восходит к работам Н.З.Шора [1], где для улучшения точности оценок предложено использовать функционально избыточные ограничения. Именно эти работы послужили основой одного из центральных результатов, который связан с задачами нахождения глобального минимума полиномиальной функции $P(x)$ от одной или нескольких переменных. Эти задачи специальным образом сводятся к квадратичным многоэкстремальным задачам (на минимум) со многими функционально избыточными квадратичными ограничениями. В [1] доказано, что лагранжева оценка для таких квадратичных задач совпадает со значением p^* полинома $P(x)$ в точке глобального минимума тогда

¹Работа выполнена при частичной финансовой поддержке гранта UKM2-2812-KV-06 (CRDF Cooperative Grants Programm)

и только тогда, когда полином $\overline{P}(x) = P(x) - p^*$ может быть представлен в виде суммы квадратов других полиномов. Эти результаты имеют отношение к классическим работам Д. Гильберта по представлению неотрицательных полиномиальных форм в виде суммы квадратов полиномов. Более того, разработанный Н.З.Шором метод дает возможность не только доказать существование такого представления (если оно существует), но и найти одно из таких возможных представлений полинома $\overline{P}(x)$. При этом удается получить значение глобального минимума полинома $P(x)$.

Значительную роль техника лагранжевых оценок играет при исследовании NP-трудных комбинаторных булевых задач. Условие булевости переменной $x \in \{0, 1\}$ можно представить квадратичным равенством $x^2 - x = 0$. Формулировка таких комбинаторных задач в виде нелинейных квадратичных задач позволяет получить более точные двойственные оценки, чем оценки, которые можно получить путем релаксации соответствующих этим задачам булевых задач линейного программирования. В монографии [1] это продемонстрировано на примерах таких экстремальных задач на графах, как задача нахождения максимального взвешенного независимого множества вершин графа, задача нахождения максимального разреза графа, оптимальная бисекция графа, минимальное разбиение графа на k частей с фиксированным числом вершин в каждой части и др. Очень интересными оказались результаты для задачи нахождения максимального устойчивого (независимого) множества вершин графа, где полученные Н. З. Шором двойственные оценки тесно связаны с известными числами Ловаса $\vartheta(G, w)$ и $\vartheta'(G, w)$ [2].

Цель этой работы – исследование новых свойств оценок Шора для взвешенного числа устойчивости неориентированного графа. Покажем, что техника лагранжевых оценок позволяет выделить такие подклассы этой NP-трудной задачи, которые разрешимы за полиномиальное время. Последовательность изложения материала будет следующей. Сначала опишем способ вычисления верхних лагранжевых оценок в квадратичной задаче на максимум, затем покажем, что функционально избыточные ограничения улучшают точность этих оценок. Затем будет описана задача нахождения взвешенного числа устойчивости неориентированного графа и охарактеризована сложность его нахождения. Далее приведена простейшая оценка Шора и ее свойства для слабо двудольных и совершенных графов (связь с числом Ловаса). Затем приведена улучшенная оценка Шора и ее свойства для известных специальных семейств графов: t -совершенных, h -совершенных и W -совершенных

графов. Будет также показано, что свойства улучшенной оценки Шора позволяют выделить семейство графов, для которого взвешенное число устойчивости можно найти за полиномиальное время.

1. Оценки ψ^* и ψ_1^* и их свойства. Рассмотрим оптимизационную квадратичную задачу в следующей формулировке: требуется найти

$$Q_0^* = \sup_{x \in E^n} Q_0(x) \quad (1)$$

при ограничениях

$$Q_i(x) = 0, \quad i = 1, \dots, m. \quad (2)$$

Здесь $Q_\nu(x)$ – квадратичные функции $Q_\nu(x) = (K_\nu x, x) + (b_\nu, x) + c_\nu$; K_ν – симметричные матрицы размером $n \times n$; $b_\nu \in E^n$; c_ν – скаляры, $\nu = 0, 1, \dots, m$. Некоторые из квадратичных функций $Q_\nu(x)$, $\nu = 0, 1, \dots, m$ могут быть и линейными.

В общем случае задача (1), (2) многоэкстремальна и относится к классу NP -трудных задач. Оценку сверху для Q_0^* можно получить следующим образом. Пусть $u = (u_1, \dots, u_m) \in E^m$ – вектор множителей Лагранжа, соответствующий ограничениям (2). Задаче (1), (2) соответствует функция Лагранжа

$$L(x, u) = Q_0(x) + \sum_{i=1}^m u_i Q_i(x) = (K(u)x, x) + (b(u), x) + c(u),$$

где

$$K(u) = K_0(x) + \sum_{i=1}^m u_i K_i(x), \quad b(u) = b_0 + \sum_{i=1}^m u_i b_i, \quad c(u) = c_0 + \sum_{i=1}^m u_i c_i.$$

Рассмотрим функцию

$$\psi(u) = \sup_{x \in E^n} L(x, u) = \sup_{x \in E^n} [(K(u)x, x) + (b(u), x) + c(u)].$$

Пусть $\Omega^- = \{u : \lambda_{\max}[K(u)] < 0\}$ – подмножество тех значений u , для которых матрица $K(u)$ отрицательно определена; Ω^0 – подмножество тех значений u , для которых $\lambda_{\max}[K(u)] = 0$. Здесь $\lambda_{\max}(K)$ – максимальное собственное число симметрической матрицы K размером $n \times n$. Область определения функции $\psi(u)$ (обозначим $\text{dom } \psi$) 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 не-тривиальная оценка сверху для Q_0^*

$$\psi^* = \inf_{u \in \text{dom } \psi} \psi(u) \quad (4)$$

(условие $\psi^* = -\infty$ означает, что система (2) несовместна). Оценка вида (4) есть оптимальной (наилучшей) оценкой сверху для Q_0^* в данном классе лагранжевых оценок. Оценку ψ^* с любой заданной точностью можно найти за полиномиальное время с помощью методов минимизации выпуклых недифференцируемых функций, например, метода эллипсоидов или r -алгоритмов [1]. Схема использования $r(\alpha)$ -алгоритма с адаптивной регулировкой шага подробно описана в [3].

Если ψ^* достигается на $u^* \in \Omega^-$, то

$$\psi^* = \psi(u^*) = Q_0^* = Q_0(x(u^*)),$$

где $x(u^*)$ — решение системы (3) при $u = u^*$. В противном случае ψ^* достигается на границе области Ω^- , при этом может существовать так называемый "разрыв двойственности"

$$\Delta^* = \psi^* - Q_0^* > 0.$$

Предложенный Н.З.Шором способ уменьшения Δ^* связан с введением функционально избыточных ограничений (при этом может увеличиться и множество переменных), добавление которых оставляет множество оптимальных решений задачи (1), (2) неизменным. Однако при этом изменяется функция Лагранжа, что может в некоторых случаях уменьшить разрыв между оптимальным значением Q_0^* целевой функции и лагранжевой (двойственной) оценкой ψ^* . Если добавить в исходную задачу (1), (2) квадратичные функционально избыточные ограничения

$$Q_{m+1}(x) \leq 0, \dots, Q_{m+r}(x) \leq 0, r \geq 1,$$

то новая квадратичная задача примет вид: найти

$$Q_0^* = \sup_{x \in E^n} Q_0(x)$$

при ограничениях

$$Q_i(x) = 0, i = 1, \dots, m,$$

$$Q_i(x) \leq 0, i = m + 1, \dots, m + r.$$

Ей соответствует более длинный вектор множителей Лагранжа

$$U = \{\{u\}, u_{m+1}, \dots, u_{m+r}\}, u_{m+1} \leq 0, \dots, u_{m+r} \leq 0,$$

а функция Лагранжа будет иметь следующий вид:

$$L_1(x, U) = Q_0(x) + \sum_{i=1}^{m+r} u_i Q_i(x) = L(x, u) + \sum_{i=m+1}^{m+r} u_i Q_i(x).$$

Поскольку $L(x, u) = L_1(x, (\{u\}, 0, \dots, 0))$, $\psi_1(\{u\}, 0, \dots, 0) = \psi(u)$, то

$$\psi_1^* = \inf_{U \in \text{dom } \psi_1} \psi_1(U) \leq \inf_{u \in \text{dom } \psi} \psi(u) = \psi^*, \quad (5)$$

и функционально избыточные ограничения могут улучшить лагранжевы оценки.

Ограничения, которые являются линейными комбинациями уже существующих ограничений, не отражаются на точности оценки ψ_1^* . Вклад таких ограничений в функцию Лагранжа эквивалентен лишь определенному изменению множителей Лагранжа при существующих ограничениях. Однако добавление функционально избыточных ограничений, которые являются нетривиальными следствиями условий задачи, в ряде случаев приводит к тому, что оценка ψ_1^* даже может стать точной для Q_0^* ($\psi_1^* = Q_0^*$).

2. Взвешенное число устойчивости графа и сложность его нахождения. Пусть $G = (V, E)$ – неориентированный граф (не содержащий петель) со множеством вершин $V(G)$ и множеством ребер $E(G)$. Для каждой вершины $i \in V(G)$ задан положительный целый вес w_i . Подмножество вершин $S \subseteq V(G)$ называется *устойчивым* (или *независимым*) множеством графа G , если для любых $i, j \in S$ ребро (i, j) не принадлежит $E(G)$. Взвешенное число устойчивости графа G определяется как $\alpha(G, w) = \max \sum_{i \in S} w_i$, где $S \subseteq V(G)$ – устойчивое множество. Подмножество S^* , на котором достигается $\alpha(G, w)$, называется *максимальным взвешенным устойчивым* (или *независимым*) множеством графа G .

В общем случае задача нахождения $\alpha(G, w)$ является NP -трудной [2]. Ее сложность легко объясняется с помощью общей формулировки задачи для $\alpha(G, w)$ через многогранник устойчивых множеств для графа G . Пусть $\chi^S \in \mathbb{R}^{|V|}$ – “инцидентный” вектор для подмножества вершин $S \subseteq V(G)$: т.е. $\chi_i^S = 1$, если $i \in S$, и $\chi_i^S = 0$, если $i \in V \setminus S$. Выпуклая оболочка “инцидентных” векторов χ^S для всех устойчивых множеств S

в графе G называется многогранником устойчивых множеств (stable set polytope) и обозначается:

$$STAB(G) := conv\{\chi^S : S \subseteq V(G) \text{ — устойчивое множество}\}.$$

Таким образом, нахождение $\alpha(G, w)$ связано с задачей максимизации на выпуклом многограннике $STAB(G)$ линейной функции такого вида

$$\alpha(G, w) = \max \sum_{i \in V(G)} w_i x_i, \quad x \in STAB(G). \quad (6)$$

Максимум в задаче (6) достигается в одной или нескольких вершинах многогранника $STAB(G)$. В общем случае многогранник $STAB(G)$ может иметь очень сложную структуру, из-за чего задача (6) и принадлежит к NP -трудным задачам.

Полиномиальная разрешимость задачи (6) связана с семействами графов, для которых многогранники $STAB(G)$ обладают специфическими свойствами. Связано это с очень простым фактом. Пусть $LSTAB(G)$ — некоторый многогранник, заданный с помощью системы линейных ограничений-неравенств, и аппроксимирующий (сверху) многогранник $STAB(G)$. Тогда решение следующей задачи линейного программирования (ЛП-задачи)

$$\alpha_L^*(G, w) = \max \sum_{i \in V(G)} w_i x_i, \quad x \in LSTAB(G) \quad (7)$$

дает оценку $\alpha_L^*(G, w)$, являющуюся оценкой сверху для $\alpha(G, w)$ (в силу того, что многогранник $LSTAB(G)$ аппроксимирует (сверху) многогранник $STAB(G)$, имеем $\alpha_L^*(G, w) \geq \alpha(G, w)$). Количество линейных ограничений-неравенств в ЛП-задаче (7) может и не быть полиномиальным. Если для некоторого семейства графов многогранник $LSTAB(G)$ совпадает со многогранником $STAB(G)$, то оценка $\alpha_L^*(G, w)$ для этого семейства графов будет точной ($\alpha_L^*(G, w) = \alpha(G, w)$). Если оценку $\alpha_L^*(G, w)$ можно найти за полиномиальное время, то для любого графа G из этого семейства задача нахождения $\alpha(G, w)$ разрешима за полиномиальное время.

Этот принцип лежит в основе обоснования полиномиальной разрешимости задачи нахождения $\alpha(G, w)$ для слабо двудольных (bipartite), совершенных (perfect), t -совершенных (t-perfect), W -совершенных (W -perfect) и ряда других графов (см. [2, гл. 9]). За каждым из этих семейств графов стоят свои многогранники вида $LSTAB(G)$, которые имеют даже свои имена, например, $FSTAB(G)$ — дробный многогранник устойчивых множеств (fractional stable set polytop), $QSTAB(G)$ —

кликковый многогранник (clique polytope), многогранник нечетных циклов (odd-cycle polytope), "колесный" многогранник (wheel polytope). Однако каждому из указанных семейств графов соответствует свой специфический способ нахождения соответствующей ему оценки вида $\alpha_L^*(G, w)$. Так, например, для семейства совершенных графов (для них $STAB(G) = QSTAB(G)$), нахождение $\alpha_Q^*(G, w)$ за полиномиальное время обеспечивает известное число Ловаса $\vartheta(G, w)$ [2], которое для произвольного графа G является более точной оценкой сверху для $\alpha(G, w)$, чем оценка $\alpha_Q^*(G, w)$. Для t -совершенных графов (для них $STAB(G) = CSTAB(G)$) полиномиальная разрешимость задачи нахождения $\alpha(G, w)$ связана со свойствами оценки $\alpha_C^*(G, w)$, которая для произвольного графа G может быть найдена за полиномиальное время. Однако всем этим полиномиально разрешимым случаям легко придать один и тот же смысл, если использовать верхние оценки для $\alpha(G, w)$, построенные Н.З.Шором в рамках невыпуклых квадратичных задач.

3. Оценка $\psi(G, w)$, слабо двудольные и совершенные графы. Простейшая верхняя оценка для $\alpha(G, w)$ (условимся называть ее $\psi(G, w)$) предложена Н.З.Шором в [1]. Она связана с формулировкой задачи для $\alpha(G, w)$ в виде следующей невыпуклой квадратичной задачи:

$$\alpha(G, w) = \max \sum_{i \in V(G)} w_i x_i \quad (8)$$

при ограничениях

$$x_i x_j = 0 \quad \forall (i, j) \in E(G), \quad (9)$$

$$x_i^2 - x_i = 0 \quad \forall i \in V(G), \quad (10)$$

где булева переменная $x_i \in \{0, 1\}$ равна единице, если вершина i включается в устойчивое множество, и равна нулю в противном случае. Здесь булевы переменные описаны равенствами (10). Ограничения (9) означают, что две вершины не могут одновременно принадлежать устойчивому множеству, если они связаны ребром в графе G .

Оценка $\psi(G, w)$ является оптимальной верхней оценкой вида ψ^* (4) для максимального значения целевой функции в квадратичной задаче (8)–(10). Оценку $\psi(G, w)$ можно найти с любой заданной точностью за полиномиальное время. Показано (см. [1]), что оценка $\psi(G, w)$ совпадает со взвешенным числом Ловаса $\vartheta(G, w)$. Для произвольного графа G имеем

$$\alpha(G, w) \leq \psi(G, w) = \vartheta(G, w).$$

Если граф G – совершенный или слабо двудольный, то

$$\alpha(G, w) = \psi(G, w) = \vartheta(G, w),$$

и оценка $\psi(G, w)$ является точной для $\alpha(G, w)$. Такой способ обоснования точности оценки $\psi(G, w)$ для этих семейств графов связан со свойствами взвешенного числа Ловаса.

Обоснование того, что простейшая оценка Шора совпадает с $\alpha(G, w)$ для слабо двудольных и совершенных графов, можно сделать более наглядным. Оно связано с тем, что из квадратичных ограничений (9)–(10) следуют семейства линейных неравенств

$$(\text{vertex constraints}) \quad 0 \leq x_i \leq 1 \quad \forall i \in V(G), \quad (11)$$

$$(\text{edge constraints}) \quad x_i + x_j \leq 1 \quad \forall (i, j) \in E(G), \quad (12)$$

$$(\text{clique constraints}) \quad \sum_{i \in V(Q)} x_i \leq 1 \quad \forall Q \in G, \quad (13)$$

которые выполняются для многогранника $STAB(G)$. Здесь Q – клика (полный подграф) в графе G . "Вершинные" неравенства (11) получаются релаксацией (ослаблением) ограничений (10). Справедливость кликовых неравенств (13) показана Н.З.Шором в [1, с.252], а справедливость реберных неравенств есть следствием кликовых неравенств в случае, когда клика в графе G состоит из двух вершин, т.е. совпадает с ребром графа G .

В результате релаксации квадратичной задачи (8)–(10) легко получить оценку $\alpha_F^*(G, w)$ для дробного многогранника устойчивых множеств:

$$FSTAB(G) = \{x \in R^{|V|} : x \text{ удовлетворяет (11) и (12)}\},$$

и оценку $\alpha_Q^*(G, w)$ для кликового многогранника

$$QSTAB(G) = \{x \in R^{|V|} : x \text{ удовлетворяет (11) и (13)}\}.$$

Поскольку $FSTAB(G) \subseteq QSTAB(G)$, то для произвольного графа G имеет место неравенство

$$\alpha_F^*(G, w) \geq \alpha_Q^*(G, w) \geq \psi(G, w) \geq \alpha(G, w),$$

из которого следует, что оценка $\psi(G, w)$ является точной как для слабо двудольных, так и для совершенных графов. Для этого достаточно лишь воспользоваться тем, что для слабо двудольных графов $\alpha_F^*(G, w) = \alpha(G, w)$ (для них $STAB(G) = FSTAB(G)$ и отсутствуют изолированные вершины), а для совершенных графов $\alpha_Q^*(G, w) = \alpha(G, w)$ (для них $STAB(G) = QSTAB(G)$).

Для t -совершенных и W -совершенных графов оценка $\psi(G, w)$ не будет точной. Ее можно улучшить за счет использования функционально избыточных ограничений.

4. Оценка $\psi_1(G, w)$ и специальные семейства графов. Улучшенная верхняя оценка Шора для $\alpha(G, w)$ (условимся обозначать ее $\psi_1(G, w)$) связана с такой невыпуклой квадратичной задачей:

$$\alpha(G, w) = \max \sum_{i \in V(G)} w_i x_i \quad (14)$$

при ограничениях

$$x_i x_j = 0 \quad \forall (i, j) \in E(G), \quad (15)$$

$$x_i^2 - x_i = 0 \quad \forall i \in V(G), \quad (16)$$

$$x_i x_k + x_j x_k \leq x_k, \quad \forall (i, j) \in E(G), k \neq i, j, \quad (17)$$

Задача (14)–(17) получена в результате добавления к задаче (8)–(10) семейства функционально избыточных квадратичных ограничений в форме неравенств (17), которые не изменяют множества оптимальных решений квадратичной задачи (8)–(10). Функционально избыточные ограничения получаются умножением реберных неравенств в форме (12) на переменные x_k , такие что $k \neq i, j$. Знак неравенств не изменится, поскольку $x_k = x_k^2 \geq 0$. Именно наличие ограничений (17) придает оценке $\psi_1(G, w)$ ряд замечательных свойств для специальных семейств графов.

Оценка $\psi_1(G, w)$ является оптимальной верхней оценкой вида ψ_1^* (5) для максимального значения целевой функции в квадратичной задаче (14)–(17). Оценку $\psi_1(G, w)$ можно найти с любой заданной точностью за полиномиальное время. Конечно, найти ее сложнее, чем оценку $\psi(G, w)$, так как количество множителей Лагранжа увеличено за счет функционально избыточных ограничений (17). Что дают эти функционально избыточные ограничения? Во-первых, для произвольного графа G оценка $\psi_1(G, w)$ удовлетворяет соотношению

$$\alpha(G, w) \leq \psi_1(G, w) \leq \psi(G, w),$$

и всегда является не менее точной оценкой сверху для $\alpha(G, w)$, чем оценка $\psi(G, w)$. Во-вторых, оценка $\psi_1(G, w)$ является точной для взвешенного числа устойчивости t -совершенных графов. Это свойство оценки $\psi_1(G, w)$ следует из того, что из квадратичных ограничений (15)–(17) вытекают линейные неравенства

$$(odd - cycle constraints) \quad \sum_{i \in V(C_{2k+1})} x_i \leq k \quad \forall C_{2k+1} \in G, \quad (18)$$

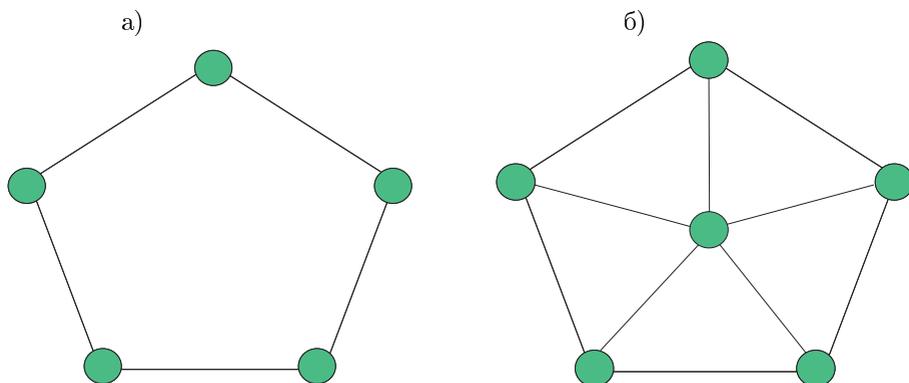


Рис. 1. Нечетный цикл C_5 и колесо W_6 (здесь $k = 2$)

которые справедливы для многогранника $STAB(G)$ (см. [1, с.252]). Здесь C_{2k+1} , $k = 1, 2, \dots$ – нечетный цикл в графе G (т.е. в нем содержится нечетное количество вершин). Пример нечетного цикла C_5 показан на рис. 1,а.

Релаксацией квадратичной задачи (15)–(17) легко получить оценку $\alpha_C^*(G, w)$ для многогранника нечетных циклов

$$CSTAB(G) = \{x \in R^{|V|} : x \text{ удовлетворяет (11), (12) и (18)}\},$$

которая для произвольного графа G удовлетворяет неравенству

$$\alpha_C^*(G, w) \geq \psi_1(G, w) \geq \alpha(G, w).$$

Из последнего следует, что оценка $\psi_1(G, w)$ совпадает с $\alpha(G, w)$ для t -совершенных графов, для которых $STAB(G) = CSTAB(G)$ и $\alpha_C^*(G, w) = \alpha(G, w)$.

Можно ли для улучшенной оценки Шора $\psi_1(G, w)$ показать более сильные результаты о ее точности? Например, будет ли она точной оценкой для $\alpha(G)$ в случае W -совершенных графов? В последних используются линейные неравенства

$$(wheel\ constraints) \quad \sum_{i \in V(C_{2k+1})} x_i + kx_{i_{2k+2}} \leq k, \quad \forall W_{2k+2} \in G, \quad (19)$$

которые справедливы для многогранника $STAB(G)$. Здесь W_{2k+2} – колесо (wheel) в графе G (оно состоит из нечетного цикла C_{2k+1} и одной вершины, связанной с каждой из вершин цикла). Пример колеса W_6 ,

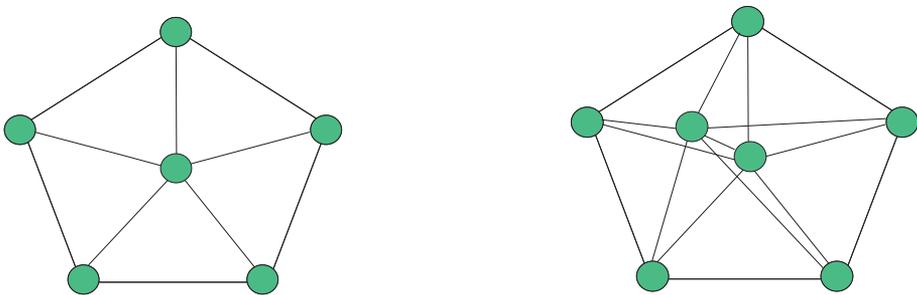


Рис. 2. Примеры 1-колеса и 2-колеса на базе нечетного цикла C_5

построенного на основе нечетного цикла C_5 , приведен на рис. 1,б. "Колесный" многогранник (wheel polytope) имеет вид

$WSTAB(G) = \{x \in R^{|V|} : x \text{ удовлетворяет (11), (12), (18) и (19)}\}$, и с ним связана оценка $\alpha_W^*(G, w)$, которая может быть найдена за полиномиальное время для произвольного графа G [2].

Оказывается, что оценка $\psi_1(G, w)$ совпадает с $\alpha(G, w)$ для W -совершенных графов. Более того, это так даже для более сложного семейства графов, чем W -совершенные. Связано это с тем, что из ограничений (15)–(17) следуют линейные неравенства

$$(p\text{-wheel constraints}) \quad \sum_{i \in V(C_{2k+1})} x_i + k \sum_{j \in V(Q_p)} x_j \leq k, \quad \forall W_{2k+1+p} \in G, \quad (20)$$

которые справедливы для многогранника $STAB(G)$ [5]. Здесь подграф W_{2k+1+p} является p -колесом (p -wheel) [4]. Вершинами p -колеса W_{2k+1+p} служат вершины непересекающихся нечетного цикла C_{2k+1} и клики Q_p (полный подграф, содержащий p вершин). Множество ребер для W_{2k+1+p} включает все ребра нечетного цикла C_{2k+1} , все ребра клики Q_p , а также ребра, связывающие каждую вершину C_{2k+1} со всеми вершинами клики Q_p . Примеры 1-колеса и 2-колеса на базе нечетного цикла C_5 показаны на рис. 2.

Неравенства (20) означают, что для каждого p -колеса из графа G в устойчивое (независимое) множество может быть включена либо одна из вершин клики Q_p , либо k вершин из нечетного цикла C_{2k+1} . В частном случае, когда клика Q_p состоит всего из одной вершины, то тогда из линейных неравенств для произвольного p -колеса в форме (20) следуют линейные неравенства (19) для обычных колес W_{2k+2} .

Комбинируя ограничения (11), (12), (13) и (18) с ограничениями (20)

легко построить более сложный класс графов, чем W -совершенные, для которых улучшенная оценка Шора будет точной для $\alpha(G, w)$. Пусть p -колесный многогранник (p -wheel polytope) задан как

$W_pSTAB(G) = \{x \in R^{|V|} : x \text{ удовлетворяет (11), (12), (18) и (20)}\}$, и ему соответствует верхняя оценка

$$\alpha_{W_p}^*(G, w) = \max \sum_{i \in V(G)} w_i x_i, \quad x \in W_pSTAB(G).$$

Справедлива следующая теорема.

Теорема 1. *Для произвольного графа G справедливо неравенство*

$$\alpha_{W_p}^*(G, w) \geq \psi_1(G, w) \geq \alpha(G, w). \quad (21)$$

Семейство графов, для которых $STAB(G) = W_pSTAB(G)$, назовем W_p -совершенными. Для них $\alpha_{W_p}^*(G, w) = \alpha(G, w)$, и неравенство (21) превращается в строгое равенство

$$\alpha_{W_p}^*(G, w) = \psi_1(G, w) = \alpha(G, w), \quad (22)$$

из которого вытекает справедливость следующей теоремы.

Теорема 2. *Если граф G является W_p -совершенным, то оценка $\psi_1(G, w)$ равна $\alpha(G, w)$.*

Поскольку оценка $\psi_1(G, w)$ с любой заданной точностью может быть найдена за полиномиальное время, из теоремы 2 следует, что для любого графа G из семейства W_p -совершенных графов задача нахождения $\alpha(G, w)$ является полиномиально разрешимой.

Выводы. Полиномиальную разрешимость задач нахождения $\alpha(G, w)$ для совершенных, t - и W -совершенных графов можно получить путем расширения (ослабления) многогранника $W_pSTAB(G)$ за счет изъятия некоторых типов неравенств. Так, например, многогранник $CSTAB(G)$ следует из многогранника $W_pSTAB(G)$ при отсутствии p -колесных ограничений. Кликовый многогранник $QSTAB(G)$ получается из многогранника $W_pSTAB(G)$ при $k = 1$ (т.е. рассматриваются только нечетные циклы, совпадающие с 3-кликкой) и произвольном p : $1 \leq p \leq |V| - 3$. Здесь неравенства для 2-клик следуют из реберных неравенств (12), неравенства для 3-клик – из неравенств для нечетных циклов (18) при $k = 1$, а неравенства для клик с большим количеством вершин – из неравенств для p -колес (20) при $k = 1$ и произвольном p . Многогранник $WSTAB(G)$ получается из $W_pSTAB(G)$ при $p = 1$. В результате справедливо неравенство

$$\alpha_W^*(G, w) \geq \psi_1(G, w) \geq \alpha(G, w),$$

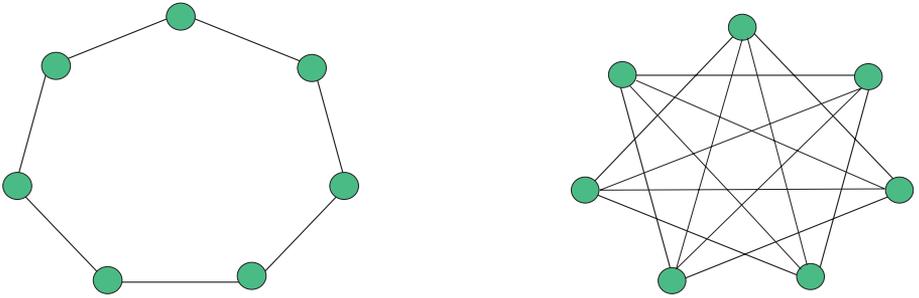


Рис. 3. Нечетный цикл C_7 и дополнительный \overline{C}_7 (здесь $k = 3$)

из которого следует, что оценка $\psi_1(G, w)$ совпадает с $\alpha(G, w)$, если граф G является W -совершенным (для него $\alpha_W^*(G, w) = \alpha(G, w)$).

В заключение отметим, что с помощью оценки Шора $\psi_1^*(G, w)$ можно обосновать полиномиальную разрешимость задачи нахождения взвешенного числа устойчивости графа для более сложного семейства графов, чем W_p -совершенные. Так, например, из ограничений (15)–(17) следуют справедливые для многогранника $STAB(G)$ линейные неравенства

$$(\text{odd} - \text{antihole constraints}) \quad \sum_{i \in V(\overline{C}_{2k+1})} x_i \leq 2, \quad \forall \overline{C}_{2k+1} \in G, \quad (23)$$

где \overline{C}_{2k+1} – дополнительный граф для нечетного цикла C_{2k+1} (в зарубежной литературе его принято называть "odd-antihole"). Пример графа \overline{C}_7 показан на рис. 3.

Аналогично тому, как и в случае p -колеса, можно определить граф \overline{C}_{2k+1+p} как расширение \overline{C}_{2k+1} за счет клики Q_p , вершины которой связаны с каждой из вершин \overline{C}_{2k+1} . Из ограничений (15)–(17) следуют линейные неравенства

$$(\text{p-odd-antihole constraints}) \quad \sum_{i \in V(\overline{C}_{2k+1})} x_i + 2 \sum_{j \in V(Q_p)} x_j \leq 2, \quad \forall \overline{C}_{2k+1+p} \in G, \quad (24)$$

которые справедливы для многогранника $STAB(G)$.

Более того, можно построить еще ряд новых семейств линейных неравенств, которые будут справедливы для многогранника $STAB(G)$ и будут характеризовать более сложные подструктуры в графе, полученные определенной комбинацией клик Q_p , нечетных циклов C_{2k+1} и их

дополнений \overline{C}_{2k+1} . Это дает возможность с помощью оценки $\psi_1(G, w)$ указать новые семейства графов, для которых задача (6) является полиномиально разрешимой и, более того, указать верхнюю границу сложности этих классов задач на основе методов недифференцируемой оптимизации. Это может в значительной мере дополнить имеющиеся результаты в [2,6] и, возможно, найти для многих семейств графов некоторую единообразную форму описания многогранника вида $LSTAB(G)$ для внешней аппроксимации многогранника устойчивых множеств $STAB(G)$.

Литература

1. SHOR N.Z. Nondifferentiable optimization and polynomial problems. – Boston; Dordrecht; London: Kluwer Academic Publishers, 1998. – 412 p.
2. GRÖTSCHEL M., LOVASZ L., SCHRIJVER A. Geometric algorithms and combinatorial optimization. – Berlin: Springer-Verlag, 1988. – 362 p.
3. ШОР Н.З., СТЕЦЮК П.И. Использование модификации г-алгоритма для нахождения глобального минимума полиномиальных функций // Кибернетика и системный анализ. – 1997. – № 4. – С. 28–49.
4. CHENG E. AND CUNNINGHAM W.H. Wheel inequalities for stable set polytopes // Math. Program. – 1997. – 77. – № 3. – P. 389–421.
5. СТЕЦЮК П.И., ЧУМАКОВ Б.М. О свойствах одной верхней оценки Н.З. Шора для взвешенного числа устойчивости графа // Пр. міжнар. симп. "Питання оптимізації обчислень (ПОО-XXXIII)". – Київ: Ін-т кібернетики ім. В.М.Глушкова НАН України, 2007. – С. 271–272.
6. SCHRIJVER A. Combinatorial Optimization: Polyhedra and Efficiency. – Berlin: Springer-Verlag, 2003. – 1881 p.

ЛП-ориентированная верхняя оценка для числа устойчивости графа на основе p -колес¹

П.И. Стецюк, С.И. Бутенко, А.П. Лиховид

Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2008. – № 7. – С. 34–44.

Для числа устойчивости неориентированного графа предложена верхняя оценка, которая базируется на аппроксимации многогранника устойчивых множеств с помощью линейных неравенств для p -колес в графе. Описан алгоритм нахождения ослабленной верхней оценки на основе решения задачи линейного программирования с конечным числом ограничений-неравенств, связанных с нечетными циклами и p -колесами в графе. Даны результаты тестовых экспериментов для DIMACS-графов и графов, связанных с максимальным объемом помехоустойчивых кодов.

Задача о максимальном независимом (устойчивом) множестве вершин графа является одной из центральных в теории графов. Она имеет много важных приложений, связанных с выбором в графе экстремальных множеств вершин с заданными свойствами. Так, например, непосредственное приложение этой задачи связано с нахождением максимального объема помехоустойчивых кодов (кодов корректирующих ошибки при передаче информации). К задаче о максимальном независимом множестве вершин графа сводятся задача о максимальной клике графа и задача о максимальной k -кликке графа. Последние могут быть использованы при нахождении взаимно-связанных подмножеств при информационном анализе социологических, биологических, телекоммуникационных и др. массивов данных. Такие подмножества в социологических данных могут характеризовать родственные, криминальные или профессиональные связи; в телекоммуникациях – группы абонентов, которые часто общаются; в биологических данных нахождение таких подмножеств может означать наличие причинно-следственных связей при функционировании отдельных частей живого организма. С помощью задачи о максимально независимом множестве вершин графа можно представить и классическую задачу о раскраске вершин неориентированного графа k -красками.

¹Работа выполнена при частичной финансовой поддержке гранта UKM2-2812-KV-06 (CRDF Cooperative Grants Programm)

1. Постановка задачи и оценка $\alpha_C^*(G)$. Пусть $G=(V, E)$ – неориентированный граф (не содержащий петель) с множеством вершин $V(G)$ и множеством ребер $E(G)$. Подмножество вершин $S \subseteq V(G)$ называется *устойчивым (или независимым)* множеством графа G , если для любых $i, j \in S$ ребро (i, j) не принадлежит $E(G)$. Мощность максимального по числу входящих в него вершин устойчивого множества в графе G называется числом устойчивости $\alpha(G)$. Подмножество S^* , на котором достигается $\alpha(G)$, называется максимальным устойчивым (*или независимым*) множеством графа G .

В общем случае задача нахождения $\alpha(G)$ является *NP*-трудной задачей [1]. Теоретический и практический интерес представляет нахождение верхних оценок, достаточно хорошо аппроксимирующих сверху $\alpha(G)$. Одной из таких верхних оценок есть оценка $\alpha_C^*(G)$, которая связана с решением следующей задачи линейного программирования (*ЛП*-задачи)

$$\alpha_C^*(G) = \max \sum_{i \in V(G)} x_i \quad (1)$$

при ограничениях

$$0 \leq x_i \leq 1 \quad \forall i \in V(G), \quad (2)$$

$$x_i + x_j \leq 1 \quad \forall (i, j) \in E(G), \quad (3)$$

$$\sum_{i \in V(C_{2k+1})} x_i \leq k \quad \forall C_{2k+1} \in G. \quad (4)$$

Здесь C_{2k+1} , $k = 1, 2, \dots$ – нечетный цикл в графе G (содержит нечетное количество вершин).

При построении оценки $\alpha_C^*(G)$ используются следующие семейства линейных неравенств: (2) – вершинные ограничения (vertex constraints), (3) – реберные ограничения (edge constraints), (4) – нечетного цикла ограничения (odd-cycle constraints). Вместе эти три семейства неравенств определяют многогранник нечетных циклов (odd-cycle polytope)

$$CSTAB(G) = \left\{ x \in R^{|V|} : x \text{ удовлетворяет (2), (3) и (4)} \right\},$$

который аппроксимирует (сверху) многогранник $STAB(G)$ – многогранник устойчивых множеств (stable set polytope). Последний является выпуклой оболочкой инцидентных векторов устойчивых множеств S в G и с помощью реберных неравенств (3) может быть представлен в следующей форме:

$$STAB(G) = \text{conv}\{x \in \{0, 1\}^{|V|} : x_i + x_j \leq 1 \quad \forall (i, j) \in E(G)\}. \quad (5)$$

С помощью многогранника $CSTAB(G)$ ЛП-задача (1)–(4) для оценки $\alpha_C^*(G, w)$ может быть сформулирована в более простой форме:

$$\alpha_C^*(G, w) = \max \sum_{i \in V(G)} x_i, \quad x \in CSTAB(G).$$

Оценка $\alpha_C^*(G, w)$ с любой заданной точностью может быть найдена за полиномиальное время [2]. Полиномиальный алгоритм для ее нахождения [2] использует метод эллипсоидов и тот факт, что для точки $\bar{x} = \{\bar{x}_1, \dots, \bar{x}_{|V|}\}$, которая удовлетворяет ограничениям (2)–(3), за полиномиальное время можно либо убедиться, что точка \bar{x} удовлетворяет ограничениям (4) для каждого нечетного цикла, либо найти такой нечетный цикл, для которого ограничение вида (4) является максимально нарушенным.

Для произвольного графа G всегда имеет место неравенство $\alpha_C^*(G) \geq \alpha(G)$. Если граф G принадлежит семейству t -совершенных графов (для них $STAB(G) = CSTAB(G)$), то оценка $\alpha_C^*(G)$ является точной для числа устойчивости графа, т.е. $\alpha_C^*(G) = \alpha(G)$. В результате задача нахождения числа устойчивости t -совершенного графа разрешима за полиномиальное время. Однако, семейство t -совершенных графов есть достаточно слабым и в него не попадают все те графы, для которых $\alpha(G) < |V|/3$ (предполагается, что граф не содержит изолированных вершин). Поэтому, если $\alpha(G) < |V|/3$, то оценка $\alpha_C^*(G)$ оказывается грубой для аппроксимации сверху $\alpha(G)$. Верхнюю оценку, которая будет более тесно аппроксимировать сверху $\alpha(G)$, чем оценка $\alpha_C^*(G)$, легко построить на основе такой подструктуры в графе G , как p -колесо.

2. Оценка $\alpha_{W_p}^*(G)$. Пусть имеется граф W_{2k+1+p} , у которого вершины состоят из вершин непересекающихся нечетного цикла C_{2k+1} и клики Q_p (полный подграф, содержащий p вершин). Если $p = 1$, то клика состоит из одной вершины. Множество ребер в графе W_{2k+1+p} включает все ребра нечетного цикла C_{2k+1} , все ребра клики Q_p , а также ребра, связывающие каждую вершину C_{2k+1} со всеми вершинами клики Q_p . Граф W_{2k+1+p} принято называть p -колесом (англ. термин – p -wheel) [3]. Примеры 1-колеса и 2-колеса на базе нечетного цикла C_5 показаны на рис. 1. С p -колесом связано семейство линейных неравенств (p -wheel constraints)

$$\sum_{i \in V(C_{2k+1})} x_i + k \sum_{j \in V(Q_p)} x_j \leq k, \quad \forall W_{2k+1+p} \in G, \quad (6)$$

которые справедливы (valid) для многогранника $STAB(G)$. Неравенства (6) означают, что для каждого p -колеса из графа G в устойчивое

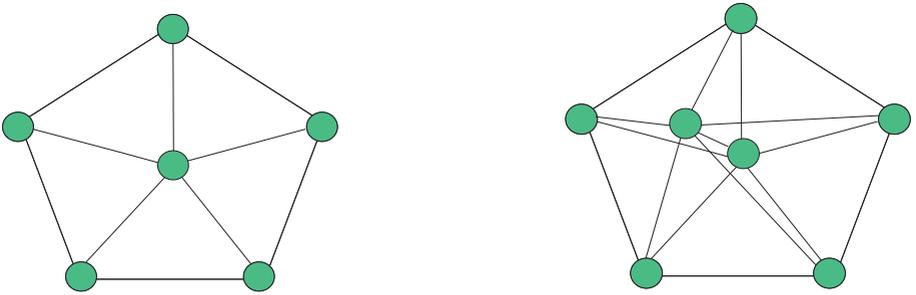


Рис. 1. Примеры 1-колеса и 2-колеса на базе нечетного цикла C_5

(независимое) множество может быть включена либо одна из вершин клики Q_p , либо k вершин из нечетного цикла C_{2k+1} . Если семейство неравенств (6) добавить к линейным неравенствам, которые определяют многогранник $CSTAB(G)$, то получим многогранник

$$W_pSTAB(G) = \left\{ x \in R^{|V|} : x \text{ удовлетворяет (2), (3), (4) и (6)} \right\},$$

который аппроксимирует (сверху) многогранник устойчивых множеств $STAB(G)$.

Оценка $\alpha_{W_p}^*(G)$ связана с решением следующей ЛП-задачи

$$\alpha_{W_p}^*(G) = \max \sum_{i \in V(G)} x_i, \quad x \in W_pSTAB(G). \quad (7)$$

Для произвольного графа G всегда имеет место неравенство

$$\alpha_C^*(G) \geq \alpha_{W_p}^*(G) \geq \alpha(G), \quad (8)$$

т.е. оценка $\alpha_{W_p}^*(G)$ является верхней оценкой (оценкой сверху) для $\alpha(G)$ и она всегда не хуже, чем оценка $\alpha_C^*(G)$. В общем случае нахождение оценки $\alpha_{W_p}^*(G)$ – NP-трудная задача. Более того, в общем случае NP-трудной является задача нахождения на основе уже имеющегося нечетного цикла такого p -колеса, для которого в точке $x^* = \{x_1^*, \dots, x_{|V|}^*\}$ максимально нарушается линейное неравенство в форме (6). Эта задача равносильна нахождению максимальной взвешенной клики для индуцированного подграфа из G с подмножеством вершин из $V(G)$, где каждая вершина связана со всеми вершинами нечетного цикла. Веса вершин в индуцированном подграфе равны соответствующим компонентам вектора x^* .

Поэтому, взамен оценки $\alpha_{W_p}^*(G)$ рассмотрим ослабленную верхнюю оценку для $\alpha(G)$ на основе линейных неравенств для таких p -колес в графе G , которые легко построить на основе имеющихся нечетных циклов. Условимся эту оценку называть ЛП-ориентированной оценкой $\alpha_{W_p}(G)$, и алгоритм для ее нахождения будем строить на основе решения ЛП-задачи с конечным количеством линейных неравенств.

3. ЛП-ориентированная оценка $\alpha_{W_p}(G)$. Алгоритм для нахождения оценки $\alpha_{W_p}(G)$ условимся называть алгоритмом LPWSTAB(p, ε). В его основу положим ЛП-задачу в следующем виде:

$$f_{W_p}^* = \max \sum_{i \in V(G)} x_i, \quad (9)$$

при ограничениях:

$$0 \leq x_i \leq 1 \quad \forall i \in V(G), \quad (10)$$

$$x_i + x_j \leq 1 \quad \forall (i, j) \in E(G), \quad (11)$$

$$\sum_{i \in V(C_{2k+1})} x_i \leq k \quad \forall C_{2k+1} \in C_{\text{odd}} \in G, \quad (12)$$

$$\sum_{i \in V(C_{2k+1})} x_i + k \sum_{j \in V(Q_{p'})} x_j \leq k, \quad \forall W_{2k+1+p'} \in W_{\text{odd}} \in G, p' \leq p \quad (13)$$

где C_{odd} и W_{odd} – конечные множества нечетных циклов и p' -колес, возможно пустые.

Пусть на k -ой итерации алгоритма получен вектор $x^* = \{x_1^*, \dots, x_{|V|}^*\}$ – оптимальная точка ЛП-задачи (9)–(13). Ее будем использовать для дополнения множеств нечетных циклов и p' -колес ($p' \leq p$) для очередной $(k+1)$ -ой итерации. На основе вектора x^* найдем множество нечетных циклов с помощью алгоритма нахождения нечетных циклов из [2], который использован в полиномиальном алгоритме для оценки $\alpha_C^*(G)$. На основе найденных нечетных циклов сформируем множество p' -колес ($p' \leq p$) следующим образом. Дополним нечетные циклы до p' -колес с помощью простейшего по количеству вычислений алгоритма (по типу алгоритма "жадного"), который последовательно включает наилучшие вершины, с помощью их можно расширить либо нечетный цикл, либо уже существующее p' -колесо. Наилучшей из списка вершин (кандидатов на расширение p' -колеса) считается та вершина, которой соответствует максимальная компонента в векторе x^* . Такой алгоритм построения p' -колеса условимся называть алгоритмом "последовательного включения наилучших вершин".

Для включения в ЛП-задачу новых ограничений будем использовать параметр ε (точность для проверки нарушенных ограничений). Он означает, что линейные неравенства, связанные с нечетным циклом C_{2k+1} и p' -колесом $W_{2k+1+p'}$, считаются нарушенными (в точке x^*) и включаются в ЛП-задачу, если для них выполняются условия:

$$\sum_{i \in V(C_{2k+1})} x_i^* \geq k + \varepsilon \quad \text{и} \quad \sum_{i \in V(C_{2k+1})} x_i^* + k \sum_{j \in V(Q_{p'})} x_j^* \geq k + \varepsilon.$$

Итеративный алгоритм LPWSTAB(p, ε) для нахождения оценки $\alpha_{W_p}(G)$ можно описать следующим образом.

Положим $itn = 0$ и $C_0 = \emptyset$, $W_0 = \emptyset$ (множества нечетных циклов и p' -колес есть пустыми). Перейдем шагу 1.

- Шаг 1. Имеем множество нечетных циклов C_{itn} и множество p' -колес W_{itn} . Положим $C_{odd} = C_{itn}$, $W_{odd} = W_{itn}$ и, решив ЛП-задачу (9)–(13), найдем $f_{W_p}^*$ и $x^* = (x_1^*, \dots, x_{|V|}^*)$.
- Шаг 2. Построим взвешенный неориентированный граф $G' = (V', E')$. Множество вершин $V'(G')$ включает вершину i и ее копию i' для всех $i \in V(G)$. Множество ребер $E'(G')$ состоит из ребер, которые соединяют все те пары вершин (i, j') и (i', j) , для которых пара вершин i и j соединена ребром (i, j) в графе G . Присвоим ребрам (i, j') и (i', j) один и тот же положительный вес (длина ребра) $l(i', j) = l(i, j) = 1 - x_i^* - x_j^*$.
- Шаг 3. В графе G' , рассматривая его как ориентированный (ребру (i, j') соответствуют дуги (i, j') и (j', i) с длинами, равными длине ребра (i, j) , т.е. $l(i, j')$), для каждой вершины $i \in V$ найдем кратчайший путь, который начинается в вершине i и заканчивается в вершине i' (копии вершины i). Число таких кратчайших путей равно $|V|$. Из них оставляем только те, которые определяют нечетный цикл. Пусть число таких кратчайших путей равно n ($n \leq |V|$). Найдем нечетные циклы $C_{2k_1+1}, \dots, C_{2k_n+1}$ (они соответствуют n найденным кратчайшим путям).
- Шаг 4. Установим множества C' и W' – пустыми и будем их заполнять следующим образом. Каждый из нечетных циклов C_{2k_i+1} для $i = 1, \dots, n$ с помощью алгоритма "последовательного включения наилучших вершин" дополняем до p' -колеса в графе G , где $p' \leq p$. Пусть такое p' -колесо получено и если линейное неравенство (6) для него нарушено, то найденное p' -колесо включаем в множество W' .

Если построить p' -колесо на базе нечетного цикла не удалось, но линейное неравенство в виде (4) для этого нечетного цикла является нарушенным, то такой нечетный цикл включаем в множество C' . Если после завершения данной процедуры для всех $i = 1, \dots, n$ оба множества C' и W' являются пустыми, то установить $\alpha_{W_p}(G) = f_{W_p}^*$ и останов. В противном случае перейдем к шагу 5.

Шаг 5. Положим $C_{itn+1} = C_{itn} \cup C'$ и $W_{itn+1} = W_{itn} \cup W'$ и $itn = itn + 1$. Перейдем к шагу 1.

О свойствах оценки $\alpha_{W_p}(G, w)$ можно сказать, что в "идеализированном" случае, т.е. когда $\varepsilon = 0$, для нее справедливо неравенство

$$\alpha_C^*(G) \geq \alpha_{W_p}(G) \geq \alpha(G),$$

которое вытекает из следующего. ЛП-задача (9)–(13) построена таким образом, что если из нее убрать p -колесные ограничения (13), то получим оценку $\alpha_C^*(G, w)$, если окажется, что отсутствуют нечетные циклы, для которых ограничения есть нарушенными. За счет добавления к ЛП-задаче (9)–(12) конечного набора линейных неравенств для p -колес в форме (13) оценку $\alpha_C^*(G)$ мы не сможем ухудшить, а сможем ли улучшить – уже зависит от того, насколько активными в ЛП-задаче окажутся включенные ограничения. Поэтому, оценку $\alpha_{W_p}(G)$ можно рассматривать как улучшение оценки $\alpha_C^*(G)$ с помощью использования линейных неравенств для p -колес в графе G . В результате, если p -колесо не удалось построить, то она будет сохранять свойства оценки $\alpha_C^*(G)$ и будет точной оценкой для числа устойчивости t -совершенных графов. В случае обнаружения p -колес, для которых линейные неравенства в форме (6) являются нарушенными, оценка $\alpha_{W_p}(G)$ может оказаться более точной оценкой сверху для $\alpha(G)$, чем оценка $\alpha_C^*(G)$.

4. Тестовые эксперименты. С целью проверки свойств оценки $\alpha_{W_p}(G)$ алгоритм $LPWSTAB(p, \varepsilon)$ реализован на языке C++ программой $LPWSTAB$, для которой p и ε – входные параметры. Для решения ЛП-задачи использована программа $SOPLEX$ [4], а для нахождения кратчайшего пути в ориентированном графе – программа $dikh$ [5]. В таблице 1 приведены результаты экспериментов, которые получены с помощью программы $LPWSTAB$ при $\varepsilon = 0.01$, для 16-ти графов из $DIMACS$ -библиотеки [6].

Здесь оценка $\alpha_C^*(G)$ найдена программой $LPWSTAB$ при $p = 0$. Первыми в таблице 1 размещены те 9 графов из 16, для которых оценка $\alpha_{W_p}(G)$ оказалась точной для $\alpha(G)$. Для них в последнем столбце

Таблица 1. Эксперименты для $\alpha_{W_p}(G)$ для DIMACS-графов

DIMACS-графы	$ V $	$ E $	$\alpha(G)$	$\alpha_C^*(G)$	$\alpha_{W_p}(G)/p_1$	$\alpha_{W_p}(G)/p_2$	$\alpha_{W_p}(G)/p^*$
c-fat200-1	200	18366	12	66,667	28,571/4	16,667/9	12,000/14
c-fat200-2	200	16665	24	66,667	40,000/2	28,571/4	24,000/6
johnson 16-2-4	120	1680	8	40,000	17,142/4	10,909/8	8,000/12
johnson8-2-4	28	210	4	9,333	7,000/1	4,667/3	4,000/4
johnson8-4-4	70	560	14	23,333	17,500/1	—	14,000/2
hamming8-4	256	11776	16	83,333	32,000/5	21,333/9	16,000/13
san200.0.9-1	200	1990	70	70,000	—	—	—
san200-0.9-2	200	1990	60	66,667	—	—	60,000/1
san200.0.9-3	200	1990	44	66,667	50,000/1	—	44,000/2
keller4	171	5100	11	57,000	21,375/5	15,103/9	14,825/13
hamming6-4	64	1312	4	21,333	10,667/3	7,111/6	5,333/9
san200-0.7-2	200	5970	18	66,667	25,000/5	20,266/7	18,890/8
brock200-1	200	5066	21	66,667	50,000/1	40,000/2	38,581/3
c-fat200-5	200	11427	58	66,667	—	—	—
mann-a27	378	702	126	135,000	—	—	—
mann-a9	45	72	16	18,000	—	—	—

приведена эта точная оценка $\alpha_{W_p}(G)$ и значение $p = p^*$, при котором она достигнута. В двух предпоследних столбцах для этих 9 графов приведены найденные значения оценки α_{W_p} при некоторых значениях p , которые меньше, чем p^* . Они иллюстрируют динамику уточнения оценок для этих 9-ти графов с ростом p . Оценка $\alpha_C^*(G)$ в графе san200.0.9-1 является точной для $\alpha(G)$, поэтому ее уточнение с помощью p -колес не понадобилось. Вторыми в табл. 1 размещены 4 графа, для которых найденная наилучшая оценка $\alpha_{W_p}(G)$ не является точной. Здесь p^* – те значения p , при которых найдены эти наилучшие оценки, и точно так же отражена динамика улучшения оценок по мере увеличения p . Из таблицы 1 видно, что для всех этих графов оценка $\alpha_{W_p}(G)$ даже при сравнительно небольших значениях p намного точнее аппроксимирует сверху $\alpha(G)$, чем оценка $\alpha_C^*(G)$. Исключение составляют лишь 3 графа, которые размещены в таблице последними. Для них оценка $\alpha_C^*(G)$ не является точной и не улучшена с помощью использования p -колес. Но разрыв здесь между оценкой $\alpha_C^*(G)$ и $\alpha(G)$ сравнительно небольшой.

Ряд экспериментов с программой LPWSTAB связан с задачами нахождения оценок максимального объема помехоустойчивых кодов из [7]. Рассматривались все семейства графов для кодов корректирующих ошибки из указанного сайта [7]: **1dc** – Graphs From Single-Deletion-Correcting Codes; **2dc** – Graphs From Two-Deletion-Correcting Codes; **1tc** – Graphs From Codes For Correcting a Single Transposition (Excluding the End-Around Transposition); **1et** – Graphs From Codes For Correcting a Single Transposition (Including the End-Around Transposition) и **1zc** – Graphs From Codes For Correcting One Error on the Z-Channel (Also Called Codes For Correcting One Unidirectional or Asymmetric Error). Для всех указанных семейств графов с числом вершин от 64 до 512 найденные оценки $\alpha_{W_p}(G)$ приведены в табл. 2. Здесь ограничений на построение p -колес не делалось, и на основе найденных нечетных циклов строились максимально возможные p -колеса согласно алгоритму "последовательного включения наилучших вершин". Для сравнения в таблице 2 приведены числа Ловаса $\vartheta(G)$, которые для этих графов вычислены Брайаном Борчерсом (Brian Borchers) и в марте 2005 года размещены на сайте Н. Слоана. Из таблицы 2 видно, что самый большой разрыв между оценкой $\alpha_{W_p}(G)$ и числом Ловаса $\vartheta(G)$ (чуть меньше двух единиц) оказался для графов 1et.256 и 1et.512. Для всех остальных графов этот разрыв меньше. А для всех графов серии 1zc, связанных с корректированием единичной ошибки в Z -канале, программа LPWSTAB нашла верхние оценки $\alpha_{W_p}(G)$ для максимального объема кода, которые в точности совпали с числами Ловаса. Нахождение

Таблица 2. Сравнение $\theta(G)$ и $\alpha_{W_p}(G)$ для графов Н. Слоана

Граф G	$\alpha(G)$	$\vartheta(G)$	$\alpha_{W_p}(G)$	Граф G	$\alpha(G)$	$\vartheta(G)$	$\alpha_{W_p}(G)$
1dc.64	6	10,0000	10,0000	1tc.256	63	63,3999	63,6538
1dc.128	7	16,4188	16,9743	1tc.512	110	113,4002	114,6800
1dc.256	30	30,0000	30,2265	1et.64	18	18,8000	19,0000
1dc.512	52	53,0307	53,3163	1et.128	28	29,2309	30,0000
2dc.128	5	5,2424	5,6000	1et.256	50	55,1148	57,0060
2dc.256	7	7,4618	7,8188	1et.512	100	104,4204	106,4000
2dc.512	11	11,7678	12,5479	1zc.128	18	20,6667	20,6667
1tc.64	20	20,0000	20,0000	1zc.256	36	38,0000	38,0000
1tc.128	38	38,0000	38,0000	1zc.512	62	68,7500	68,7500

оценки $\alpha_{W_p}(G)$ для графа 1zc.512 (включает 512 вершин и 13824 ребер) программа LPWSTAB затратила $t = 177, 78sec.$ на процессоре AMD Athlon 1,81GHz. При этом пришлось решать ЛП-задачу 10 раз и находить кратчайшие пути в ориентированном графе (содержит 1024 вершин и 54296 дуг) 10×512 раз. Количество накопленных неравенств в ЛП-задаче, исключая реберные неравенства, при этом составило 1364 – из них 140 нечетных циклов, 1196 клик (получены как p -колеса на базе нечетных циклов C_3) и 28 p -колес.

Закключение. Вычислительные эксперименты с программой LPWSTAB для указанных графов показали, что алгоритм LPWSTAB(p, ϵ) характеризует сравнительно небольшое количество итераций (порядка нескольких десятков). Этому способствует групповое включение в ЛП-задачу нарушенных неравенств для нечетных циклов и p -колес в графе, что и гарантирует сравнительно небольшие затраты по времени при работе с графами, содержащими порядка нескольких сотен вершин. Однако, это порождает и некоторые проблемы, связанные с увеличением количества ограничений в ЛП-задаче. Поэтому, отсеив лишние линейных ограничений в ЛП-задаче на каждой итерации алгоритма LPWSTAB(p, ϵ) только ускорит время нахождения оценки $\alpha_{W_p}(G)$, и является одним из резервом для разработки более быстрых реализаций LPWSTAB-программы. Вторым резервом улучшения программы является замена простейшего учета p -колес в графе G на более усовершенствованные схемы построения p -колес. Это может увеличить точность ЛП-ориентированных верхних оценок, хотя конечно здесь сложно ожидать того эффекта, который наблюдается в случае перехода от нечетных циклов к p -колесам.

Литература

1. ГЭРИ М., ДЖОНСОН Д. Вычислительные машины и труднорешаемые задачи. – М.: Мир, 1982. – 416 с.
2. GROTSCHEL M., LOVASZ L., SCHRIJVER A. Geometric Algorithms and Combinatorial Optimization. – Springer-Verlag, Berlin. – 1988. – 362 p.
3. E.CHENG AND W.H.CUNNINGHAM Wheel inequalities for stable set polytopes // Mathematical Programming, Vol. 77 (1997), No. 3, pp. 389-421.

4. R. WUNDERLING Paralleler und objektorientierter Simplex-Algorithmus, Technische Universität Berlin, 1996, <http://www.zib.de/Publications/abstracts/TR-96-09/>
5. B.V. CHERKASSKY, A.V. GOLDBERG, AND T. RADZIK Shortest Paths Algorithms: Theory and Experimental Evaluation // Math. Prog. 73 (1996), pages 129–174, <http://www.avglab.com/andrew/soft.html>
6. DIMACS (1995), Cliques, coloring, and satisfiability: second DIMACS implementation challenge, <http://dimacs.rutgers.edu/Challenges/>
7. SLOANE N. Challenge problems: Independent set in graphs. – 2005. <http://www.reseach.att.com/~njas/doc/graphs.html>.

Об ЛП-ориентированных верхних оценках для взвешенного числа устойчивости графа¹

П.И. Стецюк, А.П. Лиховид

Кибернетика и системный анализ. – 2009. – № 1. – С. 157–170.

Рассматриваются верхние оценки для взвешенного числа устойчивости графа, которые базируются на аппроксимации многогранника устойчивых множеств с помощью линейных неравенств для нечетных циклов и p -колес в графе. Построены алгоритмы нахождения верхних оценок на основе решения задачи линейного программирования с конечным числом неравенств, которые получены на основе алгоритма кратчайших путей в специальном графе. Приведены результаты тестовых экспериментов, когда граф содержит от нескольких сотен до тысячи вершин.

Введение

Задача о максимальном независимом (устойчивом) множестве вершин графа является одной из центральных в теории графов. Она имеет много важных приложений, связанных с выбором в графе экстремальных множеств вершин с заданными свойствами. Так, например, непосредственное приложение этой задачи связано с нахождением максимального объема помехоустойчивых кодов (кодов, корректирующих ошибки при передаче информации). К задаче о максимальном независимом множестве вершин графа сводится задача о максимальной клике графа и задача о k -максимальной клике графа. Последние могут использоваться при нахождении взаимосвязанных подмножеств при информационном анализе социологических, биологических, телекоммуникационных и других массивов данных. Такие подмножества в социологических данных могут характеризовать родственные, криминальные или профессиональные связи; в телекоммуникациях - группы абонентов, которые часто общаются; в биологических данных нахождение таких подмножеств может означать наличие причинно-следственных связей при функционировании отдельных частей живого организма. С помощью задачи о

¹Работа выполнена при частичной финансовой поддержке гранта UKM2-2812-KV-06 (CRDF Cooperative Grants Programm)

максимальном независимом множестве вершин графа можно представить и классическую задачу о раскраске вершин неориентированного графа k -красками.

Задача о максимальном независимом множестве вершин графа для произвольного графа является NP-трудной. Более того, в общем случае не найдено полиномиального алгоритма, который бы гарантировал для нее сколь угодно большую фиксированную относительную погрешность по оптимальному значению целевой функции (его называют числом устойчивости графа). Поэтому актуален поиск верхних оценок, достаточно хорошо аппроксимирующих сверху число устойчивости графа.

В работе обсуждаются верхние оценки для взвешенного числа устойчивости графа, частным случаем которого является обычное число устойчивости графа. Они базируются на аппроксимации многогранника устойчивых множеств с помощью линейных неравенств для нечетных циклов и p -колес в графе. Алгоритм нахождения обсуждаемых верхних оценок используют решение задачи линейного программирования с конечным числом неравенств, при построении которых используется алгоритм Дейкстры для нахождения кратчайших путей. Практическая эффективность верхних оценок для числа устойчивости графа подтверждается результатами тестовых экспериментов для графов с несколькими сотнями и тысячами вершин.

Постановка задачи и ее сложность

Пусть $G=(V, E)$ – неориентированный граф (не содержащий петель) с множеством вершин $V(G)$ и множеством ребер $E(G)$. Для каждой вершины $i \in V(G)$ задан положительный вес w_i . Подмножество вершин $S \subseteq V(G)$ называется *устойчивым (или независимым)* множеством графа G , если для любых $i, j \in S$ ребро (i, j) не принадлежит $E(G)$. Взвешенное число устойчивости графа G определяется как $\alpha(G, w) = \max \sum_{i \in S} w_i$,

где $S \subseteq V(G)$ – устойчивое множество. Подмножество S^* , на котором достигается $\alpha(G, w)$, называется максимальным взвешенным устойчивым (или независимым) множеством графа G . В частном случае, когда все веса вершин в графе равны единице, имеем обычное число устойчивости графа G , которое принято обозначать $\alpha(G)$. Число устойчивости $\alpha(G)$ характеризует мощность максимального по числу входящих в него вершин устойчивого множества в графе G .

В общем случае задачи нахождения $\alpha(G, w)$ и $\alpha(G)$ принадлежат к NP-трудным задачам [1]. Пусть $STAB(G)$ – многогранник устойчивых

множеств (stable set polytope). Он является выпуклой оболочкой инцидентных векторов устойчивых множеств S в G и может быть представлен в следующем виде:

$$STAB(G) = \text{conv}\{x \in \{0, 1\}^{|V|} : x_i + x_j \leq 1 \forall (i, j) \in E(G)\}. \quad (1)$$

Нахождение $\alpha(G, w)$ связано с задачей максимизации линейной функции на выпуклом многограннике $STAB(G)$, которая имеет вид

$$\alpha(G, w) = \max \sum_{i \in V(G)} w_i x_i, \quad x \in STAB(G). \quad (2)$$

Максимум в задаче (2) достигается в одной или нескольких вершинах многогранника $STAB(G)$. В общем случае многогранник $STAB(G)$ может иметь сложную структуру, из-за чего задача (2) принадлежит к NP -трудным задачам. Поэтому теоретический и практический интерес представляет нахождение верхних оценок, достаточно хорошо аппроксимирующих сверху $\alpha(G, w)$.

Предметом обсуждения в статье будут верхние оценки, связанные с решением задач линейного программирования (ЛП-задач). Условимся такие оценки называть ЛП-ориентированными верхними оценками. В самом общем виде их можно описать следующим образом. Пусть $\mathcal{L}STAB(G)$ – некоторый многогранник, заданный с помощью системы линейных ограничений-неравенств, и пусть он аппроксимирует (сверху) многогранник $STAB(G)$. Тогда ЛП-ориентированная верхняя оценка $\alpha_{\mathcal{L}}^*(G, w)$ связана с решением следующей ЛП-задачи

$$\alpha_{\mathcal{L}}^*(G, w) = \max \sum_{i \in V(G)} w_i x_i, \quad x \in \mathcal{L}STAB(G).$$

В силу того, что многогранник $\mathcal{L}STAB(G)$ аппроксимирует (сверху) многогранник $STAB(G)$, имеем $\alpha_{\mathcal{L}}^*(G, w) \geq \alpha(G, w)$, и, следовательно, оценка $\alpha_{\mathcal{L}}^*(G, w)$ всегда является оценкой сверху для $\alpha(G, w)$. Свойства конкретной оценки $\alpha_{\mathcal{L}}^*(G, w)$ зависят от того, с помощью каких подмножеств линейных неравенств описывается многогранник $\mathcal{L}STAB(G)$.

Оценка $\alpha_{\mathcal{C}}^*(G, w)$ и алгоритмы LPCSTAB

Пусть C_{2k+1} , $k = 1, 2, \dots$ – нечетный цикл в графе G (содержит нечетное количество вершин). Для многогранника $STAB(G)$ справедливы следующие классы линейных неравенств:

$$0 \leq x_i \leq 1 \quad \forall i \in V(G), \quad (3)$$

$$x_i + x_j \leq 1 \quad \forall (i, j) \in E(G), \quad (4)$$

$$\sum_{i \in V(C_{2k+1})} x_i \leq k \quad \forall C_{2k+1} \in G. \quad (5)$$

Семейство неравенств (3) принято называть вершинными ограничениями (vertex constraints), семейство неравенств (4) (edge constraints) – реберными ограничениями, а семейство неравенств (5) – неравенствами нечетных циклов (odd-cycle constraints). Все вместе они определяют многогранник нечетных циклов (odd-cycle polytope)

$$CSTAB(G) = \left\{ x \in R^{|V|} : x \text{ удовлетворяет (3), (4) и (5)} \right\},$$

который аппроксимирует (сверху) многогранник $STAB(G)$.

Оценка $\alpha_C^*(G, w)$ связана с решением следующей ЛП-задачи

$$\alpha_C^*(G, w) = \max \sum_{i \in V(G)} w_i x_i, \quad x \in CSTAB(G). \quad (6)$$

Для произвольного графа G всегда имеет место неравенство $\alpha_C^*(G, w) \geq \alpha(G, w)$. Если граф G принадлежит семейству t -совершенных графов, для которых $STAB(G) = CSTAB(G)$, то оценка $\alpha_C^*(G, w)$ является точной для $\alpha(G, w)$, т.е. $\alpha_C^*(G, w) = \alpha(G, w)$.

Задача (6) в самом общем случае может содержать неполиномиальное количество ограничений, что обусловлено наличием ограничений вида (5) для всех возможных нечетных циклов. Несмотря на это, задача (6) разрешима за полиномиальное время. В [2] описан полиномиальный алгоритм ее решения, базирующийся на использовании метода эллипсоидов и на том факте, что для точки \bar{x} , которая удовлетворяет ограничениям (3),(4), за полиномиальное время можно либо убедиться, что она удовлетворяет ограничениям (5) для каждого нечетного цикла, либо найти такой нечетный цикл, для которого ограничение вида (5) является максимально нарушенным. Метод из работы [2] имеет теоретическую ценность, однако он малоприменим для практического нахождения оценок $\alpha_C^*(G, w)$.

Если граф содержит несколько сотен вершин, для вычисления оценки $\alpha_C^*(G, w)$ эффективные алгоритмы можно реализовать на основе современных ЛП-программ, для которых решение ЛП-задач с сотнями переменных и десятками или сотнями тысяч ограничений не представляет особых проблем. Эти алгоритмы назовем ЛП-ориентированными алгоритмами нахождения $\alpha_C^*(G, w)$ и будем их ориентировать на решение ЛП-задач с конечным числом линейных ограничений. В основу

ЛП-ориентированного алгоритма для нахождения $\alpha_C^*(G, w)$ положим полиномиальный алгоритм из [2] для нахождения максимального нарушенного ограничения, связанного с нечетным циклом, и ЛП-задачу:

$$f_C^* = \max \sum_{i \in V} w_i x_i \quad (7)$$

при ограничениях:

$$0 \leq x_i \leq 1 \quad \forall i \in V(G), \quad (8)$$

$$x_i + x_j \leq 1 \quad \forall (i, j) \in E(G), \quad (9)$$

$$\sum_{i \in V(C_{2k+1})} x_i \leq k \quad \forall C_{2k+1} \in C_{\text{odd}} \in G, \quad (10)$$

где C_{odd} – конечное множество нечетных циклов (возможно, пустое).

Тогда ЛП-ориентированный алгоритм для нахождения $\alpha_C^*(G, w)$ (назовем его *LPCSTABA*) состоит в следующем.

Начальная установка. Положим $itn = 0$ и $C_0 = \emptyset$ (множество нечетных циклов пусто). Перейдем к шагу 1.

Шаг 1. Имеем множество нечетных циклов C_{itn} . Положим $C_{\text{odd}} = C_{itn}$ и, решив ЛП-задачу (7)–(10), найдем f_C^* и $x^* = (x_1^*, \dots, x_{|V|}^*)$.

Шаг 2. Построим взвешенный неориентированный граф $G' = (V', E')$, где множество вершин $V'(G')$ включает $2|V|$ вершин (вершина i и ее копия i' для всех $i \in V(G)$). Множество ребер $E'(G')$ состоит из ребер, соединяющих все те пары вершин (i, j') и (i', j) , для которых пара вершин i и j соединена ребром (i, j) в графе G . Присвоим ребрам (i, j') и (i', j) один и тот же положительный вес (длина ребра), равный $1 - x_i^* - x_j^*$.

Шаг 3. Рассмотрим граф G' как ориентированный (ребру (i, j') соответствуют дуги (i, j') и (j', i) с длинами, равными длине ребра (i, j') , т.е. $l(i, j')$) и найдем для каждой вершины $i \in V$ кратчайший путь, который начинается в вершине i и заканчивается в вершине i' (копии вершины i). Число таких кратчайших путей равно $|V|$. Пусть длины найденных кратчайших путей равны l_i , $i = 1, \dots, |V|$.

Шаг 4. Среди найденных $|V|$ кратчайших путей выберем кратчайший путь с минимальной (наименьшей) длиной $l^* = \min_{i=1, 2, \dots, |V|} (l_i)$ (не обязательно единственный). Если длина l^* больше либо равна единице, то $\alpha_C^*(G, w) = f_C^*$ и останов (достаточное условие того, что точка

x^* удовлетворяет ограничениям в форме (5) для всех нечетных циклов). Иначе переходим к шагу 5.

Шаг 5. Полагаем $C' = C_{2k+1}$, где C_{2k+1} – нечетный цикл, соответствующий вершинам, через которые проходит кратчайший путь минимальной длины, с учетом замены вершин " i' " вершинами " i ". Нечетный цикл C' определяет ограничение в форме (5), которое в точке x^* является максимально нарушенным (оно не обязательно единственное). Перейдем к шагу 6.

Шаг 6. Положим $C_{itn+1} = C_{itn} \cup C'$ и $itn = itn + 1$. Перейдем к шагу 1.

Одна итерация алгоритма *LPCSTABa* требует решения ЛП-задачи (шаг 1) и нахождения $|V|$ раз кратчайшего пути (шаг 3). Если граф G содержит порядка нескольких сотен вершин, то добавление к ЛП-задаче ста нечетных циклов равносильно ста итерациям алгоритма и может потребовать значительных затрат по времени. Поэтому алгоритм *LPCSTABa* легко усовершенствовать, если использовать информацию трудоемкого по вычислениям шага 3 и добавить на каждой итерации все те нечетные циклы, для которых ограничение в форме (5) нарушено. Для модифицированного ЛП-ориентированного алгоритма нахождения $\alpha_C^*(G, w)$ (назовем его *LPCSTABb*) шаги 4 и 5 заменяются следующими.

Шаг 4а. Среди найденных V кратчайших путей оставляем только те, которые определяют нечетный цикл (такая проверка требуется, так как не каждый кратчайший путь определяет нечетный цикл в графе G). Пусть число кратчайших путей, определяющих нечетные циклы, равно n ($n \leq |V|$), и пусть длины этих кратчайших путей равны l_i , $i = 1, \dots, n$. Если все l_i , $i = 1, \dots, n$ не меньше единицы, то $\alpha_C^*(G, w) = f^*$ и останов. Иначе переходим к шагу 5а.

Шаг 5а. Сформируем список нечетных циклов C' , который включает все те нечетные циклы, которые соответствуют кратчайшим путям из $1, \dots, n$, для которых $l_i < 1$. Каждый из нечетных циклов из списка C' определяет ограничение в форме (5), которое в точке x^* является нарушенным (не обязательно максимально нарушенным, хотя по крайней мере один из таких нечетных циклов дает максимально нарушенное ограничение). Перейдем к шагу 6.

Оба алгоритма *LPCSTABa* и *LPCSTABb* позволяют найти оценку $\alpha_C^*(G, w)$, и если граф G принадлежит семейству t -совершенных графов, то оценка $\alpha_C^*(G, w)$ будет точной для $\alpha(G, w)$, т.е. $\alpha_C^*(G, w) =$

$\alpha(G, w)$. Однако, время работы каждого алгоритма будет разное. Для оценки времени нахождения $\alpha_C^*(G, w)$, когда граф содержит порядка нескольких сотен вершин, оба алгоритма были программно реализованы (программы *LPCSTABa* и *LPCSTABb*). Для решения ЛП-задачи использовалась программа SOPLEX [3], а для нахождения кратчайшего пути в ориентированном графе – программа *dikh* (из библиотеки *SPLIB*) [4].

Результаты тестовых экспериментов для сравнения времени работы обоих алгоритмов приведены в табл. 1. Вычисления проводились на процессоре AMD Athlon 1,81GHz. Верхняя оценка $\alpha_C^*(G)$ для взвешенного числа устойчивости графа ($w_i = 1, \forall i \in V(G)$) вычислена для ряда тестовых наборов из DIMACS-библиотеки [5].

Здесь *Nca* – максимальное количество включенных в ЛП-задачу нечетных циклов программой *LPCSTABa*. Оно же равно и количеству итераций, затраченных на нахождение оценки $\alpha_C^*(G)$. Для программы *LPCSTABb* приведено количество итераций (*itnb*). Из таблицы 1 легко видеть, что по времени программа *LPCSTABb* всегда выигрывает, иногда существенно, у программы *LPCSTABa* (столбец t_a/t_b), незначительно проигрывая в количестве накопленных нечетных циклов (столбец *Ncb/Nca*). Например, для графа *hamming8-4* достигнут выигрыш по времени в пятьдесят раз, нечетных циклов накапливается при этом не так и много (всего в два раза больше). При этом пришлось решать ЛП-задачу всего четыре раза, а находить кратчайшие пути в ориентированном графе с 256 вершинами и 47104 дугами всего 4×256 раз.

Относительно точности оценки $\alpha_C^*(G)$ для рассмотренных примеров отметим следующее. В трех случаях, графы *san200.0.9-1*, *hamming6-2* и *hamming8-2*, оценка $\alpha_C^*(G)$ оказалась точной верхней оценкой для $\alpha(G)$, для графов *hamming6-2* и *hamming8-2* она была точной при наличии только реберных ограничений в ЛП-задаче, и не было найдено ни одного нечетного цикла (в связи с этим t_a/t_b равно единице). Однако для остальных примеров оценка $\alpha_C^*(G)$ неточна, и для графов *c-fat200-1* и *hamming8-4* она сильно завышена. В следующих разделах рассматривается улучшенный вариант ЛП-ориентированного алгоритма, который позволит в ряде случаев найти более точные верхние оценки для числа устойчивости этих графов.

Таблица 1. Эксперименты для $\alpha_C^*(G)$

DMACS-примеры	$ V $	$ E $	$\alpha(G)$	$\alpha_C^*(G)$	NCa	$itnb$	t_a/t_b	NCb/NCa
c-fat200-1	200	18366	12	66,6667	241	4	43,5	2,2
c-fat200-2	200	16665	24	66,6667	232	4	40,1	2,4
c-fat200-5	200	11427	58	66,6667	237	4	36,1	2,2
johnson 16-2-4	120	1680	8	40,0000	143	4	29,7	2,1
johnson8-2-4	28	210	4	9,3333	32	4	5,8	2,1
johnson8-4-4	70	560	14	23,3333	81	4	14,5	2,1
keller4	171	5100	11	57,0000	211	4	35,1	2,2
hamming6-2	64	192	32	32,0000	0	1	1,0	-
hamming6-4	64	1312	4	21,3333	73	4	14,7	2,2
hamming8-2	256	1024	128	128,0000	0	1	1,0	-
hamming8-4	256	11776	16	85,3333	322	4	50,6	2,1
san200-0.7-2	200	5970	18	66,6667	232	5	30,7	3,3
san200.0.9-1	200	1990	70	70,0000	175	6	14,8	4,7
san200-0.9-2	200	1990	60	66,6667	274	12	12,9	4,7
san200.0.9-3	200	1990	44	66,6667	247	8	17,4	3,7
brock200-1	200	5066	21	66,6667	237	4	31,4	2,9
mann-a27	378	702	126	135,0000	146	7	15,4	1,6
mann-a9	45	72	16	18,0000	12	2	4,2	1,7

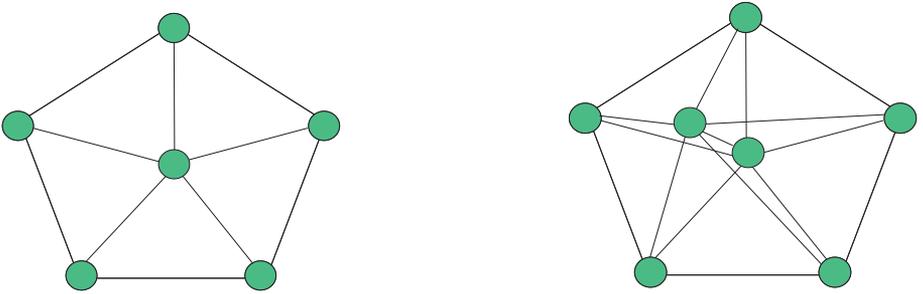


Рис. 1. Примеры 1-колеса и 2-колеса на базе нечетного цикла C_5

Оценка $\alpha_{W_p}^*(G, w)$ и W_p -совершенные графы

Рассмотрим более точную верхнюю оценку для $\alpha(G, w)$, чем оценка $\alpha_C^*(G, w)$, и обсудим ее связь с одной из верхних оценок, предложенной Н.З.Шором [6].

Пусть имеется граф W_{2k+1+p} , вершинами которого являются вершины непересекающихся нечетного цикла C_{2k+1} и клики Q_p (полный подграф, содержащий p вершин). Если $p = 1$, то клика состоит из одной вершины. Множество ребер в графе W_{2k+1+p} включает все ребра нечетного цикла C_{2k+1} , все ребра клики Q_p , а также ребра, связывающие каждую вершину C_{2k+1} со всеми вершинами клики Q_p . Граф W_{2k+1+p} принято называть p -колесом (p -wheel) [7]. Примеры 1-колеса и 2-колеса на базе нечетного цикла C_5 приведены на рис. 1,а,б, соответственно. С p -колесом связано семейство линейных неравенств

$$\sum_{i \in V(C_{2k+1})} x_i + k \sum_{j \in V(Q_p)} x_j \leq k, \quad \forall W_{2k+1+p} \in G, \quad (11)$$

которые справедливы для многогранника $STAB(G)$. Семейство неравенств (11) принято называть p -колесными ограничениями (p -wheel constraints). Они означают, что для каждого p -колеса из графа G в устойчивое (независимое) множество может быть включена либо одна из вершин клики Q_p , либо k вершин из нечетного цикла C_{2k+1} . Если семейство неравенств (11) добавить к линейным неравенствам, которые определяют многогранник $CSTAB(G)$, то получим многогранник

$$W_p STAB(G) = \left\{ x \in R^{|V|} : x \text{ удовлетворяет (3) - (11)} \right\},$$

который аппроксимирует (сверху) многогранник устойчивых множеств $STAB(G)$. Семейство графов, для которых $STAB(G) = W_p STAB(G)$, назовем W_p -совершенными (W_p -perfect).

Оценка $\alpha_{W_p}^*(G, w)$ связана с решением ЛП-задачи

$$\alpha_{W_p}^*(G, w) = \max \sum_{i \in V(G)} w_i x_i, \quad x \in W_p STAB(G). \quad (12)$$

Если граф G принадлежит семейству W_p -совершенных графов, то $\alpha_{W_p}^*(G, w) = \alpha(G, w)$. Для произвольного графа G всегда имеет место неравенство

$$\alpha_C^*(G, w) \geq \alpha_{W_p}^*(G, w) \geq \alpha(G, w), \quad (13)$$

т.е. оценка $\alpha_{W_p}^*(G, w)$ является верхней оценкой (оценкой сверху) для $\alpha(G, w)$ и она всегда не хуже, чем оценка $\alpha_C^*(G, w)$.

ЛП-задача (12) в общем случае может содержать неполиномиальное количество ограничений. В отличие от ЛП-задачи (6), она не может быть разрешена за полиномиальное время. Так, например, при $k = 1$ (т.е. рассматриваются только нечетные циклы, совпадающие с 3-кликкой) и произвольном p , $1 \leq p \leq |V| - 3$, из многогранника $W_p STAB(G)$ следует известный кликовый многогранник (clique polytope)

$$QSTAB(G) = \begin{cases} \sum_{i \in V(Q)} x_i \leq 1 & \text{для каждой клики } Q \in G, \\ 0 \leq x_i \leq 1 & \text{для каждой вершины } i \in G. \end{cases}$$

Здесь неравенства для 2-клик следуют из реберных неравенств (4), неравенства для 3-клик – из неравенств для нечетных циклов (5) при $k=1$, а неравенства для клик с большим количеством вершин – из неравенств для p -колес (11) при $k = 1$ и произвольном p . С многогранником $QSTAB(G)$ связана ЛП-ориентированная верхняя оценка

$$\alpha_Q^*(G, w) = \max \sum_{i \in V(G)} w_i x_i, \quad x \in QSTAB(G),$$

которая называется дробным взвешенным числом устойчивости графа G . В общем случае нахождение $\alpha_Q^*(G, w)$ есть NP-трудной задачей, поскольку это связано с нахождением максимальной клики в графе G , а такая задача для произвольного графа G является NP-трудной [1].

Однако когда граф G принадлежит семейству совершенных графов (для них $STAB(G) = QSTAB(G)$), то задача нахождения $\alpha_Q^*(G, w)$ разрешима за полиномиальное время. Однако, это объясняется свойствами не самой оценки $\alpha_Q^*(G, w)$, а свойствами более точной верхней оценки

для $\alpha(G, w)$, а именно известного числа Ловаса $\vartheta(G, w)$ [2]. Число Ловаса можно найти за полиномиальное время, и для произвольного графа G оно удовлетворяет неравенству

$$\alpha_Q^*(G, w) \geq \vartheta(G, w) \geq \alpha(G, w),$$

которое, если граф G принадлежит семейству совершенных графов, превращается в точное равенство

$$\alpha_Q^*(G, w) = \vartheta(G, w) = \alpha(G, w).$$

Отметим еще один интересный многогранник $WSTAB(G, w)$, с которым связано семейство W -совершенных (W -perfect) графов. Он получается из $W_pSTAB(G)$ при $p = 1$, т.е. неравенства для произвольного p колеса в форме (11) заменяются линейными неравенствами

$$\sum_{i \in V(C_{2k+1})} x_i + kx_{i_{2k+2}} \leq k, \quad \forall W_{2k+2} \in G,$$

для обычного колеса W_{2k+2} , когда клика состоит всего из одной вершины. Семейство этих неравенств принято называть колесными ограничениями (wheel constraints). С многогранником $WSTAB(G)$ связана верхняя оценка для $\alpha(G, w)$, являющаяся решением следующей ЛП-задачи

$$\alpha_W^*(G, w) = \max \sum_{i \in V(G)} w_i x_i, \quad x \in WSTAB(G).$$

Нахождение оценки $\alpha_W^*(G, w)$ может быть осуществлено за полиномиальное время для произвольного графа G [2]. Если граф G принадлежит семейству W -совершенных (W -perfect) графов (для них $STAB(G) = WSTAB(G)$), то оценка $\alpha_W^*(G, w)$ является точной для $\alpha(G, w)$.

Итак, для совершенных, t -совершенных и W -совершенных графов задача нахождения $\alpha(G, w)$ разрешима за полиномиальное время. Поэтому интерес представляет следующий вопрос: будет ли полиномиально разрешимой задача нахождения $\alpha(G, w)$ для семейства W_p -совершенных графов? Оказывается – будет. И это объясняется свойствами одной из оценок Н.З.Шора, которая для произвольного графа G является более точной верхней оценкой для $\alpha(G, w)$, чем оценка $\alpha_{W_p}^*(G, w)$.

Оценка Н.З.Шора (условимся ее обозначать $\psi_1^*(G, w)$) есть оптимальной (наилучшей) лагранжевой (двойственной) оценкой для квадратичной булевой задачи

$$\alpha(G, w) = \max \sum_{i \in V} w_i x_i \tag{14}$$

при ограничениях

$$x_i x_j = 0 \quad \forall (i, j) \in E(G), \quad (15)$$

$$x_i^2 - x_i = 0 \quad \forall i \in V(G), \quad (16)$$

$$x_i x_k + x_j x_k \leq x_k, \quad \forall (i, j) \in E(G), k \neq i, j, \quad (17)$$

в виде которой формулируется задача о максимальном взвешенном устойчивом множестве графа. Оценку $\psi_1^*(G, w)$ можно найти с любой заданной точностью с помощью методов минимизации негладких выпуклых функций [6]. Здесь булева переменная $x_i \in \{0, 1\}$ равна единице, если вершина $i \in V$ включается в устойчивое множество, и равна нулю в противном случае. Булевы переменные для всех вершин описаны квадратичными ограничениями равенствами (16). Квадратичные ограничения (15) означают, что если две вершины связаны ребром в графе G , то они обе не могут одновременно принадлежать устойчивому множеству. Квадратичные ограничения в форме неравенств (17) функционально избыточны, т.е. они не изменяют множество оптимальных решений квадратичной задачи (14)–(16). Они получаются как результат умножения реберных ограничений в виде (4) на переменные x_k , $k \neq i, j$. Знак неравенств не изменится в силу того, что $x_k = x_k^2 \geq 0$, и в результате получаем квадратичные неравенства (17).

Именно наличие ограничений (17) придает оценке $\psi_1^*(G, w)$ ряд замечательных свойств, которые связаны со специальными семействами графов. Так, например, в [6] показано, что верхняя оценка $\psi_1^*(G, w)$ является точной для $\alpha(G, w)$, когда граф G t -перфектный. Схема доказательства основана на том, что ослаблением (релаксацией) квадратичных ограничений (15)–(17) легко получить ЛП-задачу для оценки $\alpha_C^*(G, w)$. Однако из квадратичных ограничений (15)–(17) следуют неравенства в форме (11) [8]. Это означает, что для произвольного графа G справедливо неравенство

$$\alpha_{W_p}^*(G, w) \geq \psi_1^*(G, w) \geq \alpha(G, w). \quad (18)$$

Для W_p -совершенных графов неравенство (18) превращается в строгое равенство

$$\alpha_{W_p}^*(G, w) = \psi_1^*(G, w) = \alpha(G, w), \quad (19)$$

которое означает, что оценка $\psi_1^*(G, w)$ для них является точной верхней оценкой для $\alpha(G, w)$. Учитывая, что оценку $\psi_1^*(G, w)$ можно найти за полиномиальное время, получаем, что если граф G принадлежит семейству W_p -совершенных графов, то задача нахождения $\alpha(G, w)$ полиномиально разрешима.

Однако ответ, который дает равенство (19), носит скорее теоретический характер, чем практический. Вычисление оценки $\psi_1^*(G, w)$ связано с решением задачи минимизации выпуклой негладкой функции, определенной на параметрически заданном семействе неотрицательно определенных симметричных матриц размера $|V| \times |V|$, линейно зависящих от неизвестных множителей Лагранжа (соответствуют ограничениям (15), (16) и (17)). Количество неизвестных множителей Лагранжа имеет порядок $O(|V|^3)$, и когда граф содержит несколько сотен вершин, достаточно точное нахождение оценки $\psi_1^*(G, w)$ является практически неразрешимой задачей для современных методов недифференцируемой оптимизации.

Оценка $\alpha_W(G, w)$ и алгоритм LPWSTAB

Нахождение оценки $\alpha_{W_p}^*(G, w)$ – NP-трудная задача. Более того, NP-трудной является и задача нахождения на основе уже имеющегося нечетного цикла такого p -колеса, для которого в точке $x^* = \{x_1^*, \dots, x_{|V|}^*\}$ максимально нарушается линейное неравенство в виде (11). Поэтому вместо оценки $\alpha_{W_p}^*(G, w)$ рассмотрим ослабленную ЛП-ориентированную верхнюю оценку для $\alpha(G, w)$ на основе линейных неравенств для таких p -колес в графе G , которые легко построить. Эту оценку назовем оценкой $\alpha_W(G, w)$ и в основу алгоритма для ее нахождения положим следующее.

На каждой итерации алгоритма будем решать ЛП-задачу в следующей форме:

$$f_W^* = \max \sum_{i \in V} w_i x_i \quad (20)$$

при ограничениях:

$$0 \leq x_i \leq 1 \quad \forall i \in V(G), \quad (21)$$

$$x_i + x_j \leq 1 \quad \forall (i, j) \in E(G), \quad (22)$$

$$\sum_{i \in V(C_{2k+1})} x_i \leq k \quad \forall C_{2k+1} \in C_{odd} \in G, \quad (23)$$

$$\sum_{i \in V(C_{2k+1})} x_i + k \sum_{j \in V(Q_p)} x_j \leq k, \quad \forall W_{2k+1+p} \in W_{odd} \in G. \quad (24)$$

Здесь C_{odd} и W_{odd} – конечные множества нечетных циклов и p -колес (возможно, пустые). Для формирования множества нечетных циклов применим алгоритм нахождения нечетных циклов из [2], используя $x^* =$

$\{x_1^*, \dots, x_{|V|}^*\}$ – оптимальное решение ЛП-задачи (20)–(24). Для формирования множества p -колес используем уже найденные нечетные циклы и дополним их до p -колес с помощью простейшего по количеству вычислений алгоритма (по типу "жадного" алгоритма). Такой алгоритм назовем алгоритмом "последовательного включения наилучших вершин".

Пусть i_1, \dots, i_{2k+1} – вершины, через которые проходит нечетный цикл C_{2k+1} и вектор x^* – оптимальное решение ЛП-задачи (20)–(24). Чтобы на основе нечетного цикла C_{2k+1} построить p -колесо в алгоритме "последовательного включения наилучших вершин" требуется выполнение следующей последовательности действий. Найдем подмножество вершин из $V(G)$, для которого каждая из входящих в него вершин связана со всеми вершинами нечетного цикла C_{2k+1} . Если такое подмножество вершин непусто, то выберем из него ту вершину, для которой значение компоненты вектора x^* максимально, и добавим его к множеству вершин с номером $(2k+2)$. В результате получим 1-колесо с множеством вершин $i_1, \dots, i_{2k+1}, i_{2k+2}$. Для него точно также найдем подмножество вершин из $V(G)$, для которого каждая из входящих в него вершин связана со всеми вершинами 1-колеса. Если подмножество вершин непусто, то выберем вершину с максимальным значением компоненты вектора x^* и добавим к множеству вершин с номером $(2k+3)$. В результате получим 2-колесо с множеством вершин $i_1, \dots, i_{2k+1}, i_{2k+2}, i_{2k+3}$. Для него применим такую же процедуру и либо получим 3-колесо, либо прервем процесс, если 3-колеса не существует (подмножество вершин из $V(G)$, связанных с каждой из вершин 2-колеса окажется пустым). Пусть 3-колесо получено. Будем продолжать процесс далее до тех пор, пока подмножество вершин из $V(G)$, связанных с каждой из вершин уже найденного p -колеса, не окажется пустым.

Алгоритм для нахождения оценки $\alpha_W(G, w)$ (назовем его *LPWSTAB*) состоит в следующем.

Начальная установка. Положим $itn = 0$ и $C_0 = \emptyset$, $W_0 = \emptyset$ (множества нечетных циклов и p -колес пустые). Перейдем к шагу 1.

Шаг 1. Имеем множество нечетных циклов C_{itn} и множество p -колес W_{itn} . Положим $C_{odd} = C_{itn}$, $W_{odd} = W_{itn}$ и, решив ЛП-задачу (20)–(24), найдем f_W^* и $x^* = (x_1^*, \dots, x_{|V|}^*)$.

Шаг 2. Построим взвешенный неориентированный граф $G' = (V', E')$. Множество вершин $V'(G')$ включает вершину i и ее копию i' для всех $i \in V(G)$. Множество ребер $E'(G')$ состоит из ребер, которые соединяют все те пары вершин (i, j') и (i', j) , для которых пара вершин i и j соединена ребром (i, j) в графе G . Присвоим ребрам

(i, j') и (i', j) один и тот же положительный вес (длина ребра) $l(i', j) = l(i, j') = 1 - x_i^* - x_j^*$.

Шаг 3. Рассмотрим граф G' как ориентированный (ребру (i, j') соответствуют дуги (i, j') и (j', i) с длинами, равными длине ребра (i, j') , т.е. $l(i, j')$), и для каждой вершины $i \in V$ найдем кратчайший путь, который начинается в вершине i и заканчивается в вершине i' (копии вершины i). Число таких кратчайших путей равно $|V|$. Из них оставляем только те, которые определяют нечетный цикл. Пусть число таких кратчайших путей равно n ($n \leq |V|$). Найдем нечетные циклы $C_{2k_1+1}, \dots, C_{2k_n+1}$ (они соответствуют n найденным кратчайшим путям).

Шаг 4. Установим множества C' и W' пустыми и заполним их следующим образом. Каждый из нечетных циклов C_{2k_i+1} для $i = 1, \dots, n$ с помощью алгоритма "последовательного включения наилучших вершин" дополняем до p -колеса в графе G . Пусть такое p -колесо получено. Если линейное неравенство в форме (11) для него нарушено, то найденное p -колесо включаем в множество W' . Если построить p -колесо на базе нечетного цикла не удалось, но линейное неравенство в форме (5) для этого нечетного цикла является нарушенным, то такой нечетный цикл включаем в множество C' . Если после завершения этой процедуры для всех $i = 1, \dots, n$ оба множества C' и W' пустые, то $\alpha_W(G, w) = f_W^*$ и останов. Иначе перейдем к шагу 5.

Шаг 5. Положим $C_{itn+1} = C_{itn} \cup C'$ и $W_{itn+1} = W_{itn} \cup W'$, $itn = itn + 1$. Перейдем к шагу 1.

ЛП-задача (20)–(24) построена как расширение ЛП-задачи (7)–(10) за счет конечного набора справедливых для p -колес линейных неравенств в виде (24). Поэтому оценку $\alpha_W(G, w)$ можно рассматривать как улучшение оценки $\alpha_C^*(G, w)$ с помощью использования линейных неравенств для p -колес в графе G . В результате она будет сохранять свойства оценки $\alpha_C^*(G, w)$, а в случае обнаружения p -колес, для которых линейные неравенства в форме (11) являются нарушенными, оценка $\alpha_W(G, w)$ может оказаться более точной верхней оценкой для $\alpha(G, w)$, чем оценка $\alpha_C^*(G, w)$.

С целью проверки свойств оценки $\alpha_W(G, w)$ алгоритм реализован программой *LPWSTAB* на языке C++. Для решения ЛП-задачи использована программа *SOPLEX* [3], а для нахождения кратчайшего пути в ориентированном графе – программа *dikh* [4]. В программе *LPWSTAB*

используется параметр ε , который означает, что линейные неравенства, связанные с нечетным циклом C_{2k+1} и p -колесом W_{2k+1+p} , считаются нарушенными (в точке x^*) и включаются в ЛП-задачу, если для них выполняются условия:

$$\sum_{i \in V(C_{2k+1})} x_i^* \geq k + \varepsilon \quad \text{и} \quad \sum_{i \in V(C_{2k+1})} x_i^* + k \sum_{j \in V(Q_p)} x_j^* \geq k + \varepsilon.$$

Первая серия экспериментов с программой *LPWSTAB* заключалась в проверке точности оценки $\alpha_W(G, w)$, где $w_i = 1, \forall i \in V(G)$, для того же набора *DIMACS*-графов, для которых вычислялась оценка $\alpha_C^*(G)$. Рассматривались 16 графов. Исключены только два графа *hamming6-2* и *hamming8-2*, для которых не требовалось включения ни одного нечетного цикла. Для программы *LPWSTAB* использовалось значение $\varepsilon = 0.01$. Результаты эксперимента отражены в таблице 2. Здесь для каждого из 16 графов приведены значение оценки $\alpha_W(G)$ и затраты на ее нахождение – количество итераций (столбец *itn*) и время в секундах на процессоре AMD Athlon 1,81GHz (столбец *t_W*). Остальные столбцы в таблице характеризуют структуру накопленных программой *LPWSTAB* (включенных в ЛП-задачу на последней итерации) линейных неравенств, исключая "реберные" неравенства. Здесь столбец *N* задает полное количество включенных в ЛП-задачу линейных неравенств, а в столбцах *NC*, *NW* и *NQ* указано, сколько среди них было нечетных циклов (сюда же включены и 3-клики), p -колес с $k \geq 2$ и клик, которые получены как p -колеса при $k = 1$ и произвольных p .

Из таблицы 2 видно, что оценка $\alpha_W(G)$ является точной оценкой для $\alpha(G)$ для девяти графов из 16. Структура накопленных неравенств для графа *san200.0.9-1* позволяет утверждать, что он не может быть t -совершенным, хотя для него оценка $\alpha_C^*(G)$ является точной верхней оценкой для $\alpha(G)$ (см. таблицу 1). Такой вывод дают возможность сделать 16 накопленных 4-клик (см. столбец *NQ* в таблице 2), и, следовательно, для графа *san200.0.9-1* можно подобрать такие веса вершин, при которых оценка $\alpha_C^*(G, w)$ не будет точной для $\alpha(G, w)$.

Затраты по времени нахождения оценки $\alpha_W(G)$ превышают (иногда существенно) затраты по времени на нахождение оценки $\alpha_C^*(G)$ программой *LPCSTABb*. Одну из причин этого объясняют накопленные неравенства для графа *san200.0.9-1*, что подсказывает путь для усовершенствования алгоритма *LPWSTAB*. Так для оценки $\alpha_W(G)$ накоплено 663 неравенства (из них нечетных циклов – 646, и 4-клик – 16). Их меньше, чем количество накопленных нечетных циклов программой *LPCSTABb* (равно 828) при той же точности $\varepsilon = 0.01$, и суще-

Таблица 2. Эксперименты для $\alpha_W(G)$

DMACS-примеры	$ V $	$\alpha(G)$	$\alpha_W(G)$	itn	t_W	N	NC	NW	NQ
c-fat200-1	200	12	12,00	31	485,8	5787	0	0	5787
c-fat200-2	200	24	24,00	17	219,5	3342	0	12	3330
c-fat200-5	200	58	66,66	3	29,0	533	533	0	0
johnson 16	120	8	8,00	3	2,3	128	36	7	85
johnson8-2	28	4	4,00	3	0,7	41	8	7	26
johnson8-4	70	14	14,00	3	1,0	139	12	4	123
keller4	171	11	14,82	14	32,8	1151	41	36	1074
hamming6-4	64	4	5,33	5	2,1	301	0	16	285
hamming8-4	256	16	16,00	5	54,8	1144	48	48	1048
san200-0.7	200	18	19,35	18	58	1569	134	1	1434
san200.0.9-1	200	70	70,00	4	6,4	663	646	1	16
san200-0.9-2	200	60	60,00	11	16	1369	1299	0	70
san200.0.9-3	200	44	44,00	21	23,4	1526	1378	0	148
brock200-1	200	21	38,58	32	88,5	1543	227	28	1288
mann-a27	378	126	135,00	7	5,1	245	245	0	0
mann-a9	45	16	18,00	1	0,5	21	21	0	0

ственно больше, чем число нечетных циклов, накопленных программой *LPCSTABa* (равно 175, см. табл. 1). Это означает, что процедура включения неравенств в ЛП-задачу в программе *LPWSTAB* требует доработки и дополнительного исследования, с целью уменьшить количество включаемых в ЛП-задачу нарушенных неравенств. Так, например, можно устроить их ранжирование по точности нарушения неравенств, ограничить количество включаемых неравенств на одной итерации. Создание более конструктивного критерия для включения нарушенных неравенств может способствовать и уменьшению общего количества итераций *LPWSTAB*-алгоритма, что может сделать затраты по времени нахождения оценки $\alpha_W(G, w)$ более близкими к затратам на нахождение оценки $\alpha_C^*(G, w)$ с помощью алгоритма *LPCSTABb*.

Кроме тестовых примеров из DIMACS-библиотеки программа *LPWSTAB* проверена на тестовых задачах для помехоустойчивых кодов из сайта², для которых получены помехозащищенные коды максимального объема [9]. Например, для всех тестовых задач, связанных с корректированием единичной ошибки в Z -канале, программа *LPWSTAB* нашла верхние оценки для максимального объема кода, которые совпали с числом Ловаса. Для тестовой задачи, где граф включает 1024 вершин и 33280 ребер, программа затратила время $t = 1139sec.$ на процессоре AMD Athlon. При этом пришлось решать ЛП-задачу 12 раз и находить кратчайшие пути в ориентированном графе (содержит 2048 вершин и 133120 дуг) 12×1024 раз. Количество накопленных неравенств при этом составило 2768 – из них 249 нечетных циклов, 2468 клик и 51 p -колес.

Вторая серия экспериментов для программы *LPWSTAB* связана с ее выходом, а именно с ЛП-задачей, которая соответствует оценке $\alpha_W(G)$. Целью экспериментов было выяснить: насколько быстрее с помощью программ для решения задачи целочисленного линейного программирования можно решить булеву ЛП-задачу, подготовленную программой *LPWSTAB*, чем булеву ЛП-задачу с реберными ограничениями (включает только линейные неравенства для всех ребер графа). Для этой цели была выбрана свободно распространяемая программа *SCIP* (Solving Constraint Integer Programming) [10]. Для тех же наборов *DIMACS*-графов результаты экспериментов отражены в табл. 3. Здесь N_{SCIP} – количество переменных в обеих булевых ЛП-задачах (равно количеству вершин графа), M_{ESCIPI} – количество ограничений в булевой ЛП-задаче с реберными ограничениями (равно количеству вершин и ребер в графе), M_{WSCIPI} – количество ограничений в буле-

²<http://www.research.att.com/njas/doc/graphs.html>

Таблица 3. Эксперименты для программы SCIP

DIMACS-примеры	N_{SCIP}	M_{ESCIIP}	t_{ESCIIP}	$M_{W_{SCIP}}$	M_{SCIP}	$t_{W_{SCIP}}$	$\frac{t_{ESCIIP}}{t_{W_{SCIP}}}$
c-fat200-1	200	18366	1041	24353	5420	125	8,33
c-fat200-2	200	16665	410	20207	5284	91	4,51
c-fat200-5	200	11427	1194	12160	10831	1472	0,81
johnson 16-2-4	120	1680	1,41	1928	48	0,14	10,07
johnson8-2-4	28	210	0,04	237	16	0,03	1,33
johnson8-4-4	70	560	1,64	769	135	0,09	18,22
keller4	171	5100	829	6422	1155	259	3,20
hamming6-4	64	1312	10,1	1677	209	2,31	4,39
hamming8-4	256	11776	746	13176	1324	26,04	28,00
san200-0.7-2	200	5970	410	7739	2213	112	3,66
san200-0.9-1	200	1990	5,5	2853	2312	2,39	2,30
san200-0.9-2	200	1990	13,3	3559	2638	4,09	3,25
san200-0.9-3	200	1990	111,9	3716	2562	346	0,32
brock200-1	200	5066	22766	6809	2515	8065	2,82
mann-a27	378	702	359	1325	597	55	6,50
mann-a9	45	72	0,21	138	58	0,26	0,91

вой ЛП-задаче для оценки $\alpha_W(G, w)$ (равно сумме количества вершин и ребер в графе и накопленных программой LPWSTAB линейных неравенств), M_{SCIP} – количество ограничений в булевой ЛП-задаче, которое осталось после препроцессинга программы *SCIP* (задача с этим количеством ограничений является стартовой для метода ветвей и границ, реализованной в программе *SCIP*), t_{ESCI} и t_{WSCI} – времена решения программой *SCIP* булевой ЛП-задачи с реберными ограничениями и булевой ЛП-задачи для оценки $\alpha_W(G)$ соответственно. Последний столбец характеризует отношение затрат по времени решения обеих булевых ЛП-задач: указано, во сколько раз быстрее программа *SCIP* решила булеву ЛП-задачу для оценки $\alpha_W(G)$, чем булеву ЛП-задачу для реберных ограничений.

Из табл. 3 видно, что булева ЛП-задача для оценки $\alpha_W(G)$ проиграла только в четырех случаях: незначительно для графов *s-fat200-5*, *san200-0.9-2* и *mapn-a9* и существенно (почти в три раза) для графа *san200.0.9-3*. В остальных случаях булева ЛП-задача для оценки $\alpha_W(G)$ выиграла, и даже очень существенно (более, чем в 20 раз для графа *hamming8-4*, более, чем в 8 раз для графа *s-fat200-1*, более, чем в четыре раза для графа *s-fat200-2*, почти в 4 раза для графа *keller4*, более, чем в два раза для графа *san200-0.7-2*, более, чем в 6 раз для графа *mapn-a27*). Наиболее существенный выигрыш по реальному времени получился для графа *brock200-1* – 2 часа 13 минут против 6 часов 19 минут. Для графов *johnson 16-2-4*, *johnson8-2-4* и *johnson8-4-4* как незначительный, так и существенный выигрыш по времени ничего не значат. Для них общее время решения обеих булевых ЛП-задач очень мало. В этом случае представляет интерес количество линейных неравенств, которые оставлены в булевых ЛП-задачах после препроцессинга. Оно невелико по сравнению с количеством вершин в этих графах, и это объясняется нахождением алгоритмом LPWSTAB достаточно "хороших" кликовых ограничений, которые сравнительно небольшим количеством покрывают вершины графа.

С помощью программы *SCIP* и подготовленной программой LPWSTAB ЛП-задачи для оценки $\alpha_W(G)$ найдена точная верхняя оценка для максимального объема кода, корректирующего одно удаление бита, для графа *1dc1024*. Здесь затраты по времени составили 55 часов на процессоре AMD Athlon 1,81GHz. Это значительно меньше, чем время, указанное на сайте³. Оно равно 298 часов и было достигнуто с помощью метода ветвей и границ и оценок Ловаса на мощном по сравнению с нашим компьютером IBM P690. Для графа *1dc1024* оценка $\alpha_W(G)$ рав-

³<http://www.reseach.att.com/njas/doc/graphs.html>

на 96.412 (точное решение $\alpha(G) = 94$). Программа LPWSTAB накопила 3483 неравенства (из них 276 – нечетных циклов, 3154 – клик, 53 – p -колес) и затратила 1128,7 секунд. Булева ЛП-задача содержала 28570 линейных неравенств до препроцессинга и 2162 линейных неравенства после препроцессинга программы SCIP.

При работе с большими графами (порядка тысяч вершин) проявился основной недостаток алгоритма LPWSTAB. Он связан с тем, что лежащая в его основе ЛП-задача только накапливает линейные ограничения, не отсеивая заведомо лишние. Из-за этого в программе SCIP значительное время занимает препроцессинг ЛП-задачи. В результате в булевой ЛП-задаче для оценки $\alpha(G, w)$ остается намного меньше ограничений. Главным образом по этой причине приостановлены попытки решать задачи, связанные с графами больших размеров. Очевидно, что отсеивание лишних линейных ограничений в алгоритме LPWSTAB только ускорит время нахождения оценки $\alpha_W(G, w)$, и это является резервом для разработки более быстрых реализаций LPWSTAB-программы. Так, например, простейшие усовершенствования здесь очевидны: те реберные неравенства, для которых пара вершин входит в какую-либо из клик для ограничений, входящих в ЛП-задачу, следует отбросить. Более тонкие отсеивания связаны уже с анализом двойственных переменных для ЛП-задачи на каждой из итераций LPWSTAB-алгоритма.

Заключение

Отметим ряд важных моментов для задачи нахождения взвешенного числа устойчивости графа $\alpha(G, w)$, которые следуют из вычислительных экспериментов для оценок $\alpha_C^*(G, w)$ и $\alpha_W(G, w)$:

1. Современные общедоступные ЛП-программы позволяют создавать для современных ПЭВМ практически пригодные (по времени) алгоритмы для нахождения ЛП-ориентированных верхних оценок для $\alpha(G, w)$, для графов, содержащих несколько сотен вершин. Это подтверждают тестовые эксперименты с алгоритмами LPCSTABb и LPWSTAB;
2. Точность ЛП-ориентированных верхних оценок существенно зависит от использования того набора неравенств, с помощью которых аппроксимируется многогранник $STAB(G)$. Например, даже простейший учет p -колес в графе G позволяет очень много сделать для улучшения точности ЛП-ориентированной оценки по сравне-

нию с использованием только нечетных циклов. Усовершенствованные схемы построения p -колес могут усилить точность ЛП-ориентированных верхних оценок;

3. На основе современных ЦЛП-программ (целочисленного линейного программирования) можно получить и обосновать точную верхнюю границу для $\alpha(G, w)$, и более того, найти одно из булевых решений, на котором достигается эта граница. При этом качество ЛП-ориентированных оценок играет важную роль для ускорения работы ЦЛП-программ. Это доказывают численные эксперименты с программой SCIP.

Можно ли расширить границы применимости ЛП-ориентированных оценок на случай графов с несколькими тысячами вершин, используя коммерческие ЛП-программы и ЦЛП-программы? Ответ требует более детального исследования. Пример с графом 1dc1024 указывает на то, что с помощью программы CPLEX доказательство максимального объема помехоустойчивого кода для этого графа потенциально можно осуществить за время $t * 0.08$, где t – время доказательства с помощью программы SCIP. Это следует из того, что работа программы SCIP оценивается по времени в соотношении 1 к 0.08 по отношению к промышленному варианту программы CPLEX⁴. При этом можно ожидать, что на процессоре AMD Athlon 1,81GHz обоснование точности верхней оценки максимального объема помехоустойчивого кода для графа 1dc1024 можно осуществить с помощью CPLEX за время, меньшее пяти часов (сравн.: 55 часов с помощью программы SCIP). Подтвердится ли это предположение? Ответ на этот вопрос может дать только численный эксперимент.

Литература

1. ГЭРИ М., ДЖОНСОН Д. Вычислительные машины и труднорешаемые задачи. – М.: Мир, 1982. – 416 с.
2. GROTSCHEL M., LOVASZ L., SCHRIJVER A. Geometric Algorithms and Combinatorial Optimization. – Berlin, Springer-Verlag. – 1988. – 362 p.
3. WUNDERLING R. Paralleler und objektorientierter Simplex-Algorithmus, Technische Universität Berlin, 1996, <http://www.zib.de/Publications/abstracts/TR-96-09/>

⁴<http://scip.zib.de>

4. CHERKASSKY B.V., GOLDBERG A.V., AND RADZIK T. Shortest Paths Algorithms: Theory and Experimental Evaluation // Math. Prog. 73 (1996), pages 129–174, <http://www.avglab.com/andrew/soft.html>
5. DIMACS (1995), Cliques, coloring, and satisfiability: second DIMACS implementation challenge, <http://dimacs.rutgers.edu/Challenges/>
6. SHOR N.Z. Nondifferentiable Optimization and Polynomial Problems. – Boston; Dordrecht; London: Kluwer Acad. Publ., 1998. – 412 p.
7. CHENG E. AND CUNNINGHAM W.H. Wheel inequalities for stable set polytopes // Mathematical Programming, Vol. 77 (1997), No. 3, pp. 389–421.
8. СТЕЦЮК П.И., ЧУМАКОВ Б.М. О свойствах одной верхней оценки Н.З.Шора для взвешенного числа устойчивости графа // Праці міжнародного симпозіума "Питання оптимізації обчислень (ПОО-XXXIII)". – К.: Ін-т кібернетики ім. В.М.Глушкова НАН України, 2007. – С.271–272.
9. СЕРГИЕНКО И.В., ШИЛО В.П. Проблемы дискретной оптимизации: сложные задачи, основные подходы к их решению // Кибернетика и системный анализ. – 2006. – №4. – С. 3–25.
10. ACHTERBERG T. Constraint Integer Programming, Technische Universität Berlin, 2007, <http://opus.kobv.de/tuberlin/volltexte/2007/1611/>

Поступила 03.03.2008

Оценки Н.З. Шора в квадратичных экстремальных задачах и их применение в комбинаторной оптимизации

П.И. Стецюк

Российская конференция "Дискретная оптимизация и исследование операций": Материалы конференции (Алтай, 27 июня – 3 июля 2010). – Новосибирск: Изд-во Ин-та математики, 2010. – С. 60–64.

Многие из задач комбинаторной оптимизации могут быть сформулированы с помощью многоэкстремальных квадратичных моделей с булевыми переменными. Условие булевости переменной $x \in \{0, 1\}$ представляется квадратичным равенством $x^2 - x = 0$. При разработке алгоритмов решения таких задач можно использовать предложенную Н.З. Шором [1,2] методику лагранжевых двойственных оценок для оптимального значения целевой функции в экстремальной квадратичной задаче. Условимся такие оценки называть оценками Шора. В квадратичных задачах на минимум они будут оценками снизу для минимального значения целевой функции, а в квадратичных задачах на максимум – оценками сверху для максимального значения целевой функции.

Методика верхних (нижних) оценок Шора включает два важных аспекта. Первый аспект связан с алгоритмами нахождения оценок с помощью методов минимизации недифференцируемых выпуклых функций. Второй аспект связан с использованием функционально избыточных квадратичных ограничений (не изменяют множества допустимых решений исходной квадратичной задачи) для улучшения точности оценок. Он придает оценкам Шора замечательное свойство, благодаря которому можно обосновать алгоритмы решения ряда комбинаторных задач за полиномиальное время.

В работе рассмотрим методику верхних оценок Шора на примере квадратичной задачи на максимум (с целью упрощения будем рассматривать только ограничения-равенства). Дадим примеры двух верхних оценок для NP-трудной задачи о числе независимости неориентированного графа.

1. Оценка Ψ^* . Пусть R^n – евклидово пространство со скалярным произведением (\cdot, \cdot) . Рассмотрим квадратичную экстремальную задачу:

$$\text{найти } Q^* = \max_{x \in R^n} Q_0(x) \quad (1.1)$$

при ограничениях $Q_i(x) = 0, i = 1, \dots, m$.

Здесь $Q_i(x)$ – квадратичные функции вида $Q_i(x) = (K_i x, x) + (b_i, x) + c_i$, где K_i – симметричные вещественные матрицы размера $n \times n$, b_i – n -мерные векторы из R^n , c_i – вещественные числа, $i = 0, \dots, m$. Некоторые из функций $Q_i(x), i = 0, 1, \dots, m$ могут быть и линейными. В общем случае задача (1.1) многоэкстремальна и относится к классу *NP*-трудных задач.

Оценки сверху для Q^* можно получить путем следующей лагранжевой релаксации. Пусть $u \in R^m$ – вектор множителей Лагранжа, соответствующий ограничениям задачи (1.1). Функция Лагранжа для задачи (1.1) имеет следующий вид

$$L(x, u) = Q_0(x) + \sum_{i=1}^m u_i Q_i(x) \equiv (K(u)x, x) + (b(u), x) + c(u),$$

где

$$K(u) = K_0 + \sum_{i=1}^m u_i K_i, \quad b(u) = b_0 + \sum_{i=1}^m u_i b_i, \quad c(u) = c_0 + \sum_{i=1}^m u_i c_i.$$

Рассмотрим функцию

$$\Psi(u) = \max_{x \in R^n} L(x, u) \equiv \max_{x \in R^n} [(K(u)x, x) + (b(u), x) + c(u)].$$

Пусть $\lambda_1(K)$ – максимальное собственное число симметричной $n \times n$ -матрицы K .

Функция $\Psi(u)$ есть выпуклой функцией от переменных u (как результат взятия операции максимума по переменным x для семейства линейных по переменным u функций). Эффективное множество функции $\Psi(u)$ обозначим $\text{dom } \Psi$. Оно состоит из $\Omega^- = \{u \in R^m : \lambda_1(K(u)) < 0\}$ (подмножество $u \in R^m$, для которых матрица $K(u)$ отрицательно определена) и подмножества тех точек $u \in \Omega^0 = \{u \in R^m : \lambda_1(K(u)) = 0\}$, для которых имеет решение система линейных уравнений $2K(u)x + b(u) = 0$. Для всех других точек $\Psi(u) = +\infty$.

Если $\text{dom } \Psi \neq \emptyset$, то для любого $u \in \text{dom } \Psi$ значение функции $\Psi(u)$ служит нетривиальной (т.е. неравной $+\infty$) оценкой сверху для Q^* – оптимального значения целевой функции в задаче (1.1). Оценка Шора Ψ^* есть наилучшей оценкой сверху для Q^* в классе лагранжевых оценок вида $\Psi(u)$ и связана с решением задачи негладкой оптимизации $\Psi^* = \min_{u \in \text{dom } \Psi} \Psi(u)$ (точки негладкости функции $\Psi(u)$ есть точками границы множества Ω^- , где система линейных уравнений $2K(u)x + b(u) = 0$ имеет неединственное решение). Оценку Ψ^* с любой заданной точностью можно найти за полиномиальное время методом эллипсоидов, итерация которого требует $O(m^2) + O(n^3)$ арифметических операций. Из них $O(m^2)$ операций связано с растяжением пространства двойственных переменных (множителей Лагранжа), а $O(n^3)$ операций требуются для вычисления вектора субградиентного поля при фиксированных значениях множителей Лагранжа (решение системы линейных уравнений с симметричной $n \times n$ -матрицей, нахождение максимального собственного числа симметричной $n \times n$ -матрицы и соответствующего этому числу собственного вектора).

Оценка Ψ^* обладает следующими свойствами. Если Ψ^* достигается на $u^* \in \Omega^-$, то $\Psi^* = Q^*$ (т.е. оценка точная). При этом находится и точка глобального минимума. Иначе Ψ^* достигается на границе области Ω^- , при этом может существовать так называемый "разрыв двойственности" $\Delta^* = \Psi^* - Q^* > 0$. Один из способов уменьшения Δ^* связан с введением функционально избыточных ограничений (при этом может увеличиться и количество переменных в задаче).

2. Улучшенная оценка Ψ_1^* . Функционально избыточными ограничениями есть ограничения, добавления которых оставляет множество допустимых решений начальной задачи неизменным. Однако при этом изменяется функция Лагранжа, что в некоторых случаях разрешает уменьшить Δ^* . Если к исходной задаче (1.1) прибавить r функционально избыточных квадратичных ограничений $Q_{m+1}(x) = 0, \dots, Q_{m+r}(x) = 0, r \geq 1$, то новая квадратичная задача примет вид:

$$\text{найти } Q^* = \max_{x \in R^n} Q_0(x) \quad (2.1)$$

при ограничениях $Q_i(x) = 0, i = 1, \dots, m + r$.

Лемма 1. Пусть Ψ_1^* – оценка вида Ψ^* для задачи (2.1). Тогда $\Psi_1^* \leq \Psi^*$.

Доказательство. Задаче (2.1) соответствует вектор множителей Лагранжа $U \in R^{m+r}$ и функция Лагранжа для нее имеет следующий

вид:

$$L_1(x, U) = Q_0(x) + \sum_{i=1}^{m+r} u_i Q_i(x) = L(x, u) + \sum_{i=m+1}^{m+r} u_i Q_i(x).$$

Поскольку $L_1\left(x, (\{u\}, 0, \dots, 0)\right) = L(x, u)$ и $\Psi_1(\{u\}, 0, \dots, 0) = \Psi(u)$, то

$$\Psi_1^* = \min_{U \in \text{dom } \Psi_1} \Psi_1(U) \leq \min_{u \in \text{dom } \Psi} \Psi(u) = \Psi^*,$$

откуда следует доказательство леммы 1.

Добавление функционально избыточных ограничений, которые являются нетривиальными следствиями из условий задачи, в ряде случаев приводит к тому, что верхняя оценка Ψ_1^* даже может стать точной для Q^* [2,3]. Ограничения, которые являются линейными комбинациями уже существующих ограничений, не отражаются на точности оценки Ψ_1^* . Вклад таких ограничений в функцию Лагранжа эквивалентен лишь определенному изменению множителей Лагранжа при существующих ограничениях. Справедлива следующая лемма.

Лемма 2. Пусть $Q_{m+k}(x) = \sum_{i=1}^m \lambda_{ki} Q_i(x)$, $k=1, \dots, r$. Тогда $\Psi_1^* = \Psi^*$.

Лемма 2 помогает в решении двух вопросов. Во-первых, если в квадратичной задаче имеется ограничение, которое есть линейной комбинацией других ограничений, то его можно убрать, уменьшив тем самым число множителей Лагранжа, а значит, и трудоемкость метода эллипсоидов для нахождения оценки вида Ψ^* . Во-вторых, с ее помощью для комбинаторных задач легко строится аппроксимация целочисленного многогранника допустимых решений с помощью неполиномиального количества линейных неравенств. Ниже рассмотрим такую аппроксимацию на примере обоснования одной из верхних оценок для числа независимости графа.

3. Число независимости графа. Пусть $G=(V, E)$ – неориентированный граф (не содержащий петель) с множеством вершин V и множеством ребер E . Подмножество вершин $S \subseteq V$ называется независимым множеством графа G , если для любых $i, j \in S$ ребро $e = (i, j)$ не принадлежит E . Число независимости графа принято обозначать $\alpha(G)$ и оно характеризует мощность максимального по числу входящих в него вершин независимого множества в графе G . В общем случае задача нахождения $\alpha(G)$ принадлежит к NP -трудным задачам [4].

Задача о нахождении $\alpha(G)$ может быть сформулирована в форме следующей квадратичной булевой задачи [1,2]:

$$\text{найти } \alpha(G) = \max \sum_{i \in V} x_i \quad (3.1)$$

при ограничениях $x_i x_j = 0 \forall (i, j) \in E$, $x_i^2 - x_i = 0 \forall i \in V$.

Здесь булева переменная $x_i \in \{0, 1\}$ равна единице, если вершина $i \in V$ включается в независимое множество, и нулю – в противоположном случае. Квадратичные ограничения для ребер графа G означают, что две вершины, связанные ребром в графе G , не могут одновременно принадлежать независимому множеству.

Задаче (3.1) соответствует оценка $\Psi^*(G)$ и ее нахождение связано с решением задачи негладкой оптимизации с $(|E| + |V|)$ двойственными переменными. Оценка $\Psi^*(G)$ аппроксимирует сверху число независимости графа $\alpha(G)$, т.е. $\Psi^*(G) \geq \alpha(G)$. Н.З. Шор доказал [2], что $\Psi^*(G) = \vartheta(G)$, где $\vartheta(G)$ – известное число Ловаса [4]. Другими словами, оценка Шора $\Psi^*(G)$ имеет такие же свойства, как и число Ловаса $\vartheta(G)$. Так, например, если граф G принадлежит семейству совершенных графов, то $\Psi^*(G) = \alpha(G)$, т.е. оценка является точной.

Покажем, как обосновать этот же факт с помощью леммы 2. Пусть Q – произвольная клика (полный подграф) в графе G и $V(Q)$ – соответствующее ей подмножество вершин из V . Клике Q соответствует линейная комбинация ограничений задачи (3.1)

$$\sum_{i \in V(Q)} (x_i^2 - x_i) + 2 \sum_{i, j \in V(Q), i < j} x_i x_j = 0,$$

которую можно записать в форме квадратичного равенства

$$\left(\sum_{i \in V(Q)} x_i \right)^2 - \sum_{i \in V(Q)} x_i = 0. \quad (3.2)$$

Ограничение вида (3.2) справедливо и для ребра графа G , которое можно рассматривать как клику из двух вершин.

Если ограничения в задаче (3.1) заменить на ограничения

$$\left(\sum_{i \in V(Q)} x_i \right)^2 - \sum_{i \in V(Q)} x_i = 0 \forall Q \in G, \quad x_i^2 - x_i = 0 \forall i \in V, \quad (3.3)$$

то согласно лемме 2 новой квадратичной задаче будет соответствовать оценка $\Psi^*(G)$. В общем случае, новая задача может содержать неполиномиальное количество квадратичных равенств, учитывая, что с их помощью описываются все возможные клики в графе, в том числе и ребра графа. Оценка $\Psi^*(G)$ для новой задачи аппроксимируется сверху величиной $\alpha^*(G)$, которая равна оптимальному значению целевой функции в задаче линейного программирования:

$$\text{найти } \alpha^*(G) = \max \sum_{i \in V} x_i \quad (3.4)$$

$$\text{при ограничениях } \sum_{i \in V(Q)} x_i \leq 1 \quad \forall Q \in G, \quad 0 \leq x_i \leq 1 \quad \forall i \in V,$$

полученной релаксацией квадратичных ограничений-равенств (3.3). В общем случае задача (3.4) может содержать неполиномиальное количество линейных неравенств.

Теорема 1. Для произвольного графа G имеем $\alpha^*(G) \geq \Psi^*(G) \geq \alpha(G)$.

Система ограничений в задаче (3.4) задает многогранник, который называют кликовым многогранником $QSTAB(G)$. Он аппроксимирует сверху многогранник независимых множеств $STAB(G)$ (stable set polytope), определяемый как выпуклая оболочка булевых индикаторных векторов независимых множеств S в графе G :

$$STAB(G) = \text{conv}\{x^S, S \text{ — независимое множество в графе } G.\}$$

Индикаторный вектор множества S определяется как $x^S = (x_i^S, i \in V) \in \{0, 1\}^{|V|}$, где $x_i^S = 1$, если $i \in S$ и $x_i^S = 0$ в противном случае.

Граф G называют совершенным, если кликовый многогранник $QSTAB(G)$ и многогранник независимых множеств $STAB(G)$ совпадают. Если граф G — совершенный, то задача (3.4) связана с максимизацией линейной функции на многограннике $QSTAB(G) = STAB(G)$ и, следовательно, величина $\alpha^*(G)$ совпадает с величиной $\alpha(G)$. Из теоремы 1 для совершенного графа G имеем $\alpha^*(G) = \Psi^*(G) = \alpha(G)$, откуда следует, что оценка $\Psi^*(G)$ есть точной оценкой для числа независимости совершенных графов.

Более сильные свойства имеет оценка Н.З. Шора (условимся ее обозначать $\Psi_1^*(G)$), которая соответствует квадратичной задаче, полученной из (3.1) прибавлением семейства функционально избыточных ограничений

$$x_i x_k + x_j x_k \leq x_k, \quad \forall i, j, k : (i, j) \in E, k \neq i, j. \quad (3.5)$$

Они получаются умножением линейных неравенств $x_i + x_j \leq 1$ для ребер $e = (i, j) \in E$ на неотрицательную переменную $x_k = x_k^2 \geq 0, k \neq i, j$.

Наличие ограничений (3.5) придает оценке $\Psi_1^*(G)$ ряд замечательных свойств, которые связаны со специальными семействами графов. Так, например, в [2] показано, что оценка $\Psi_1^*(G)$ является точной для $\alpha(G)$, когда граф G есть t -совершенным. В [5] найдены новые семейства линейных неравенств для аппроксимации многогранника $STAB(G)$, которые связаны с подструктурами в графе, полученными определенной комбинацией клик, нечетных циклов и дополнений к нечетным циклам (odd-antihole). В частности, эти семейства включают известные колесные и p -колесные неравенства (wheel и p -wheel constraints). Это позволяет дополнить семейства графов из [4], для которых задача нахождения числа независимости $\alpha(G)$ может быть решена за полиномиальное время. Как частный случай сюда попадает и известное семейство W -совершенных графов, для которого полиномиальный алгоритм обоснован с помощью метода эллипсоидов.

Линейные неравенства для p -колес использованы в итерационном алгоритме для нахождения эффективных (по типу Ловаса-Шора) ЛП-ориентированных верхних оценок для числа независимости неориентированных графов с тысячами вершин [6]. Алгоритм реализован программой LPWSTAB, где использована известная ЛП-программа SoPlex-1.4.1. С ее помощью найдены верхние оценки для размера максимальной клики в графах keller6 (3361 вершин и 4,619,898 ребер) и keller7 (14190 вершин и 87,091,347 ребер), которые связаны с доказательством гипотезы Келлера (Keller's conjecture). Численные эксперименты проводились на 64-разрядных компьютерах Института автоматки и процессов управления ДВО РАН. Для графа keller7 на компьютере с 36Гб оперативной памяти программа LPWSTAB работала 13 часов (сделала 8 итераций, каждая из которых была связана с решением ЛП-задач с 14190 переменных и максимально 13,593,798 ограничений).

Работа поддержана совместным украинско-российским проектом ДФФД Украины – Ф28.1/005 и РФФИ – 09-01-90413-Укр.

Литература

1. Н.З. ШОР, С.И. СТЕЦЕНКО. Квадратичные экстремальные задачи и недифференцируемая оптимизация. Киев: Наукова думка, 1989.
2. N.Z. SHOR. Nondifferentiable Optimization and Polynomial Problems. Boston; Dordrecht; London: Kluwer Academic Publishers, 1998.

3. N.Z. SHOR, P.I. STETSYUK. Lagrangian bounds in multiextremal polynomial and discrete optimization problems. // J. of Global Optimization. 2002. V. 23. P. 1–41.
4. M. GROTSCHEL, L. LOVASZ, A. SCHRIJVER. Geometric Algorithms and Combinatorial Optimization. – Berlin: Springer-Verlag, 1988.
5. П.И.СТЕЦЮК. О новых свойствах оценок Шора для взвешенного числа устойчивости графа // Праці міжнародної конференції “50 років Інституту кібернетики ім.В.М.Глушкова НАН України”, К: Інститут кібернетики ім. В.М.Глушкова НАН України, 2008. С. 164–173
6. П.И. СТЕЦЮК, А.П. ЛИХОВИД. Об ЛП-ориентированных верхних оценках для взвешенного числа устойчивости графа // Кибернетика и системный анализ. 2009. № 1. С. 157–170.

Р А З Д Е Л V

Специальные квадратичные задачи (часть I)

В раздел включены статьи:

- 5.1. Двойственные оценки для специальной оптимизационной задачи квадратичного типа на многообразии Штифеля (*Шор, Стецюк, Березовский, 2004*) 342
- 5.2. Лагранжевая оценка для максимального разреза графа с заданными количествами вершин в обоих подмножествах разбиения (*Стецюк, Березовский, 2008*) 351
- 5.3. Квадратичная оптимизационная задача для максимального k -плекса в неориентированном графе (*Стецюк, 2006*) 360
- 5.4. Квадратична задача для максимального k -плекса в неорієнтованому графі (*Стецюк, Бардадим, Ляшко, 2017*) 378
- 5.5. Властивості квадратичної задачі про максимальний k -плекс у неорієнтованому графі (*Стецюк, Ляшко, Бардадим, 2017*) 388

Двойственные оценки для специальной оптимизационной задачи квадратичного типа на многообразии Штифеля

Шор Н.З., Стецюк П.И., Березовский О.А.

*Теория оптимальных решений. – Киев: Ин-т кибернетики
им. В.М. Глушкова НАН Украины, 2004. – С. 3–10.*

Рассматривается задача глобальной оптимизации, связанная с нахождением минимума квадратичной однородной функции специального вида на многообразии Штифеля. Рассмотрены лагранжевые двойственные оценки для трех вариантов этой задачи, которые различаются ограничениями на компоненты ортонормированных векторов. На тестовых примерах показано, что их можно эффективно использовать при нахождении глобального минимума.

Многообразие Штифеля (Stiefel manifold) состоит из всех ортонормированных систем векторов $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_k \in R^n$, где R^n – n -мерное евклидово пространство и $k \leq n$. В работе [1] рассмотрена следующая оптимизационная задача на многообразии Штифеля:

$$\min \left\{ f(\vec{x}_1, \dots, \vec{x}_k) = \sum_{i=1}^k \vec{x}_i^T A_i \vec{x}_i \right\} \quad (1)$$

при ограничениях

$$\vec{x}_i^T \vec{x}_j = \delta_{ij}, \quad 1 \leq i, j \leq k, \quad (2)$$

где A_i , $i=1, \dots, k$ – заданные $n \times n$ вещественные симметричные матрицы, δ_{ij} – символ Кронекера. Ограничения (2) задают многообразие Штифеля, которое принято обозначать $M_{n,k}$.

В зависимости от условий, наложенных на компоненты векторов из $M_{n,k}$, выделим следующие варианты задачи (1)–(2):

- (R) компоненты векторов \vec{x}_i , $i=1, \dots, k$ принадлежат множеству действительных чисел. Этот вариант совпадает с постановкой задачи (1)–(2);
- (Z) компоненты векторов \vec{x}_i , $i=1, \dots, k$ принадлежат множеству целых чисел. Для этого варианта векторы \vec{x}_i , $i=1, \dots, k$ могут содержать только компоненты либо 0, либо +1, либо –1;

(N) компоненты векторов $\vec{x}_i, i=1, \dots, k$ принадлежат множеству натуральных чисел, т.е. векторы $\vec{x}_i, i=1, \dots, k$ содержат только компоненты либо 0, или 1.

Варианты (Z) и (N) определяются только диагональными коэффициентами матриц A_i и не зависят от коэффициентов вне диагонали, что обусловлено спецификой целевой функции вида (1). Кроме того варианты (Z) и (N) близки в том смысле, что им соответствует одно и то же значение целевой функции в точках глобальных экстремумов. Отличие состоит лишь в числе глобальных экстремумов, т. е. если у варианта (N) их количество равно L , то у варианта (Z) их будет $2^k L$.

В общем случае задача (1)-(2) для каждого из трех вариантов является многоэкстремальной задачей нелинейного программирования (пусть даже и квадратичного типа), и ее решение представляет проблемы даже при небольших размерностях $n = 2$ и $n = 3$. Так, например, в [2] сообщается, что простой пример на $M_{2,2}$ с диагональными коэффициентами матриц $A_i, i = 1, 2$ требует три миллиона вычислений значений функции при использовании программы GlobSol [3 – 4]. Однако, используя технику двойственных лагранжевых квадратичных оценок [5], нахождение глобального минимума задачи (1)-(2) можно значительно ускорить в ряде случаев.

Рассмотрим оптимизационные задачи квадратичного типа, соответствующие каждому из указанных трех вариантов этой задачи, и лагранжевые двойственные оценки для каждой из задач.

Для каждого $i = \overline{1, k}$ обозначим коэффициенты симметричной $n \times n$ -матрицы A_i через $a_{ijl}, j = \overline{1, n}, l = \overline{1, n}$, а компоненты n -мерного вектора \vec{x}_i через $x_{ij}, j = \overline{1, n}$.

Варианту (R) соответствует оптимизационная задача квадратичного типа:

$$Q_1^* = \min \sum_{i=1}^k \sum_{j=1}^n \sum_{l=1}^n a_{ijl} x_{ij} x_{il} \quad (3)$$

при ограничениях

$$\sum_{j=1}^n x_{ij}^2 = 1, \quad i = \overline{1, k}, \quad (4)$$

$$\sum_{l=1}^n x_{il} x_{jl} = 0, \quad i = \overline{1, k-1}, \quad j = \overline{i+1, k}. \quad (5)$$

Задача (3)–(5) – переписанный с учетом расшифровки символа Кронекера δ_{ij} аналог задачи (1)–(2). Здесь ограничения (4) задают условие

нормированности векторов \vec{x}_i , $i = \overline{1, k}$, а ограничения (5) – условие их взаимной ортогональности. Указанных ограничений достаточно для задания многообразия Штифеля $M_{n, k}$, состоящего из всех ортонормированных систем векторов $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_k \in R^n$ ($k \leq n$).

Пусть $u \in R^m$, где $m = k(k+1)/2$ – вектор множителей Лагранжа, состоящий из двух частей: $\{u_{ii}, i = \overline{1, k}\}$ – множители Лагранжа, соответствующие ограничениям (4) и $\{u_{ij}, i = \overline{1, k-1}, j = \overline{i+1, k}\}$ – множители Лагранжа, соответствующие ограничениям (5). Согласно [5], задаче (3)–(5) соответствует следующая двойственная лагранжева оценка:

$$\psi_1^* = \psi_1(u^*) = \sup_{u: K(u) \geq 0} \psi_1(u), \quad (6)$$

где функция $\psi_1(u) = - \sum_{i=1}^k u_{ii}$, а $K(u) \geq 0$ обозначает неотрицательную определенность семейства симметричных $kn \times kn$ -матриц вида:

$$K(u) = K_0 + \sum_{i=1}^k u_{ii} K_{ii} + \frac{1}{2} \sum_{i=1}^{k-1} \sum_{j=i+1}^k u_{ij} (K_{ij} + K_{ji}).$$

Здесь матрицы K_0 и K_{ij} состоят из k^2 блоков размером $n \times n$ и задаются следующим образом. По диагонали матрицы K_0 размещены блоки, заданные симметричными $(n \times n)$ -матрицами A_i , $i = \overline{1, k}$, а все внедиагональные блоки матрицы K_0 равны нулевым $n \times n$ -матрицам. K_{ij} – матрица, (i, j) -й блок которой равен единичной $n \times n$ -матрице, а все остальные блоки равны нулевым $n \times n$ -матрицам.

Варианту (Z) поставим в соответствие следующую оптимизационную задачу квадратичного типа:

$$Q_2^* = \min \sum_{i=1}^k \sum_{j=1}^n \sum_{l=1}^n a_{ijl} x_{ij} x_{il} \quad (7)$$

при ограничениях

$$\sum_{j=1}^n x_{ij}^2 = 1, \quad i = \overline{1, k} \quad (8)$$

$$\sum_{l=1}^n x_{il} x_{jl} = 0, \quad i = \overline{1, k-1}, \quad j = \overline{i+1, k}, \quad (9)$$

$$x_{ij} x_{il} = 0, \quad i = \overline{1, k}, \quad j = \overline{1, n-1}, \quad l = \overline{j+1, n}, \quad (10)$$

$$x_{ij} x_{lj} = 0, \quad j = \overline{1, k}, \quad i = \overline{1, k-1}, \quad l = \overline{i+1, k}. \quad (11)$$

Здесь ограничения (8) и (9) такие же, как и в предыдущей задаче, т. е. задают многообразие Штифеля $M_{n,k}$ на множестве вещественных чисел. Тот факт, что компоненты векторов должны принадлежать множеству целых чисел, уточняется посредством двух групп дополнительных ограничений. Так группа ограничений (10) обозначает, что для любой пары компонент вектора их произведение должно равняться нулю. Это связано с тем фактом, что для варианта (Z) каждый из векторов \vec{x}_i содержит только одну ненулевую компоненту, равную $+1$, либо -1 . Последнее следует из условий нормировки векторов \vec{x}_i , $i = \overline{1, k}$, заданных ограничениями (8), и того, что компоненты векторов должны принадлежать множеству целых чисел. Ограничения (11) означают, что для двух различных векторов компоненты с одним и тем же номером не могут одновременно быть ненулевыми, т. е. ± 1 для разных векторов должны находиться на разных местах. Последнее обусловлено уже не только нормированностью, но и взаимной ортогональностью векторов.

Указанные группы ограничений идентифицируют вариант (Z) в том смысле, что множество допустимых решений системы (8)-(11) будет состоять только из целых чисел: 0 , $+1$ и -1 . Отметим, что последнее было бы справедливо и тогда, если рассмотреть для варианта (Z) задачу только с одной группой ограничений, т. е. либо (10), либо (11). Однако, согласно [5], включение в систему ограничений функционально избыточных ограничений способствует улучшению нижней двойственной лагранжевой оценки.

Задаче (7)–(11) соответствует двойственная лагранжева оценка:

$$\psi_2^* = \psi_2(u^*, v^*, V^*) = \sup_{u, v, V : K(u, v, V) \geq 0} \psi_2(u, v, V), \quad (12)$$

где $\psi_2(u, v, V) = \psi_1(u) = - \sum_{i=1}^k u_{ii}$.

Здесь $u \in R^m$ – вектор множителей Лагранжа, соответствующих ограничениям (8) и (9); $v \in R^{m_1}$, $m_1 = k \frac{n(n-1)}{2}$ – вектор множителей Лагранжа, соответствующих ограничениям (10); $V \in R^{m_2}$, $m_2 = n \frac{k(k-1)}{2}$ – вектор множителей Лагранжа, соответствующих ограничениям (11); $K(u, v, V) \geq 0$ обозначает неотрицательную определенность семейства симметричных матриц

$$K(u, v, V) = K(u) + \frac{1}{2} \sum_{i=1}^k \sum_{j=1}^{n-1} \sum_{l=j+1}^n \nu_{ijl} (K_{ijl} + K_{ilj}) +$$

$$+ \frac{1}{2} \sum_{i=1}^{k-1} \sum_{l=i+1}^k \sum_{j=1}^n V_{ijl} \left(\tilde{K}_{ilj} + \tilde{K}_{lij} \right),$$

где K_{ijl} – матрица размера $nk \times nk$, в i -м диагональном блоке которой содержится $(n \times n)$ -матрица с (j, l) -м коэффициентом, равным единице, и всеми остальными коэффициентами, равными нулю; \tilde{K}_{ilj} – матрица размера $nk \times nk$, для которой блок (i, l) содержит $(n \times n)$ -матрицу с j -м диагональным коэффициентом, равным единице, и всеми остальными коэффициентами, равными нулю. Все остальные блоки матрицы \tilde{K}_{ilj} равны нулевым $(n \times n)$ -матрицам.

Как видим, функция $\psi_2(u, v, V)$ имеет такой же вид, как и функция ψ_1 для первой задачи. Это обусловлено однородностью добавленных ограничений (10)–(11) и тем, что правая часть в этих ограничениях тождественно равна нулю.

Далее рассмотрим вариант (N). Ему соответствует оптимизационная задача квадратичного типа:

$$Q_3^* = \min \sum_{i=1}^k \sum_{j=1}^n \sum_{l=1}^n a_{ijl} x_{ij} x_{il} \quad (13)$$

при ограничениях

$$\sum_{j=1}^n x_{ij}^2 = 1, \quad i = \overline{1, k}, \quad (14)$$

$$\sum_{l=1}^n x_{il} x_{jl} = 0, \quad i = \overline{1, k-1}, \quad j = \overline{i+1, k}, \quad (15)$$

$$x_{ij} x_{il} = 0, \quad i = \overline{1, k}, \quad j = \overline{1, n-1}, \quad l = \overline{j+1, n}, \quad (16)$$

$$x_{ij} x_{lj} = 0, \quad j = \overline{1, k}, \quad i = \overline{1, k-1}, \quad l = \overline{i+1, k}. \quad (17)$$

$$x_{ij}^2 - x_{ij} = 0, \quad i = \overline{1, k}, \quad j = \overline{1, n}. \quad (18)$$

Задача (13)–(18) отличается от задачи (7)–(11) тем, что к последней добавлены булевы ограничения (18) на неизвестные компоненты векторов \vec{x}_i , $i = \overline{1, k}$. Эти ограничения исключают из множества допустимых точек $M_{n,k}$ векторы с компонентами, равными -1, и, соответственно, компоненты векторов могут быть равны либо нулю, либо единице. Следовательно, среди допустимых решений системы ограничений (14)–(18) будут векторы, содержащие только одну компоненту равную единице,

а уже какая конкретно компонента будет определяться глобальным экстремумом задачи (13)-(18).

Задача (13)-(18) существенно отличается от двух предыдущих. В обеих предыдущих задачах квадратичные функции, задающие как функцию цели, так и ограничения, были однородными и не содержали линейных членов. В последней же задаче булевы ограничения приводят к линейным членам. В силу этого нахождение двойственной лагранжевой оценки для задачи (13)-(18) будет сложнее и потребует решения системы линейных уравнений.

Пусть $w \in R^m$ ($m_3 = nk$) – вектор множителей Лагранжа, соответствующий ограничениям (18), а $u \in R^m$, $v \in R^{m_1}$, $V \in R^{m_2}$ – векторы множителей Лагранжа для ограничений (14)-(15), (16) и (17), соответственно.

Задаче (13)-(18) соответствует такая двойственная лагранжевая оценка:

$$\psi_3^* = \sup_{u,v,V,w:K(u,v,V,w) \succeq 0} \psi_3(u, v, V, w),$$

где

$$\psi_3(u, v, V, w) = \inf_{x \in R^{kn}} \left\{ \left(K(u, v, V, w)x, x \right) - \sum_{i=1}^k \sum_{j=1}^n w_{ij} x_{ij} - \sum_{i=1}^k u_{ii} \right\}.$$

Здесь $K(u, v, V, w) \succeq 0$ обозначает неотрицательную определенность семейства симметричных матриц

$$K(u, v, V, w) = K(u, v, V) + \text{diag} \{ w_{11}, \dots, w_{1n}, \dots, w_{k1}, \dots, w_{kn} \}.$$

Когда ψ_3^* достигается внутри области положительной определенности семейства матриц $K(u, v, V, w)$ (т. е. $K(u^*, v^*, V^*, w^*) \succeq 0$, точку глобального экстремума x^* для задачи (13)-(18) можно найти по следующей формуле

$$x^* = (x_{11}^*, \dots, x_{1n}^*, \dots, x_{k1}^*, \dots, x_{kn}^*)^T = \frac{1}{2} K^{-1}(u^*, v^*, V^*, w^*) w^*.$$

В данном случае оценка ψ_3^* будет точной нижней оценкой для Q_3^* (т. е. $\psi_3^* = Q_3^*$).

Эффективность двойственных оценок ψ_1^* , ψ_2^* , ψ_3^* исследовалась на ряде тестовых примеров (для их нахождения использовалась программа DSQTPR [6] с использованием модификации г-алгоритма [7]). Ниже приведем два из них.

Пример 1. Рассмотрим задачу (1)–(2), где

$$A_1 = \text{diag}(-1, 2, 3), \quad A_2 = \text{diag}(4, -5, 6), \quad A_3 = \text{diag}(7, 8, -9).$$

Целевая функция имеет вид

$$f(x) = -x_{11}^2 + 2x_{12}^2 + 3x_{13}^2 + 4x_{21}^2 - 5x_{22}^2 + 6x_{23}^2 + 7x_{31}^2 + 8x_{32}^2 - 9x_{33}^2,$$

множество допустимых решений состоит из $8 = 2^3$ точек следующего вида:

$$x^* = ((\pm 1, 0, 0)^T, (0, \pm 1, 0)^T, (0, 0, \pm 1)^T).$$

Оптимальное значение целевой функции равно сумме отрицательных компонент матриц A_i , т. е. $f^* = f(x^*) = -15$. Следовательно, для вариантов (R) и (Z) имеем $Q_1^* = Q_2^* = -15$. Для варианта (N) $Q_3^* = -15$, но решением будет единственная точка $x^* = ((1, 0, 0)^T, (0, 1, 0)^T, (0, 0, 1)^T)$.

Для первого примера все лагранжевые двойственные оценки являются точными, т.е.

$$\psi_1^* = -15 = Q_1^*, \quad \psi_2^* = -15 = Q_2^*, \quad \psi_3^* = -15 = Q_3^*.$$

Кроме того, для варианта (N) программа DSQTPR обеспечивает нахождение точного оптимального решения задачи (13)–(18) по переменным x^* . Для этого ей потребовалось 208 итераций г-алгоритма, и, следовательно, затраты на нахождение и строгое обоснование точки глобального минимума для варианта (N) небольшие (время счета 1 сек на компьютере IBM PC/AT/486DX/66MHz).

Недостаток первого примера состоит в том, что глобальный минимум для варианта (N) однозначен, и значение целевой функции в нем сильно отличается от значений целевой функции в других допустимых точках. Поэтому ниже опишем второй, более сложный пример, когда глобальный минимум для варианта (N) неоднозначен.

Пример 2. Заданы матрицы

$$A_1 = \text{diag}(1, 2, 3), \quad A_2 = \text{diag}(4, 5, 6), \quad A_3 = \text{diag}(7, 8, 9),$$

т.е. целевая функция имеет вид

$$f(x) = x_{11}^2 + 2x_{12}^2 + 3x_{13}^2 + 4x_{21}^2 + 5x_{22}^2 + 6x_{23}^2 + 7x_{31}^2 + 8x_{32}^2 + 9x_{33}^2.$$

Для данного примера $Q_1^* = Q_2^* = Q_3^* = 15$, а множество допустимых решений будет иметь следующий вид. Для варианта (N) имеем шесть допустимых решений:

$$x_1^* = ((1, 0, 0)^T, (0, 1, 0)^T, (0, 0, 1)^T); x_2^* = ((1, 0, 0)^T, (0, 0, 1)^T, (0, 1, 0)^T);$$

$$x_3^* = ((0, 1, 0)^T, (1, 0, 0)^T, (0, 0, 1)^T); x_4^* = ((0, 1, 0)^T, (0, 0, 1)^T, (1, 0, 0)^T);$$

$$x_5^* = ((0, 0, 1)^T, (1, 0, 0)^T, (0, 1, 0)^T); x_6^* = ((0, 0, 1)^T, (0, 1, 0)^T, (1, 0, 0)^T).$$

Для варианта (Z) количество допустимых решений равно $6 \cdot 2^3 = 48$, где решения будут отличаться тем, что на соответствующих единице местах могут находиться как единица, так и минус единица. Множество допустимых решений для варианта (R) содержит бесконечное число точек.

Для второго примера лагранжевые двойственные оценки $\psi_1^* = 12$ и $\psi_2^* = 12$, и ни одна из них не является точной нижней оценкой для Q_1^* и Q_2^* . Заметим, что разрыв двойственности $\Delta_i = Q_i^* - \psi_i^*$, $i = 1, 2$, равен трем и есть достаточно большим ($\approx 20\%$). Оценка $\psi_3^* = 15$ является точной. Из-за неоднозначности решений программа DSQTPR не гарантирует нахождения оптимального решения задачи (13)–(18) по переменным x^* . Для того, чтобы получить одно из решений требуется соответственно "возмутить" коэффициенты матриц A_i , $i = 1, 2, 3$. Так, например, для "возмущенных" матриц $\bar{A}_1 = \text{diag}(0.9999, 2, 3)$, $\bar{A}_2 = \text{diag}(4, 4.9999, 6)$, $\bar{A}_3 = \text{diag}(7, 8, 9)$ получим решение x_1^* , а для "возмущенных" матриц $\bar{A}_1 = \text{diag}(1, 2, 3)$, $\bar{A}_2 = \text{diag}(4, 4.9999, 6)$, $\bar{A}_3 = \text{diag}(6.9999, 8, 9)$, получим решение x_6^* . Для этого программе DSQTPR потребовалось 499 итераций г-алгоритма в первом случае и 487 итераций – во втором.

В заключение отметим, что описанный в работе аппарат нахождения двойственных оценок с использованием функционально избыточных ограничений можно перенести на случай, когда квадратичная функция цели имеет общий вид (при этом ограничения остаются неизменными). То есть предлагаемый подход для получения нижних оценок функционала (а в некоторых случаях и оптимальных значений компонент вектора переменных) возможно использовать для оценок NP-трудных задач, например, в квадратичной задаче о назначениях.

Литература

1. RAPCSAK T. On minimization on Stiefel manifold // European J. of Operation Research, 2002. – 143. – P. 365–376.
2. BALOGH J., CSENDES T., RAPCSAK T. Global optimization problems on Stiefel manifold // NMCM-2002 Book of Abstracts, Miskolc, Hungary, 2002. – P. 19–21.
3. CORLISS G.F., KEARFOTT R.B. Rigorous global search: Industrial applications. – In: T. Csendes (ed.): Development in Reliable Computing, Kluwer, Dordrecht, 1999. – P. 1–16.
4. KEARFOTT R.B. Rigorous Global Search: Continuous Problems, Kluwer, Dordrecht, 1996. – 98 p.
5. SHOR N.Z. Nondifferentiable optimization and polynomial problems. – Dordrecht: Kluwer, 1998. – 394 p.
6. SHOR N. Z., STETSYUK P.I. Dual Solution of Quadratic-Type Problems by r -algorithm (subroutine DSQTPr) // Abstracts of Second International Workshop "Recent Advances in Nondifferentiable Optimization" (October, 1–4, 2001, Kyiv, Ukraine). – P. 36.
7. 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.

Получено 06.08.2004

Лагранжевая оценка для максимального разреза графа с заданными количествами вершин в обоих подмножествах разбиения¹

Стецюк П.И., Березовский О.А.

*Теория оптимальных решений. – Киев: Ин-т кибернетики
им. В.М. Глушкова НАН Украины, 2006. – С. 31–38.*

Рассматривается задача нахождения верхней Лагранжевой оценки Ψ^* для квадратичной формулировки задачи о максимальном взвешенном разрезе графа с заданными количествами вершин в обоих подмножествах разбиения. Показано, что нахождение Ψ^* сводится к решению безусловной задачи минимизации выпуклой функции.

Имеется взвешенный неориентированный граф $G(V, E)$ с множеством вершин $V = \{1, \dots, n\}$ и множеством ребер E . Вес ребра $(i, j) \in E$ задан произвольным вещественным числом w_{ij} . Пусть V_1 и V_2 – непустые и непересекающиеся подмножества вершин графа G , такие что $V_1 \cup V_2 = V$. Разрезом графа G есть совокупность тех ребер, концевые вершины которых лежат в разных подмножествах. Вес разреза определяется как сумма весов его ребер.

Рассмотрим задачу нахождения максимального взвешенного разреза графа с заданными количествами вершин в обоих подмножествах разбиения. Не ограничивая общности, будем считать, что количество вершин в подмножестве V_1 равно k , а в подмножестве V_2 – $(n - k)$, где $k < n$. Задача нахождения максимального взвешенного разреза графа с заданными количествами вершин в обоих подмножествах разбиения состоит в нахождении разбиения множества вершин V на два таких подмножества V_1^* и V_2^* ($|V_1^*| = k$), чтобы соответствующий этому разбиению вес разреза был максимальным. Максимальный вес разреза обозначим $mc(G, w, k)$.

В случае, когда $k = 1, 2, 3$, задача нахождения максимального разреза графа с заданными количествами вершин в обоих подмножествах

¹Работа выполнена при частичной финансовой поддержке гранта UKM2-2812-KV-06 (CRDF Cooperative Grants Program)

разбиения является достаточно простой. Так, даже чисто переборные алгоритмы для нее будут иметь сложность $O(|E|)$, $O(|E|^2)$, $O(|E|^3)$ соответственно. Однако с увеличением k сложность задачи возрастает, и при больших $k \leq \frac{n}{2}$ задача становится NP -трудной. Так, например, когда $k = \frac{n}{2}$, а n – четное, то эта задача является NP -трудной задачей нахождения оптимальной бисекции графа, которая есть NP -трудной ([1]). Поэтому вопросы, связанные с уточнением верхних оценок для $mc(G, w, k)$, есть актуальными и представляют интерес.

Предметом обсуждения в работе будут верхние Лагранжевые оценки для $mc(G, w, k)$, которые получаются из формулировки задачи нахождения максимального взвешенного разреза графа с заданными количествами вершин в обоих подмножествах разбиения в форме оптимизационной задачи квадратичного типа.

1. Формулировка оптимизационной квадратичной задачи.

Пусть вершине $i \in V$ ($i \in \{1, 2, \dots, n\}$) соответствует бинарная переменная $x_i \in \{-1, 1\}$, такая что

$$x_i = \begin{cases} -1, & \text{если } i \in V_1, \\ +1, & \text{если } i \in V_2. \end{cases}$$

Задача нахождения максимального разреза графа с заданными количествами вершин в обоих подмножествах разбиения формулируется такой оптимизационной задачей квадратичного типа:

$$mc(G, w, k) = \max_{x \in E^n} \left\{ Q(x) = \frac{1}{4} \sum_{(i,j) \in E} w_{ij} (x_i - x_j)^2 \right\} \quad (1)$$

при ограничениях:

$$x_i^2 = 1, \quad \forall i = 1, \dots, n, \quad (2)$$

$$\left(\sum_{i=1}^n x_i \right)^2 = (n - 2k)^2. \quad (3)$$

Здесь квадратичная функция $Q(x)$ в целевой функции (1) задает вес разреза графа для произвольного разбиения вершин V на подмножества V_1 и V_2 . Функцию $Q(x)$ можно записать в такой форме

$$Q(x) = \frac{1}{4} \sum_{(i,j) \in E} w_{ij} (x_i - x_j)^2 = \sum_{(i,j) \in E} w_{ij} \left(\frac{x_i - x_j}{2} \right)^2 = \sum_{(i,j) \in E} w_{ij} y_{ij}, \quad (4)$$

где

$$y_{ij} = \left(\frac{x_i - x_j}{2} \right)^2.$$

Учитывая, что $x_i^2 = 1$ и $x_j^2 = 1$, имеем

$$y_{ij} = \frac{(x_i - x_j)^2}{4} = \frac{x_i^2 - 2x_i x_j + x_j^2}{2} = \frac{2 - 2x_i x_j}{4} = \frac{1 - x_i x_j}{2}.$$

Отсюда

$$y_{ij} = \frac{1 - x_i x_j}{2} = \begin{cases} 1, & \text{если } x_i = 1 \text{ и } x_j = -1, \\ 1, & \text{если } x_i = -1 \text{ и } x_j = 1, \\ 0, & \text{если } x_i = 1 \text{ и } x_j = 1, \\ 0, & \text{если } x_i = -1 \text{ и } x_j = -1, \end{cases}$$

и суммироваться в функции $Q(x)$ будут веса только тех ребер, вершины которых лежат в разных подмножествах V_1 и V_2 .

Итак, целевая функция (1) означает максимизацию веса разреза графа для произвольного разбиения вершин V на подмножества V_1 и V_2 . Ограничения (2)–(3) задают условие: разбиение должно быть таким, чтобы одно из подмножеств содержало k вершин из V , а второе – оставшиеся $n - k$ вершин. Ограничения (2) есть записью бинарности переменных $x_i, i = 1, \dots, n$, и означают, что x_i принадлежит либо подмножеству V_1 ($x_i = -1$), либо подмножеству V_2 ($x_i = 1$). Ограничение (3) задает условие: разность между количеством вершин в двух подмножествах V_1 и V_2 может быть равной либо $k - (n - k) = 2k - n$, либо $(n - k) - k = n - 2k$, т.е.

$$\sum_{i=1}^n x_i = \pm (n - 2k). \quad (5)$$

Возведение в квадрат обеих частей равенства (5) есть квадратичным ограничением (3). В случае оптимальной бисекции графа равенство (5) и его квадратичный аналог (3) упрощаются и переходят соответственно в ограничения

$$\sum_{i=1}^n x_i = 0 \quad \text{и} \quad \left(\sum_{i=1}^n x_i \right)^2 = 0.$$

1. Анализ задачи (1)–(3). Задача (1)–(3) является многоэкстремальной задачей нелинейного программирования, где целевая функция (1) и ограничения (2) и (3) есть квадратичные функции. Она попадает в класс невыпуклых оптимизационных задач квадратичного типа.

Чтобы оценить сверху значение глобального максимума $mc(G, w, k)$, целесообразно применить методiku Лагранжевых двойственных оценок для задач квадратичного типа [2, 3]. Пусть Ψ^* – наилучшая оценка из класса лагранжевых двойственных квадратичных оценок [3] для задачи (1)–(3). Для нее справедливо $\Psi^* \geq mc(G, w, k)$, а точность оценки Ψ^* (т.е. близость ее к $mc(G, w, k)$) зависит от структуры графа G и от весов его ребер.

Для самого общего вида оптимизационной задачи квадратичного типа на максимум нахождение оценки вида Ψ^* сводится к задаче минимизации выпуклой недифференцируемой функции, определенной на параметрическом (линейно зависящем от множителей Лагранжа) семействе неположительно определенных симметричных матриц. Оценка Ψ^* может быть найдена с помощью методов недифференцируемой оптимизации [3], однако ее нахождение связано с трудностями учета области неположительной определенности семейства симметричных матриц. Для задачи (1)–(3) этих трудностей можно избежать, если учесть специфику квадратичных функций цели и ограничений.

Специфика задачи (1)–(3) состоит в том, что целевая функция (1) и ограничения (2) и (3) есть квадратичными "однородными" функциями, т.е. не содержат линейных членов от переменных x_i , $i = 1, \dots, n$. Свойство "однородности" квадратичных функций в задаче (1)–(3) позволяет избежать решения задачи минимизации выпуклой функции, которая определена на классе неположительно определенных матриц. Нахождение оценки Ψ^* для задачи (1)–(3) можно свести к безусловной задаче минимизации выпуклой функции, т.е. условие на неположительную определенность семейства симметричных матриц можно учесть автоматически. Так, например, этот факт был использован в [4] для задачи о максимальном разрезе графа. Здесь применим аналогичный подход для нахождения оценки Ψ^* для задачи (1)–(3). При этом понадобятся следующие вспомогательные результаты о точных штрафных функциях.

3. Теорема Пшеничного о точных штрафных функциях. Пусть имеется следующая задача оптимизации:

$$\min_{y \in M} f_0(y) \quad (6)$$

$$f_i(y) \leq 0, \quad i = 1, \dots, m, \quad (7)$$

где $f_i(y)$, $i = 0, 1, \dots, m$, – непрерывные функции, $M \subset R^n$ – некоторое множество. Для учета ограничений (7) используем негладкую штраф-

фнующую функцию в виде функции максимума:

$$\begin{aligned}\Phi_N(y) &= f_0(y) + N \times F(y), \\ F(y) &= \max\{0, f_1(y), \dots, f_m(y)\}.\end{aligned}\quad (8)$$

В работе [5] приводится теорема, которая для специального вида задач (6)–(7) дает возможность устанавливать точное значение штрафного множителя N таким, чтобы решение безусловной задачи минимизации функции (8) совпадало с решением задачи (6)–(7).

Этот результат заключается в следующем. Рассмотрим семейство параметрических задач

$$V(z) = \inf \{f_0(y) : f_i(y) \leq z_i; i = 1, \dots, m; y \in M\}, \quad (9)$$

зависящее от вектора $z \in E^m$. Очевидно, что

$$V(0) = \inf \{f_0(y) : f_i(y) \leq 0; i = 1, \dots, m; y \in M\}$$

совпадает с решением задачи (6)–(7). Справедлива следующая теорема [5].

Теорема 1. Пусть $\inf_{t>0} \frac{V(te) - V(0)}{t} = -L > -\infty$, где e – m -мерный вектор, все компоненты которого равны единице. Если $N > L$, то тогда точки минимума задач $V(0)$ и $\inf_{y \in M} \Phi_N(y)$, где $\Phi_N(y)$ определяется по формуле (8), совпадают.

Теорема 1 позволяет установить точное значение штрафного множителя при использовании для решения задачи (6)–(7) негладкой штрафной функции в виде функции максимума.

4. Нахождение верхней оценки Ψ^* . Основной результат статьи сформулируем в форме следующей теоремы.

Теорема 2. Пусть $u \in R^{n+1}$ и $K(u)$ – симметричная $(n \times n)$ -матрица, элементы которой $K_{ij}(u) \forall i, j = 1, \dots, n$ линейно зависят от компонент вектора u и определяются по правилу

$$K_{ij}(u) = K_{ji}(u) = \begin{cases} -u_i - u_{n+1}, & \text{если } i = j, \\ -u_{n+1} - w_{ij}, & \text{если } (i, j) \in E, \\ -u_{n+1}, & \text{если } (i, j) \notin E. \end{cases}$$

Пусть $\lambda_{\max}(A)$ – максимальное собственное число симметричной матрицы A .

Тогда глобальному максимуму $mc(G, w, k)$ в задаче (1)–(3) соответствует верхняя оценка

$$\Psi^* = \frac{1}{2} \sum_{(i,j) \in E} w_{ij} + \frac{1}{4} \psi^*, \quad (10)$$

где ψ^* – решение безусловной задачи минимизации выпуклой функции

$$\psi^* = \min_{u \in R_{n+1}} \left(\sum_{i=1}^n u_i + (n - 2k)^2 u_{n+1} + S \max\{0, \lambda_{max}(K(u))\} \right) \quad (11)$$

при $S > n$.

Доказательство. Доказательство теоремы 2 проведем в два этапа. На первом этапе покажем справедливость (10), предполагая, что ψ^* – лагранжева оценка для некоторой новой задачи квадратичного типа. На втором этапе покажем, что ψ^* вычисляется согласно (11), где нижняя граница на негладкий штраф S , равная n , устанавливается с помощью теоремы 1.

Первый этап. Учитывая (4), функцию $Q(x)$ можно записать в таком виде:

$$Q(x) = \frac{1}{2} \sum_{(i,j) \in E} w_{ij} - \frac{1}{2} \sum_{(i,j) \in E} w_{ij} x_i x_j = \frac{1}{2} \sum_{(i,j) \in E} w_{ij} + \frac{1}{4} \sum_{(i,j) \in E} (-2) w_{ij} x_i x_j.$$

Следовательно, задача (1)–(3) эквивалентна задаче

$$mc(G, w, k) = \frac{1}{2} \sum_{(i,j) \in E} w_{ij} + \frac{1}{4} q^*, \quad (12)$$

где q^* – решение такой задачи квадратичного типа

$$q^* = \max_{x \in R^n} \sum_{(i,j) \in E} (-2) w_{ij} x_i x_j \quad (13)$$

при ограничениях:

$$x_i^2 = 1, \quad \forall i = 1, \dots, n, \quad (14)$$

$$\left(\sum_{i=1}^n x_i \right)^2 = (n - 2k)^2. \quad (15)$$

Пусть ψ^* – верхняя лагранжева оценка для задачи (13)–(15). Тогда из (12) имеем

$$mc(G, w, k) = \frac{1}{2} \sum_{(i,j) \in E} w_{ij} + \frac{1}{4} q^* \leq \Psi^* = \frac{1}{2} \sum_{(i,j) \in E} w_{ij} + \frac{1}{4} \psi^*,$$

что означает справедливость равенства (10) при условии, что ψ^* – верхняя лагранжева оценка, соответствующая задаче (13)–(15).

Второй этап будет состоять в нахождении оценки ψ^* для задачи (13)–(15), согласно методике Н.З.Шора [3], и доказательстве того, что ψ^* вычисляется согласно (11).

Пусть $u \in R^{n+1}$ – вектор множителей Лагранжа, соответствующий ограничениям (14)–(15). Для задачи (13)–(15) запишем функцию Лагранжа

$$\begin{aligned} L(x, u) &= \sum_{(i,j) \in E} (-2)w_{ij}x_i x_j + \sum_{i=1}^n u_i(1-x_i^2) + \\ &+ u_{n+1} \left((n-2k)^2 - \left(\sum_{i=1}^n x_i \right)^2 \right) = \\ &= x^T K(u)x + \sum_{i=1}^n u_i + u_{n+1} (n-2k)^2. \end{aligned}$$

Рассмотрим функцию $\psi(u) = \max_{x \in R^n} L(x, u)$. Она имеет следующую форму

$$\psi(u) = \begin{cases} +\infty, & \text{если } \lambda_{max}(K(u)) > 0, \\ \sum_{i=1}^n u_i + u_{n+1} (n-2k)^2, & \text{если } \lambda_{max}(K(u)) \leq 0, \end{cases}$$

где $\lambda_{max}(K(u))$ – максимальное собственное число матрицы $K(u)$. Наилучшая двойственная Лагранжева оценка – $\psi^* = \min_{u \in R^{n+1}} \psi(u)$, которая может быть получена в результате решения задачи выпуклого программирования:

$$\psi^* = \min_{u \in R^{n+1}} \left(\sum_{i=1}^n u_i + u_{n+1} (n-2k)^2 \right) \quad (16)$$

при ограничении

$$\lambda_{max}(K(u)) \leq 0. \quad (17)$$

Заменим задачу (16)–(17) задачей безусловной оптимизации, используя для учета ограничения (17) негладкую штрафную функцию в виде функции максимума:

$$\psi(S) = \min_{u \in R^{n+1}} \left(\sum_{i=1}^n u_i + (n-2k)^2 u_{n+1} + S \cdot \max\{0, \lambda_{max}(K(u))\} \right), \quad (18)$$

где S – штрафной множитель ($S > 0$).

Покажем, что если штрафной множитель S установить больше чем n , то решения задачи (18) и задачи (16)–(17) совпадают. Согласно (9) функция $V(te)$ для задачи (16)–(17) имеет вид

$$\begin{aligned} V(t) &= \inf \left\{ \sum_{i=1}^n u_i + (n-2k)^2 u_{n+1} : \lambda_{max}(K(u)) \leq t \right\} = \\ &= \inf \left\{ \sum_{i=1}^n u_i + (n-2k)^2 u_{n+1} : \lambda_{max}(K(u) - tI) \leq 0 \right\} = \\ &= \inf \left\{ \sum_{i=1}^n (u_i + t) + (n-2k)^2 u_{n+1} - nt : \lambda_{max}(K(u) - tI) \leq 0 \right\} = \\ &= \inf \left\{ \sum_{i=1}^n \tilde{u}_i + (n-2k)^2 \tilde{u}_{n+1} - nt : \lambda_{max}(K(\tilde{u})) \leq 0 \right\} = V(0) - nt, \end{aligned}$$

где $\tilde{u}_i = u_i + t$, а $\tilde{u}_{n+1} = u_{n+1}$. Тогда

$$\inf_{t>0} \frac{V(t) - V(0)}{t} = \inf_{t>0} \frac{V(0) - nt - V(0)}{t} = -n = -L > -\infty.$$

и согласно теореме 1 при $S > L = n$

$$\psi(S) = \min_{u \in R^{n+1}} \left(\sum_{i=1}^n u_i + (n-2k)^2 u_{n+1} + S \max\{0, \lambda_{max}(K(u))\} \right) = \psi^*.$$

Теорема доказана. \square

Таким образом, теорема 2 дает возможность нахождения верхней оценки взвешенной мощности максимального разреза графа с заданными количествами вершин в обоих подмножествах разбиения свести к задаче минимизации выпуклой недифференцируемой функции. Для ее решения целесообразно применять r -алгоритм [3], который на практике зарекомендовал себя эффективным методом решения подобных задач.

Литература

1. ГЭРИ М., ДЖОНСОН Д. Вычислительные машины и труднорешаемые задачи. – М.: Мир, 1982. – 416 с.

2. ШОР Н.З., СТЕЦЕНКО С.И. Квадратичные экстремальные задачи и недифференцируемая оптимизация. – Киев: Наукова думка, 1989. – 208 с.
3. SHOR N.Z. Nondifferentiable optimization and polynomial problems. – Boston; Dordrecht; London: Kluwer Academic Publishers, 1998. – 412 p.
4. ШОР Н.З., БЕРЕЗОВСКИЙ О.А. Новые алгоритмы решения задачи о максимальном взвешенном разрезе графа // Кибернетика и системный анализ. – 1995. – № 2. – С. 100–106.
5. ПШЕНИЧНЫЙ Б.Н. Метод линеаризации. – М.: Наука, 1983. – 136 с.

Получено 05.06.2006

Квадратичная оптимизационная задача для максимального k -плекса в неориентированном графе

П.И. Стецюк

Рукопись на 19.08.2006

В статье формулируется квадратичная оптимизационная задача для нахождения максимального k -плекса неориентированного графа. Анализируются лагранжевые двойственные оценки для этого класса задач и их улучшение за счет добавления определенного вида функционально избыточных ограничений.

Введение

Подмножества из множества вершин графа, которые обладают определенным, наперед заданным свойством, описывают важные структурные свойства графа. Задачи нахождения максимальных по мощности таких подмножеств имеют разнообразные приложения [1]. Одной из таких задач есть задача нахождения максимального k -плекса неориентированного графа [2, 3]. Она возникает и активно используется при анализе социальных, телекоммуникационных и других сетей.

Понятие k -плекса для неориентированного графа введено в [2]. Здесь k – некоторое натуральное число, $k = 1, 2, \dots$. Когда $k = 1$, то k -плекс совпадает с кликой (полный подграф) графа. Когда $k > 1$, то k -плекс есть ослаблением понятия клики графа и подразумевает более слабые требования на включение вершины в k -плекс, чем требования на включение вершины в клику. Понятие, противоположное k -плексу, есть ко- k -плекс неориентированного графа. Если $k > 1$, то ко- k -плекс есть ослаблением понятия независимого множества вершин графа (известного также как внутренне устойчивое множество). Когда $k = 1$, то ко- k -плекс совпадает с независимым множеством вершин графа.

Пусть $G = (V, E)$ – неориентированный граф с множеством вершин $V = \{1, \dots, n\}$ и множеством ребер E . Ребро графа G , связывающее вершины $i \in V$ и $j \in V$, условимся обозначать $(i, j) \in E$. Для графа G будем использовать и другую форму его задания: $G = (V, \Gamma)$, где $\Gamma = \{\Gamma(i), i = 1, \dots, n\}$, где $\Gamma(i)$ – конечные вершины тех дуг, у которых начальной вершиной есть вершина i . Количество ребер графа G в обеих

формах связаны соотношением: $|E| = \frac{1}{2} \sum_{i \in V} |\Gamma(i)|$. Дополнительный граф к G будем обозначать $\bar{G} = (V, \bar{E})$, и $\bar{G} = (V, \bar{\Gamma})$, где $(i, j) \in \bar{E}$ и $\bar{\Gamma} = \{\bar{\Gamma}(i), i = 1, \dots, n\}$.

Подмножество вершин S из V называется k -плексом графа G , если степень каждой вершины в индуцированном подграфе $G[S]$ (подграфе, порожденном подмножеством S) есть не меньше, чем $|S| - k$. Т.е. $S \subset V$ есть k -плексом, если выполняется следующее условие

$$\deg_{G[S]}(i) = |\Gamma(i) \cap S| \geq |S| - k \quad \forall i \in S.$$

k -плекс есть максимальным по включению (maximal), если он не содержится ни в каком другом k -плексе. Наибольший из максимальных по включению k -плексов называется максимальным, его размер есть k -плексным числом графа G , которое обозначается $\rho_k(G)$ [3]. Очевидно, что 1-плекс есть кликой графа G , так как степень каждой вершины в индуцированном подграфе $G[S]$ есть не меньше, чем $|S| - 1$, что означает, что каждая из вершин в подграфе $G[S]$ связана со всеми остальными вершинами, т.е. подграф $G[S]$ есть полным подграфом (кликой) графа G . В данном случае $\rho_1(G) = \omega(G)$, где $\omega(G)$ – кликовое число графа G , которое соответствует мощности максимальной клики в графе G .

Ко- k -плекс графа G [3] есть обобщением понятия независимого множества вершин графа G . Подмножество вершин S из V называется ко- k -плексом графа G , если выполняется следующее условие

$$\deg_{G[S]}(i) = |\Gamma(i) \cap S| \leq k - 1 \quad \forall i \in S.$$

Другими словами, $S \subset V$ есть ко- k -плексом, если степень каждой вершины в индуцированном подграфе $G[S]$ есть не больше, чем $(k - 1)$. Очевидно, что ко-1-плекс есть независимым множеством вершин графа G , так как степень каждой вершины в индуцированном подграфе $G[S]$ есть равной нулю, что означает, что каждая из вершин в подграфе $G[S]$ не связана ни с одной из остальных вершин подграфа $G[S]$. Заметим, что ко- k -плекс и k -плекс для графа G находятся в такой же связи как клика графа G и независимое множество вершин графа G . Другими словами, подмножество S есть ко- k -плексом графа G тогда и только тогда, когда S есть k -плексом для дополнительного графа \bar{G} .

В работе [3] проблема нахождения максимального k -плекса ($\rho_k(G)$) сформулирована в форме задачи булевого линейного программирования. В данной работе проблему нахождения $\rho_k(G)$ сформулируем в форме квадратичной оптимизационной задачи и проведем ее анализ, ориентированный на применение к квадратичной задаче техники лагранжевых двойственных оценок [4, 5].

Последовательность изложения материала статьи следующая. В разделе 2 опишем множество допустимых решений для k -плекса графа G с помощью системы квадратичных ограничений. В разделе 3 дадим квадратичную задачу для нахождения $\rho_k(G)$ и проанализируем ее связь с булевой линейной постановкой из [3]. Здесь же рассмотрим семейства функционально избыточных ограничений для уточнения лагранжевых двойственных оценок в квадратичной задаче, которые базируются на использовании ограничений линейной булевой задачи для максимального k -плекса [3]. В разделе 3 проанализируем связь квадратичной задачи для 1-плекса с известной формулировкой квадратичной задачи для нахождения максимальной клики графа. В разделе 4 рассмотрим лагранжевые двойственные оценки для простейших квадратичных задач нахождения $\rho_2(G)$ и покажем, что лагранжевые двойственные оценки можно улучшить посредством добавления функционально избыточных ограничений, рассмотренных в разделе 2.

1. Квадратичные ограничения для k -плекса.

Пусть вершине $i \in V$ ($i \in \{1, 2, \dots, n\}$) соответствует булева переменная $x_i \in \{0, 1\}$ такая, что

$$x_i = \begin{cases} 1, & \text{если } i \in S, \\ 0, & \text{если } i \in V/S. \end{cases}$$

Булевые переменные x_i , $i = 1, \dots, n$ опишем с помощью квадратичных ограничений равенств

$$x_i^2 - x_i = 0, \quad \forall i \in V. \quad (1)$$

Построим квадратичные ограничения, чтобы подмножество S было k -плексом. Эти ограничения должны задавать требования на то, чтобы степень каждой вершины $i \in S$ в подграфе $G[S]$ была не меньше, чем $|S| - k$, или другими словами, чтобы в подграфе $G[S]$ количество дуг, выходящих из каждой вершины $i \in S$ было не меньше, чем $|S| - k$.

Пусть вершина i принадлежит подмножеству S , т.е. $x_i = 1$. Обозначим $N_e(i)$ – степень вершины i в подграфе $G[S]$, т.е. количество дуг, которые выходят из вершины $i \in S$. Тогда в подграфе $G[S]$ степени вершин из подмножества S задаются с помощью семейства соотношений:

$$N_e(i) = \sum_{j \in \Gamma(i)} x_j, \quad \forall i \in S. \quad (2)$$

Учитывая, что $|S| = \sum_{j \in V} x_j$, и требуя, чтобы множество S было k -плексом имеем неравенства

$$N_e(i) \geq |S| - k = \sum_{j \in V} x_j - k, \quad \forall i \in S. \quad (3)$$

Из соотношений (2) и (3) имеем семейство неравенств

$$\sum_{j \in \Gamma(i)} x_j \geq \sum_{j \in V} x_j - k, \quad \forall i \in S, \quad (4)$$

которые задают требования на то, чтобы все те вершины $i \in V$, для которых $x_i = 1$, образовывали k -плекс.

Однако, неравенства по типу (4) не будут выполняться для тех вершин $i \in V$, для которых $x_i = 0$, т.е. для всех $i \in V/S$. Для того, чтобы получить квадратичные неравенства, которые будут справедливыми и для переменных $x_i = 0$, достаточно обе части неравенства вида (4), соответствующего вершине i умножить на переменную x_i . Учитывая, положительность переменной x_i , знак неравенства после умножения не изменится, и в результате мы получим такие неравенства

$$x_i \left(\sum_{j \in \Gamma(i)} x_j \right) \geq x_i \left(\sum_{j \in V} x_j - k \right), \quad \forall i \in V,$$

которые перепишем в следующей форме

$$\sum_{j \in \Gamma(i)} x_i x_j \geq \sum_{j \in V} x_i x_j - k x_i, \quad \forall i \in V. \quad (5)$$

Квадратичные неравенства (5) в сочетании с ограничениями (1) полностью описывают требования, чтобы вершины i принадлежали k -плексу. На самом деле, неравенства (5) будут справедливы для тех вершин i , для которых $x_i = 1$, так как они переходят в ограничения (4). Неравенства (5) будут справедливы и для всех вершин i , для которых $x_i = 0$, так как они переходят во всегда справедливое неравенство $0 \geq 0$.

Понятно, что с помощью ограничений в форме равенств (1) и форме неравенств (5) можно описать допустимые булевы решения, соответствующие k -плексу. Эти ограничения можно использовать при построении квадратичной оптимизационной задачи для нахождения максимального k -плекса. При этом содержательный смысл ограничений (5) будет связан с интерпретацией степени вершины, как этого требует понятие k -плекса, т.е. степень вершины $i \in S$ больше либо равна, чем

$|S| - k$. На самом деле, правая часть ограничения (5) для вершины i , принадлежащей k -плексу, содержит количество ребер, которые выходят из i -ой вершины, учитывая, что $x_i x_j = 1$, тогда когда обе переменные x_i и x_j равны единице. По этой же причине правая часть неравенства (5) содержит полное количество вершин в S из которого вычитается натуральное число k , т.е. $|S| - k$.

Однако, неравенство (5) можно упростить, сделав его более экономным. Заметим, что

$$\sum_{j \in V} x_j = \sum_{j \in \Gamma(i)} x_j + \sum_{j \in \bar{\Gamma}(i)} x_j + x_i.$$

Тогда, неравенства (5) можно переписать следующим образом:

$$\sum_{j \in \Gamma(i)} x_i x_j \geq \left(\sum_{j \in \Gamma(i)} x_i x_j + \sum_{j \in \bar{\Gamma}(i)} x_i x_j + x_i^2 \right) - k x_i, \quad \forall i \in V,$$

откуда имеем

$$\sum_{j \in \bar{\Gamma}(i)} x_i x_j \leq k x_i - x_i^2 \quad \forall i \in V,$$

Учитывая, что $x_i^2 = x_i$ из (1), последние неравенства можно переписать как

$$\sum_{j \in \Gamma(i)} x_i x_j \leq (k - 1) x_i, \quad \forall i \in V, \quad (6)$$

которые есть более экономной заменой неравенствам (5). Неравенства (6) в сочетании с равенствами (1) мы положим в основу квадратичной модели для нахождения максимального k -плекса графа G .

Отметим, что содержательный смысл квадратичным неравенствам (6) можно придать иной, чем неравенствам (5). Неравенства (6) связаны с дополнительным графом \bar{G} и задают описание такого подмножества вершин S , чтобы степень вершины в индуцированном этим подмножеством подграфе $\bar{G}[S]$ была меньше либо равна, чем $(k - 1)$. На самом деле, для тех вершин $i \in V$, для которых $x_i = 1$, неравенства (6) равносильны таким неравенствам

$$\sum_{j \in \bar{\Gamma}(i)} x_j \leq (k - 1), \quad \forall i \in S,$$

а для тех вершин i , для которых $x_i = 0$, они равносильны правильным неравенствам $0 \leq 0$. Поэтому, описание подмножества S согласно неравенств (6) и равенств (1) более логично интерпретировать как описание k -плекса для дополнительного графа \bar{G} .

Итак, совершенно формальный переход, который мы осуществили от неравенств (5) к неравенствам (6), означает простой факт, а именно, подмножество S есть k -плексом графа G тогда и только тогда, когда S есть ко- k -плексом дополнительного графа \bar{G} . В силу этого квадратичная постановка задачи для нахождения максимального k -плекса графа G будет равносильна задаче нахождения максимального ко- k -плекса графа \bar{G} . И это не зависит от того на какую форму (5) либо (6) для квадратичных ограничений-неравенств ориентироваться при формулировке квадратичной задачи.

2. Квадратичная задача для $\rho_k(G)$.

Учитывая, что ограничения (1) и (6) описывают множество допустимых решений для k -плекса, для нахождения максимального k -плекса графа G сформулируем оптимизационную квадратичную задачу в следующей форме:

$$\rho_k(G) = \max_x \sum_{i \in V} x_i \quad (7)$$

при ограничениях:

$$\sum_{j \in \bar{\Gamma}(i)} x_i x_j \leq (k-1)x_i, \quad \forall i \in V \quad (8)$$

$$x_i^2 - x_i = 0, \quad \forall i \in V. \quad (9)$$

Понятно, что задачу (7)–(9) можно интерпретировать как задачу нахождения максимального ко- k -плекса графа \bar{G} .

Из формулировки (7)–(9) легко получить формулировку квадратичной оптимизационной задачи для ко- k -плекса графа G . Для этого достаточно в сумме левой части ограничения (8) взамен суммирования по $j \in \bar{\Gamma}(i)$ использовать суммирование по $j \in \Gamma(i)$. Другими словами, если ограничения (8) заменить на ограничения

$$\sum_{j \in \Gamma(i)} x_i x_j \leq (k-1)x_i, \quad \forall i \in V \quad (8')$$

то мы приходим к формулировке квадратичной оптимизационной задачи для нахождения максимального ко- k -плекса графа G .

Формулировку задачи (7)–(9) можно получить из задачи булевого линейного программирования, предложенной в [3],

$$\rho_k(G) = \max_x \sum_{i \in V} x_i \quad (10)$$

при ограничениях:

$$\sum_{j \in \bar{\Gamma}(i)} x_j \leq (k-1)x_i + \bar{d}_i(1-x_i), \quad \forall i \in V \quad (11)$$

$$x_i \in \{0, 1\}, \quad \forall i \in V. \quad (12)$$

Здесь $\bar{d}_i = |\bar{\Gamma}(i)|$. Линейные ограничения (11) построены по подобной схеме, которую в предыдущем разделе мы использовали при переходе от ограничений (4), справедливых для $i \in S$, т.е. когда $x_i = 1$, к ограничениям (5), которые есть справедливыми и для $i \in V/S$, т.е. когда $x_i = 0$. Однако, правило для того чтобы ограничения (11) были справедливы для всех $i \in V$ здесь иное. Так, когда $x_i = 1$, то ограничения (11) выполняются, как неравенства

$$\sum_{j \in \Gamma(i)} x_j \leq (k-1), \quad \forall i \in S$$

которые должны быть справедливы для ко- k -плекса графа \bar{G} , совпадающего с k -плексом графа G . Когда $x_i = 0$, то ограничения (11) переходят в неравенства

$$\sum_{j \in \bar{\Gamma}(i)} x_j \leq |\bar{\Gamma}(i)| \quad \forall i \in V/S,$$

которые есть справедливыми из-за использования в качестве верхней границы на степени вершин, не входящих в ко- k -плекс графа \bar{G} , максимально возможного количества ребер, выходящих из каждой из вершин графа \bar{G} . Те из этих ограничений, где не все переменные под знаком суммы равны единице, будут избыточными.

Из задачи линейного булевого программирования (10)–(12) легко получить квадратичную задачу (7)–(9). Для этого следует ограничение (12) заменить на их нелинейный аналог (9), а соответствующее i -ой вершине ограничение из (11) умножить на переменную x_i . Учитывая положительность переменных x_i , $i = 1, \dots, n$ знаки неравенств при умножении не изменятся, и в результате получим

$$\sum_{j \in \bar{\Gamma}(i)} x_i x_j \leq (k-1)x_i^2 + \bar{d}_i(1-x_i)x_i, \quad \forall i \in V,$$

откуда с учетом того, что $x_i(1-x_i) = x_i - x_i^2 = 0$ для всех $i \in V$ приходим к ограничениям (8).

Понятно, что каждая из задач (7)–(9) и (10)–(12) имеет свои преимущества и свои недостатки. Так, например, самое значительное преимущество квадратичной задачи над линейной булевой задачей состоит в том, что незначительным усовершенствованием задачи (7)–(9) можно сформулировать квадратичные оптимизационные задачи для нахождения максимальных по размеру подмножеств графа, с более сильными свойствами, чем k -плекс или ко- k -плекс, а именно, ужесточив требование на включение вершин в эти подмножества. Так, например, условимся под "строгим" k -плексом графа G понимать подмножество его вершин, для которых степень вершины равна $|S| - k$. Аналогично, под "строгим" ко- k -плексом – когда степень вершины равна $(k - 1)$. Тогда, для того чтобы сформулировать квадратичные задачи для нахождения максимальных таких подмножеств, достаточно в ограничениях (8) и (8') взамен неравенств использовать равенства. Другими словами, для нахождения "строгого" k -плекса графа G ограничения (8) следует заменить на ограничения

$$\sum_{j \in \bar{\Gamma}(i)} x_i x_j = (k - 1)x_i, \quad \forall i \in V,$$

а для нахождения "строгого" ко- k -плекса графа G ограничения (8') следует заменить на такие ограничения

$$\sum_{j \in \Gamma(i)} x_i x_j = (k - 1)x_i, \quad \forall i \in V.$$

К преимуществу, линейной булевой задачи над квадратичной можно отнести тот факт, что для задачи (10)–(12) легко считать верхние оценки для $\rho_k(G)$, релаксируя ограничения (12). В ряде случаев, например, когда $\rho_k(G)$ будет больше $n/3$, то эти оценки, как правило, могут оказаться достаточно хорошими по качеству оценками сверху для $\rho_k(G)$. В результате для некоторых специальных графов на базе метода ветвей и границ можно реализовать быстрые алгоритмы для нахождения $\rho_k(G)$.

Нахождение верхних оценок для квадратичной задачи (7)–(9) более трудоемко, чем для релаксированной задачи (10)–(12). Так, например, если в качестве таковых использовать лагранжевые двойственные оценки [4, 5], базирующиеся на методах недифференцируемой оптимизации, то нахождение таких оценок требует больше затрат по времени, чем в случае задач линейного программирования. Однако, для ряда графов лагранжевые двойственные оценки могут оказаться значительно более точными верхними оценками, чем линейные оценки. Более того, если к задаче (7)–(9) добавлять функционально избыточные ограничения [5], то имеется резерв для уточнения лагранжевых двойственных оценок.

Линейные ограничения (11) можно использовать для построения функционально избыточных ограничений с тем, чтобы улучшить точность лагранжевых двойственных оценок в многоэкстремальных квадратичных задачах (7)–(9). Одна из возможных схем здесь может быть аналогичной той, которая использовалась Н.З.Шором для задачи о максимальном независимом множестве вершин графа [5], стр. 250. Она состоит в добавлении к задаче (7)–(9) функционально избыточных ограничений, которые получены если каждое из линейных ограничений в (11) домножить на те переменные x_l , которые не входят в это ограничение. Так, например, в результате имеем

$$\sum_{j \in \bar{\Gamma}(i)} x_j \leq (k-1)x_i + \bar{d}_i(1-x_i) \times x_l, \quad \forall l, i \in V : l \neq i,$$

откуда следует такое семейство из $n(n-1)$ функционально избыточных квадратичных ограничений

$$\sum_{j \in \bar{\Gamma}(i)} x_l x_j \leq (k-1)x_l x_i + \bar{d}_i(1-x_i)x_l, \quad \forall l, i \in V : l \neq i. \quad (**)$$

Другая форма функционально избыточных ограничений на основе линейных ограничений (11) может быть получена умножением их на $(1-x_l)$, $l = 1, \dots, n$. Здесь уже можно использовать $l = i$, учитывая что они дают новые квадратичные ограничения в форме неравенств. В результате имеем n^2 ограничений

$$\sum_{j \in \bar{\Gamma}(i)} x_j \leq (k-1)x_i + \bar{d}_i(1-x_i) \times (1-x_l), \quad \forall l, i \in V,$$

откуда следует такое семейство из n^2 функционально избыточных квадратичных ограничений

$$\sum_{j \in \bar{\Gamma}(i)} x_l x_j - \sum_{j \in \bar{\Gamma}(i)} x_j \leq (k-1)x_l x_i + \bar{d}_i(1-x_i)x_l, \quad \forall l, i \in V. \quad (***)$$

В разделе 5 на простейших примерах квадратичных задач вида (7)–(9) покажем, что использование функционально избыточных ограничений в форме (**) и (***) позволяет улучшить точность лагранжевых двойственных оценок.

3. Квадратичные задачи для максимального 1-плекса и максимальной клики графа

Учитывая, что для графа G 1-плекс совпадает с кликой, то естественно, что нахождение максимального 1-плекса графа G согласно задаче (7)–(9) соответствует нахождению максимальной клики графа G . Когда $k = 1$, то правая часть неравенств (11) равна нулю и задача (7)–(9) переходит в такую квадратичную оптимизационную задачу

$$\rho_1(G) = \omega(G) = \max \sum_{i \in V} x_i \quad (13)$$

при ограничениях:

$$\sum_{j \in \Gamma(i)} x_i x_j \leq 0, \quad \forall i \in V \quad (14)$$

$$x_i^2 - x_i = 0, \quad \forall i \in V. \quad (15)$$

Задачу (13)–(15) можно рассматривать как одну из возможных квадратичных формулировок задачи о нахождении максимальной клики графа G . Задаче (13)–(15) соответствует ровно n квадратичных ограничений неравенств (14).

Более известной квадратичной формулировкой задачи для максимальной клики графа G есть квадратичная оптимизационная задача [5]:

$$\omega(G) = \max \sum_{i \in V} x_i \quad (16)$$

при ограничениях:

$$x_i x_j = 0, \forall (i, j) \in \bar{E} \quad (17)$$

$$x_i^2 - x_i = 0, \quad \forall i \in V. \quad (18)$$

Задачу (16)–(18) можно интерпретировать как очевидное следствие из задачи (13)–(15), учитывая, что все переменные x_i , $i = 1, \dots, n$ положительны. На самом деле, на основании квадратичных ограничений (14), которые требуют чтобы сумма положительных произведений пар переменных была меньше либо равна нулю, можно записать квадратичные равенства, где каждое отдельное произведение двух переменных, которое входит в ограничения (14), равно нулю. Именно это и отражают ограничения (17).

Какая из квадратичных моделей лучше для максимальной клики графа G ? Формально, задача (13)–(15) выигрывает тем, что она содержит меньшее количество квадратичных ограничений, чем задача (16)–(18). Однако она проигрывает в точности лагранжевой двойственной оценки.

Пусть $\psi_{\rho_1}^*(G)$ – оптимальная лагранжева оценка для задачи (13)–(15), а $\psi_{\omega}^*(G)$ – оптимальная лагранжева оценка для задачи (16)–(18). Каждая из них будет верхней оценкой для $\omega(G)$. Верхняя оценка $\psi_{\rho_1}^*(G)$ будет, как правило, хуже, чем оценка $\psi_{\omega}^*(G)$. Это и понятно, так как для задачи (16)–(18) имеется полная свобода на выбор множителей Лагранжа, соответствующих ограничениям (17), т.е. каждому из ребер дополнительного графа \bar{E} будет соответствовать свой множитель Лагранжа. Для ограничений (14) таких множителей Лагранжа всего n , т.е. свой множитель Лагранжа будет соответствовать каждой из вершин графа G . Кроме того, свобода множителей Лагранжа для ограничений (14) ограничена еще и тем, что для каждого из этих множителей требуется учесть его неположительность. Учет неположительности дает свободы множителям Лагранжа в задаче (13)–(14) меньше, чем было бы тогда, когда ограничения (14) заменить на такие

$$\sum_{j \in \bar{\Gamma}(i)} x_i x_j = 0, \quad \forall i \in V,$$

которые следуют из ограничений (17), если их сгруппировать по вершинам графа G . Заметим, что квадратичную задачу с этими ограничениями взамен ограничений (14) можно проинтерпретировать как задачу нахождения максимального "строгого" 1-плекса графа G . И, естественно, что лагранжева двойственная оценка для квадратичной задачи, соответствующей максимальному "строгому" 1-плексу графа G , будет во многих случаях более точной, чем оценка $\psi_{\rho_1}^*(G)$.

Проиллюстрируем поведение лагранжевых двойственных оценок $\psi_{\rho_1}^*(G)$ и $\psi_{\omega}^*(G)$ на примере двух простейших графов G_1 и G_2 (рис.1), содержащих всего пять вершин. Здесь ребра графов G_1 и G_2 даны сплошной линией, а ребра дополнительных графов \bar{G}_1 и \bar{G}_2 – пунктирной линией. Так, для графа G_1 на рис 1а) системы ограничений для задачи (13)–(15) и задачи (16)–(18) будут следующими:

$$\begin{array}{ll} x_1 x_2 + x_1 x_5 \leq 0, & x_1 x_2 = 0, \\ x_1 x_2 + x_2 x_3 \leq 0, & x_2 x_3 = 0, \\ x_2 x_3 + x_3 x_4 \leq 0, & x_3 x_4 = 0, \\ x_3 x_4 + x_4 x_5 \leq 0, & x_4 x_5 = 0, \\ x_4 x_5 + x_1 x_5 \leq 0, & x_1 x_5 = 0. \end{array}$$

Лагранжевые двойственные оценки $\psi_{\rho_1}^*(G_1)$ и $\psi_{\omega}^*(G_1)$ одинаковы и равны $\sqrt{5}$. Они есть неточными оценками сверху, так как $\omega(G_1) = 2$. Пример для графа G_1 демонстрирует, что лагранжевые двойственные оценки $\psi_{\rho_1}^*(G)$ и $\psi_{\omega}^*(G)$ могут совпадать.

В то же время это не так для графа G_2 на рис 16), который отличается от G_1 тем, что из того удалено ребро (1,3). Для него системы ограничений задачи (13)–(15) и задачи (16)–(18) будут следующими:

$$\begin{array}{ll} x_1x_2 + x_1x_3 + x_1x_5 \leq 0, & x_1x_2 = 0, \\ x_1x_2 + x_2x_3 \leq 0, & x_1x_3 = 0, \\ x_1x_3 + x_2x_3 + x_3x_4 \leq 0, & x_2x_3 = 0, \\ x_3x_4 + x_4x_5 \leq 0, & x_3x_4 = 0, \\ x_4x_5 + x_1x_5 \leq 0, & x_4x_5 = 0, \\ & x_1x_5 = 0, \end{array}$$

Здесь уже оценки $\psi_{\rho_1}^*(G_2)$ и $\psi_{\omega}^*(G_2)$ различные. Так, оценка $\psi_{\omega}^*(G_2) = 2.0$ и есть точной верхней оценкой для $\omega(G_2)$. Оценка $\psi_{\rho_1}^*(G_2) = 2.15222$ не есть точной, и нуждается в уточнении. Следовательно, пример графа G_2 показывает, что лагранжевая двойственная оценка для задачи (16)–(18) есть предпочтительней, чем для задачи (13)–(15). Заметим, что если для графа G_2 ограничения неравенства заменить на ограничения в форме равенств, что соответствует нахождению максимального "строго" 1-плекса графа G_1 , то лагранжева двойственная оценка станет равной двум и оказывается точной верхней оценкой для $\omega(G_1)$.

Если требуется найти более точную лагранжевую двойственную оценку для максимальной клики графа G в рамках рассмотренных двух квадратичных задач, то лучше пользоваться квадратичной задачей в форме (16)–(18), чем квадратичной задачей в форме (13)–(15), которая для 1-плекса следует из квадратичной формулировки (7)–(9). Аналогичная ситуация будет иметь место и для квадратичной задачи [5], которая связана с нахождением максимального независимого множества вершин графа. Ей также следует отдавать предпочтение перед квадратичной задачей, которая получается для максимального ко-1-плекса графа G как следствие из квадратичной задачи для нахождения максимального ко- k -плекса графа G .

Однако, ситуация может измениться если речь пойдет не об нахождении оценок $\psi_{\rho_1}^*(G)$ и $\psi_{\omega}^*(G)$, а об их улучшении посредством добавления функционально избыточных ограничений. Тогда, квадратичная задача (13)–(15) может оказаться предпочтительней, хотя бы потому, что к ней можно добавить большее количество функционально избыточных ограничений. Кроме того, когда $k \geq 2$, то для нахождения максималь-

ного k -плекса вряд-ли можно найти альтернативу квадратичной формулировке (7)–(9). Поэтому исследование точности лагранжевых двойственных оценок для задачи (7)–(9), которая усилена за счет добавления функционально избыточных квадратичных ограничений представляет определенный интерес. В результате может оказаться, что прямолинейный перенос таких оценок на частный случай $k = 1$ для ряда графов может привести к достаточно хорошим по точности верхним оценкам для $\omega(G)$.

Для графов G_1 и G_2 из рис. 1 ниже покажем, что с помощью функционально избыточных ограничений из семейств (**) и (***) лагранжевые двойственные оценки можно сделать точными верхними оценками для $\rho_2(G_1)$ и $\rho_2(G_2)$.

4. Об улучшении лагранжевых двойственных оценок для максимального 2-плекса

Рассмотрим задачу нахождения максимального 2-плекса для графа G_1 из рис. 1а). Она формулируется следующей квадратичной задачей:

$$\rho_2(G_1) = \max_{x \in R^5} x_1 + x_2 + x_3 + x_4 + x_5 \quad (19)$$

при ограничениях:

$$\begin{aligned} x_1 x_2 + x_1 x_5 &\leq x_1, \\ x_1 x_2 + x_2 x_3 &\leq x_2, \\ x_2 x_3 + x_3 x_4 &\leq x_3, \\ x_3 x_4 + x_4 x_5 &\leq x_4, \\ x_4 x_5 + x_1 x_5 &\leq x_5, \end{aligned} \quad (20)$$

$$x_i^2 - x_i = 0, \quad i = 1, 2, 3, 4, 5. \quad (21)$$

Здесь $\rho_2(G_1) = 3$. В то же время лагранжева двойственная оценка для задачи (19)–(21) равна 3.618034. Она есть менее точной, чем линейная верхняя оценка, которая получается в результате релаксации задачи булевого программирования (10)–(12) и равна $10/3 = 3.33333$. Задаче линейного программирования соответствуют такие линейные ограничения

$$\begin{aligned} x_1 + x_2 + x_5 &\leq 2, \\ x_1 + x_2 + x_3 &\leq 2, \\ x_2 + x_3 + x_4 &\leq 2, \\ x_3 + x_4 + x_5 &\leq 2, \\ x_1 + x_4 + x_5 &\leq 2, \end{aligned} \quad (22)$$

которые следуют из ограничений (11).

Как лагранжевая двойственная оценка, так и линейная оценка есть неточными верхними оценками. Улучшить лагранжевую двойственную оценку, соответствующую задаче (19)–(21), можно, добавляя к квадратичной задаче (19)–(21) функционально избыточные ограничения из семейства (**). Они следуют из домножения линейных ограничений (22) на переменные x_i , $i = 1, \dots, n$. Так, например, последовательное добавление таких ограничений делает лагранжевую двойственную оценку все более точной. Вот такие оценки ψ^* получаются, если добавлять функционально избыточные квадратичные ограничения из семейства (**) в следующем порядке:

$+n$	<i>linear constraint</i>	$\times x_i$	<i>quadratic constraint</i>	ψ^*
+1	$x_1 + x_2 + x_5 \leq 2$,	$\times x_2$	$x_1x_2 + x_2x_5 \leq x_2$,	3.56878
+2	$x_1 + x_2 + x_3 \leq 2$,	$\times x_3$	$x_1x_3 + x_2x_3 \leq x_3$,	3.52727
+3	$x_2 + x_3 + x_4 \leq 2$,	$\times x_4$	$x_2x_4 + x_3x_4 \leq x_4$,	3.42929
+4	$x_3 + x_4 + x_5 \leq 2$,	$\times x_5$	$x_3x_5 + x_4x_5 \leq x_5$,	3.21298
+5	$x_1 + x_4 + x_5 \leq 2$,	$\times x_1$	$x_1x_4 + x_1x_5 \leq x_1$,	3.00000

В результате добавления четырех ограничений оценка ψ^* становится более точной верхней оценкой, чем линейная оценка. А в результате добавления всех пяти ограничений оценка ψ^* становится точной верхней оценкой для $\rho_2(G_1)$.

Похожая, но немного иная, ситуация имеет место и для задачи нахождения максимального 2-плекса для графа G_2 из рис. 1б). Она формулируется в форме квадратичной задачи

$$\rho_2(G_2) = \max_{x \in R^5} x_1 + x_2 + x_3 + x_4 + x_5 \quad (23)$$

при ограничениях:

$$\begin{aligned} x_1x_2 + x_1x_3 + x_1x_5 &\leq x_1, \\ x_1x_2 + x_2x_3 &\leq x_2, \\ x_1x_3 + x_2x_3 + x_3x_4 &\leq x_3, \\ x_3x_4 + x_4x_5 &\leq x_4, \\ x_4x_5 + x_1x_5 &\leq x_5, \end{aligned} \quad (24)$$

$$x_i^2 - x_i = 0, \quad i = 1, 2, 3, 4, 5. \quad (25)$$

и ей соответствует $\rho_2(G_2) = 3$. Лагранжева двойственная оценка для задачи (23)–(25) равна 3.37332 и она есть более точной, чем линейная

верхняя оценка, которая равна 3.4. Задаче линейного программирования соответствуют такие линейные ограничения вида (11)

$$\begin{aligned} 2x_1 + x_2 + x_3 + x_5 &\leq 3, \\ x_1 + x_2 + x_3 &\leq 2, \\ x_1 + x_2 + 2x_3 + x_4 &\leq 3, \\ x_3 + x_4 + x_5 &\leq 2, \\ x_1 + x_4 + x_5 &\leq 2. \end{aligned} \quad (26)$$

Учитывая, что лагранжева двойственная оценка есть более точной, чем линейная оценка, то, казалось бы, меньшими усилиями можно придти к точной оценке ψ^* посредством добавления функционально избыточных ограничений (**), т.е. тех, которые получены из линейных ограничений (26) умножением на переменные $x_i, i = 1, \dots, 5$. Однако, это оказывается не так, и максимальная по точности лагранжева двойственная оценка, которой можно достигнуть на этом пути, равна 3.02316. Однако, если добавлять функционально избыточные ограничения в форме (***), то можно добиться того, чтобы лагранжевая двойственная оценка стала точной. Так, например, после добавления к квадратичной задаче (23)–(25) функционально избыточного ограничения вида (**)

$$x_1x_1 + x_1x_4 + x_1x_5 - 2x_1 \leq 0,$$

которое получено из пятого ограничения (26) умножением на x_1 , оценка ψ^* становится равной 3.21576. После прибавления ограничения вида (***)

$$-x_1x_3 - x_1x_4 - x_1x_5 + 2x_1 + x_3 + x_4 + x_5 \leq 2,$$

которое получено умножением четвертого ограничения из (26) на $(1 - x_1)$ оценка ψ^* равна 3.01571. А после прибавления еще и ограничения вида (***)

$$-2x_1x_2 - x_2^2 - x_2x_3 - x_2x_5 + 2x_1 + 4x_2 + x_3 + x_5 \leq 3,$$

которое получено умножением первого ограничения из (26) на $(1 - x_2)$ оценка ψ^* становится точной и равна 3.0.

Из примеров квадратичных задач (19)–(21) и (23)–(25) для нахождения максимальных 2-плексов графов G_1 и G_2 видим, что функционально избыточные ограничения вида (**) и (***) играют значительную роль для улучшения точности лагранжевых двойственных оценок. Так, с их помощью лагранжеву двойственную оценку ψ^* можно сделать точной для $\rho_2(G_1)$ и $\rho_2(G_2)$. Однако, это не значит, что для произвольной

квадратичной задачи (7)–(9) с помощью ограничений вида (**) и (***) можно добиться того, чтобы лагранжева двойственная оценка была точной. Однако, для повышения точности лагранжевых двойственных оценок можно задействовать и другие способы для построения функционально избыточных ограничений из [5].

Заключение

В работе построена квадратичная оптимизационная задача для нахождения максимального k -плекса неориентированного графа. Показано, что подобного типа квадратичную задачу можно сформулировать и для нахождения максимального k -плекса графа. Проанализированы лагранжевые двойственные оценки для простейших задач этого типа, в частности и для максимального 1-плекса, который совпадает с максимальной кликой графа. Показано, что лагранжевые двойственные оценки можно улучшить посредством добавления функционально избыточных ограничений, которые следуют из ограничений линейной булевой задачи для максимального k -плекса [3], если их домножать на x_i , либо $(1 - x_i)$, $i = 1, \dots, n$.

Отметим, что задачи вида (7)–(9) могут служить генератором достаточно непростых тестовых примеров для квадратичных невыпуклых задач оптимизации. Более того, если рассмотреть задачи соответствующие нахождению максимального "строгого" k -плекса, т.е. квадратичные задачи вида (7)–(9), где ограничения неравенства (8) заменяются на ограничения в форме равенств, то эти тестовые примеры будут еще сложнее. Им будет соответствовать более значительный разрыв между лагранжевой двойственной оценкой и точным оптимальным значением квадратичной задачи.

Покажем, что даже для графов небольших размеров разрыв между лагранжевой двойственной оценкой и оптимальным значением функции может быть существенным. Так, например, нахождению максимального "строгого" 2-плекса для графа G_1 из рис. 1а) соответствует такая квадратичная оптимизационная задача

$$f^* = \max_{x \in R^5} x_1 + x_2 + x_3 + x_4 + x_5 \quad (27)$$

при ограничениях:

$$\begin{aligned} x_1 x_2 + x_1 x_5 &= x_1, \\ x_1 x_2 + x_2 x_3 &= x_2, \\ x_2 x_3 + x_3 x_4 &= x_3, \\ x_3 x_4 + x_4 x_5 &= x_4, \\ x_4 x_5 + x_1 x_5 &= x_5, \end{aligned} \quad (28)$$

$$x_i^2 - x_i = 0, \quad i = 1, 2, 3, 4, 5. \quad (29)$$

Здесь $f^* = 2$, а оценка $\psi^* = 3.618034$. Относительная точность оценки ψ^* определяется величиной

$$\frac{\psi^* - f^*}{f^*} = (3.618034 - 2)/2 = 0.809017$$

и составляет около 81%. С помощью добавления к задаче (27)–(29) функционально избыточных ограничений вида (**) оценку ψ^* можно понизить только до $\psi^* = 3.0$. Используя еще и функционально избыточные ограничения вида (***) оценку ψ^* можно понизить до $\psi^* = 2.40$. Для дальнейшего ее понижения требуется использовать другие приемы для построения функционально избыточных квадратичных ограничений.

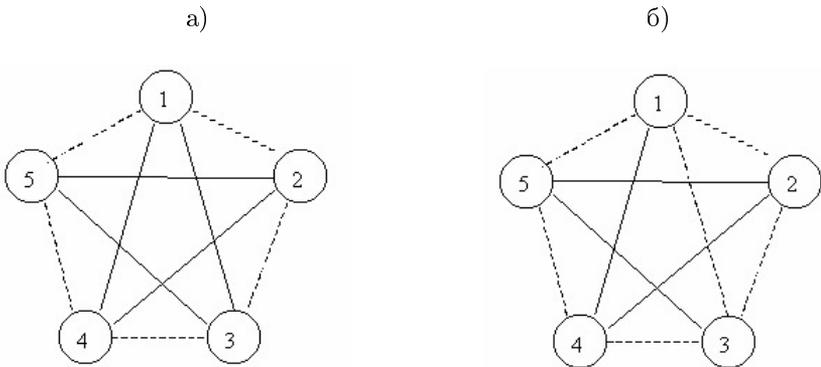


Рис. 1. Примеры графов G_1 и G_2

Применение техники лагранжевых двойственных оценок в сочетании с использованием функционально избыточных ограничений к задаче (7)–(9) и исследование эффективности таких оценок требует более детального исследования. Такое исследование предполагается провести в рамках гранта UKM2-2812-KV-06 (CRDF Cooperative Grants Program).

Литература

1. КРИСТОФИДЕС Н. Теория графов. Алгоритмический подход. – Издательство "Мир", Москва. – 1978. – 432 с.
2. SEIDMAN S.B., FOSTER B.L. A graph theoretic generalization of the clique concept // *Journal of Mathematical Sociology* **6**, 139–154 (1978).
3. V.BALANSUNDARAM, S.BUTENKO, I.V.HICKS, S.SACHDEVA Clique Relaxations in Social Network Analysis: The Maximum k -plex Problem // *Journal of Global Optimization* (to appear).
4. ШОР Н.З., СТЕЦЕНКО С.И. Квадратичные экстремальные задачи и недифференцируемая оптимизация. – Киев: Наукова думка, 1989. – 208 с.
5. SHOR N.Z. Nondifferentiable Optimization and Polynomial Problems. – Dordrecht, Kluwer. – 1998. – 394 p.

Квадратична задача для максимального k -плекса в неорієнтованому графі

П. І. Стецюк, Т. О. Бардадим, В. І. Ляшко

*Журнал обчислювальної та прикладної математики. – 2017. –
№ 1(124). – С. 71–78.*

У статті сформульовано квадратичну оптимізаційну задачу для знаходження максимального k -плекса у неорієнтованому графі. Показано, що квадратичну задачу можна отримати з відомої лінійної булевої задачі для максимального k -плекса. Наведено два сімейства функціонально надлишкових квадратичних обмежень, які отримано за допомогою обмежень булевої задачі.

ВСТУП

Структурні властивості графа суттєво залежать від того, чи мають окремі підмножини з множини вершин графа певні, наперед задані властивості. Задачі знаходження максимальних за потужністю таких підмножин мають різноманітні застосування [1]. Окремим випадком таких підмножин є k -плекси. Це поняття для неорієнтованого графа було введене в [2], а задача знаходження максимального k -плекса неорієнтованого графа [2, 3] виникає та активно використовується при аналізі соціальних, телекомунікаційних та інших мереж [4]. При $k = 1$ k -плекс збігається з клікою (повним підграфом) графа. При $k > 1$ k -плекс є ослабленням поняття кліки графа і відповідає слабкішим вимогам на включення вершини в k -плекс, ніж вимоги на включення вершини до кліки.

У статті [3] задачу знаходження максимального k -плекса сформульовано у формі задачі булевого лінійного програмування. Нижче цю задачу сформульовано як квадратичну оптимізаційну задачу та проведено її аналіз, орієнтований на застосування техніки лагранжевих двоїстих оцінок [5].

Послідовність викладу матеріалу буде такою: у розділі 1 наведено загальні відомості про k -плекс, у розділі 2 описано множину допустимих розв'язків для k -плекса графа G за допомогою системи квадратичних обмежень. У розділі 3 розглянуто квадратичну задачу знаходже-

ння максимального k -плекса та проаналізовано її зв'язок із булевою лінійною постановкою з [3]. Там же розглянуто сімейства функціонально надлишкових обмежень для уточнення лагранжевих двоїстих оцінок у квадратичній задачі, що базуються на використанні обмежень лінійної булевої задачі для максимального k -плекса.

1. ЗАГАЛЬНІ ВІДОМОСТІ ПРО k -ПЛЕКС

Нехай $G = (V, E)$ — неорієнтований граф із множиною вершин $V = \{1, \dots, n\}$ та множиною ребер E . Ребро графа G , що зв'язує вершини $i \in V$ та $j \in V$, будемо позначати $(i, j) \in E$. Для графа G буде використовуватися також інша форма його представлення: $G = (V, \Gamma)$, де $\Gamma = \{\Gamma(i), i = 1, \dots, n\}$, а $\Gamma(i)$ — кінцеві вершини тих дуг, у яких початковою вершиною є вершина i . Кількість ребер графа G в обох представленнях зв'язані співвідношенням: $|E| = \frac{1}{2} \sum_{i \in V} |\Gamma(i)|$. Комплементарний до G граф будемо позначати $\overline{G} = (V, \overline{E})$ або $\overline{G} = (V, \overline{\Gamma})$, де $(i, j) \in \overline{E}$ і $\overline{\Gamma} = \{\overline{\Gamma}(i), i = 1, \dots, n\}$.

Означення 1. Підмножина вершин S із V називається k -плексом графа G , якщо ступінь кожної вершини в індукованому підграфі $G[S]$ (підграфі, породженому підмножиною S) є не меншою, ніж $|S| - k$.

Підмножина $S \subset V$ є k -плексом, якщо виконується така умова

$$\deg_{G[S]}(i) = |\Gamma(i) \cap S| \geq |S| - k \quad \forall i \in S.$$

k -Плекс є максимальним за включенням (maximal), якщо він не міститься ні в якому іншому k -плексі. Найбільший з максимальних за включенням k -плексів називається максимальним (maximum), його розмір називається k -плексним числом графа G та позначається $\rho_k(G)$ [3]. Очевидно, що 1-плекс є клікою графа G , тому що ступінь кожної вершини в індукованому підграфі $G[S]$ не менше, ніж $|S| - 1$, а це означає, що кожна з вершин у підграфі $G[S]$ зв'язана з усіма іншими вершинами, тобто підграф $G[S]$ є повним підграфом (клікою) графа G . У даному випадку $\rho_1(G) = \omega(G)$, де $\omega(G)$ — клікове число графа G (розмір його максимальної кліки).

Поняття so - k -плекса графа G , також введене в [3], є узагальненням поняття незалежної множини вершин графа G . При $k = 1$ so - k -плекс збігається з незалежною множиною вершин графа. При $k > 1$ so - k -плекс є ослабленням поняття незалежної множини вершин графа (відомої також як внутрішньо стійка множина).

Означення 2. Підмножина вершин S з V називається $co-k$ -плексом графа G , якщо виконується така умова

$$\deg_{G[S]}(i) = |\Gamma(i) \cap S| \leq k - 1 \quad \forall i \in S.$$

Отже, $S \subset V$ є $co-k$ -плексом, якщо ступінь кожної вершини в індукованому підграфі $G[S]$ є не більшою, ніж $k - 1$. Очевидно, що $co-1$ -плекс є незалежною множиною вершин графа G , оскільки ступінь кожної вершини в індуційованому підграфі $G[S]$ дорівнює нулю, а це означає, що жодна з вершин у підграфі $G[S]$ не зв'язана з жодною з інших вершин підграфа $G[S]$. Відмітимо, що $co-k$ -плекс і k -плекс для графа G знаходяться в такому ж зв'язку як кліка графа G і незалежна множина вершин графа G . Тому підмножина S є $co-k$ -плексом графа G тоді й тільки тоді, коли S є k -плексом для комплементарного графа \overline{G} .

2. КВАДРАТИЧНІ ОБМЕЖЕННЯ ДЛЯ k -ПЛЕКСА

Нехай вершині $i \in V$ ($i = 1, 2, \dots$) відповідає булева змінна $x_i \in \{0, 1\}$ така, що

$$x_i = \begin{cases} 1, & \text{якщо } i \in S, \\ 0, & \text{якщо } i \in V \setminus S. \end{cases}$$

Булеві змінні x_i , $i = 1, \dots, n$ будуть описуватися за допомогою квадратичних обмежень-рівностей

$$x_i^2 - x_i = 0 \quad \forall i \in V. \quad (1)$$

Побудуємо такі квадратичні обмеження, щоб підмножина S була k -плексом. Ці обмеження повинні задавати вимоги на те, щоб ступінь кожної вершини $i \in S$ у підграфі $G[S]$ була не меншою за $|S| - k$, тобто, щоб у підграфі $G[S]$ кількість дуг, що виходять із кожної вершини $i \in S$ була не меншою за $|S| - k$.

Нехай вершина i належить підмножині S , тобто $x_i = 1$. Позначимо через $N_e(i)$ ступінь вершини i у підграфі $G[S]$, тобто кількість дуг, що виходять з вершини $i \in S$. Тоді в підграфі $G[S]$ ступені вершин із підмножини S задаються за допомогою сімейства співвідношень

$$N_e(i) = \sum_{j \in \Gamma(i)} x_j \quad \forall i \in S. \quad (2)$$

З рівняння $|S| = \sum_{j \in V} x_j$ та умови, що множина S є k -плексом, одержуємо нерівності

$$N_e(i) \geq |S| - k = \sum_{j \in \Gamma(i)} x_j - k \quad \forall i \in S. \quad (3)$$

Із співвідношень (2) та (3) одержуємо сімейство нерівностей

$$\sum_{j \in \Gamma(i)} x_j \geq \sum_{j \in V} x_j - k \quad \forall i \in S, \quad (4)$$

при виконанні яких всі ті вершини $i \in V$, для яких $x_i = 1$, будуть утворювати k -плекс.

Однак, нерівності типу (4) не будуть виконуватися для тих вершин $i \in V$, для яких $x_i = 0$, тобто для всіх $x_i \in V \setminus S$. Для того, щоб одержати квадратичні нерівності, що будуть справедливими і для змінних $x_i = 0$, досить обидві частини нерівності вигляду (4), що відповідає вершині i , помножити на змінну x_i . З огляду на невід'ємність змінної x_i знак нерівності після множення не зміниться, і в результаті одержимо такі нерівності

$$x_i \left(\sum_{j \in \Gamma(i)} x_j \right) \geq x_i \left(\sum_{j \in V} x_j - k \right) \quad \forall i \in V,$$

які можна переписати у вигляді

$$\sum_{j \in \Gamma(i)} x_i x_j \geq \sum_{j \in V} x_i x_j - k x_i \quad \forall i \in V. \quad (5)$$

Квадратичні нерівності (5) разом із обмеженнями (1) повністю описують умови, за яких вершини i належать k -плексу. Дійсно, нерівності (5) будуть справедливими для тих вершин i , для яких $x_i = 1$, оскільки вони переходять в обмеження (4). Нерівності (5) будуть справедливими також для усіх вершин i , для яких $x_i = 0$, тому що вони переходять у тривіальну нерівність $0 \geq 0$.

Зрозуміло, що за допомогою обмежень у вигляді рівностей (1) та у вигляді нерівностей (5) можна описати допустимі булеві розв'язки, які відповідають k -плексу. При цьому зміст обмежень (5) буде пов'язаний із інтерпретацією ступеня вершини, як цього вимагає поняття k -плекса, — в ньому ступінь вершини $i \in S$ більше або дорівнює $|S| - k$. Дійсно, права частина обмеження (5) для вершини i , що належить k -плексу, вказує кількість ребер, що виходять з i -ї вершини, з урахуванням того, що $x_i x_j = 1$ лише тоді, коли обидві змінні x_i та x_j дорівнюють одиниці.

Нерівність (5) можна спростити. Помітимо, що

$$\sum_{j \in V} x_j = \sum_{j \in \Gamma(i)} x_j + \sum_{j \in \bar{\Gamma}(i)} x_j + x_i.$$

Тоді нерівності (5) можна переписати так

$$\sum_{j \in \Gamma(i)} x_i x_j \geq \left(\sum_{j \in \Gamma(i)} x_i x_j + \sum_{j \in \bar{\Gamma}(i)} x_i x_j + x_i^2 \right) - kx_i, \quad \forall i \in V,$$

звідки

$$\sum_{j \in \bar{\Gamma}(i)} x_i x_j \leq kx_i - x_i^2 \quad \forall i \in V.$$

З урахуванням того, що $x_i = x_i^2$ (див. формулу (1)), останні нерівності можна переписати так

$$\sum_{j \in \bar{\Gamma}(i)} x_i x_j \leq (k-1)x_i \quad \forall i \in V. \quad (6)$$

Нерівності (6) разом із рівностями (1) ми покладемо в основу квадратичної моделі для знаходження максимального k -плекса графа G .

Відмітимо, що квадратичним нерівностям (6) можна надати інший зміст, ніж нерівностям (5). Нерівності (6) пов'язані з комплементарним графом \bar{G} і описують таку підмножину вершин S , що ступінь вершини в індуційованому цією підмножиною підграфі $\bar{G}[S]$ не більша за $(k-1)$. Дійсно, для тих вершин $i \in V$, для яких $x_i = 1$, нерівності (6) рівносильні таким нерівностям

$$\sum_{j \in \bar{\Gamma}(i)} x_j \leq (k-1) \quad \forall i \in S,$$

а для тих вершин i , для яких $x_i = 0$, вони рівносильні тривіальним нерівностям $0 \leq 0$. Тому опис підмножини S за допомогою нерівностей (6) та рівностей (1) логічно інтерпретувати як опис co - k -плекса для комплементарного графа \bar{G} .

3. КВАДРАТИЧНА БУЛЕВА ЗАДАЧА ДЛЯ $\rho_k(G)$

Ураховуючи, що обмеження (1) та (6) описують множину допустимих варіантів утворення k -плекса, для знаходження максимального k -плекса графа G оптимізаційну квадратичну задачу можна сформулювати в такій формі

$$\rho_k(G) = \max_x \sum_{i \in V} x_i \quad (7)$$

при обмеженнях

$$\sum_{j \in \overline{\Gamma}(i)} x_i x_j \leq (k-1)x_i \quad \forall i \in V, \quad (8)$$

$$x_i^2 - x_i = 0 \quad \forall i \in V. \quad (9)$$

Зрозуміло, що задачу (7)–(9) можна інтерпретувати як задачу знаходження максимального co - k плекса графа \overline{G} . Із (7)–(9) легко одержати формулювання квадратичної оптимізаційної задачі для co - k -плекса графа \overline{G} . Для цього досить у сумі лівої частини обмеження (8) замість підсумовування по $j \in \overline{\Gamma}(i)$ використовувати підсумовування по $j \in \Gamma(i)$. Тобто, якщо обмеження (8) замінити на

$$\sum_{j \in \Gamma(i)} x_i x_j \leq (k-1)x_i \quad \forall i \in V, \quad (10)$$

то ми одержимо формулювання квадратичної оптимізаційної задачі знаходження максимального co - k -плекса графа \overline{G} .

Формулювання задачі (7)–(9) можна одержати із задачі булевого лінійного програмування, запропонованої в [3]:

$$\rho_k(G) = \max_x \sum_{i \in V} x_i \quad (11)$$

при обмеженнях

$$\sum_{j \in \overline{\Gamma}(i)} x_j \leq (k-1)x_i + \overline{d}_i(1-x_i) \quad \forall i \in V, \quad (12)$$

$$x_i \in \{0, 1\}, \quad \forall i \in V, \quad (13)$$

де $\overline{d}_i = |\overline{\Gamma}(i)|$. Лінійні обмеження (12) побудовано за схемою, подібною тій, що використовувалася у попередньому розділі при переході від обмежень (4), справедливих для $i \in S$ (при $x_i = 1$), до обмежень (5), що є справедливими також для $i \in V \setminus S$ (при $x_i = 0$). Однак, правило для того, щоб обмеження (12) виконувалися для будь-яких $i \in V \setminus S$, тут буде іншим. При $x_i = 1$ обмеження (12) виконуються як нерівності

$$\sum_{j \in \overline{\Gamma}(i)} x_j \leq (k-1) \quad \forall i \in S,$$

які повинні бути справедливими для $so-k$ -плекса графа \overline{G} , що збігається з k -плексом графа G . При $x_i = 0$ обмеження (12) переходять у нерівності

$$\sum_{j \in \overline{\Gamma}(i)} x_j \leq \overline{d}_i = |\overline{\Gamma}(i)| \quad \forall i \in V \setminus S,$$

які є справедливими, бо в якості верхньої границі на ступені вершин, що не входять у $so-k$ -плекс графа \overline{G} , використовується максимально можлива кількість ребер, що виходять з кожної з вершин графа \overline{G} . Ті з цих обмежень, де не всі змінні під знаком суми дорівнюють одиниці, будуть надлишковими.

Із задачі лінійного булевого програмування (11)–(13) легко одержати квадратичну задачу (7)–(9). Для цього слід обмеження (13) замінити на відповідний нелінійний аналог (9), а обмеження, яке відноситься до i -ї вершини з (12), помножити на змінну x_i . В силу невід'ємності змінних x_i , $i = 1, 2, \dots$ знаки нерівностей при множенні не зміняться, і в результаті одержимо

$$\sum_{j \in \overline{\Gamma}(i)} x_i x_j \leq (k-1)x_i^2 + \overline{d}_i(1-x_i)x_i \quad \forall i \in V,$$

звідки, з урахуванням того, що $(1-x_i)x_i = x_i - x_i^2 = 0$ для всіх $i \in V$, приходимо до обмежень (8).

Зрозуміло, що кожна з задач (7)–(9) та (11)–(13) має свої переваги та свої недоліки. Так, наприклад, найсуттєвіша перевага квадратичної задачі над лінійною булевою полягає в тому, що незначним удосконаленням задачі (7)–(9) можна сформулювати квадратичні оптимізаційні задачі знаходження максимальних за розміром підмножин графа із сильнішими властивостями, ніж k -плекс або $so-k$ -плекс, — достатньо лише посилити вимогу на включення вершин у ці підмножини. Так, наприклад, умовимся під „строгим“ k -плексом графа G розуміти підмножину його вершин, для яких ступінь вершини дорівнює $|S| - k$. Аналогічно введемо поняття „строгого“ $so-k$ -плекса: в ньому ступінь вершини дорівнює рівно $k - 1$. Для того щоб сформулювати квадратичні задачі знаходження максимальних із цих підмножин, досить в обмеженнях (8) та (10) замість нерівностей використовувати рівності. Для знаходження „строгого“ k -плекса графа G обмеження (8) слід замінити на обмеження

$$\sum_{j \in \overline{\Gamma}(i)} x_i x_j = (k-1)x_i, \quad \forall i \in V,$$

а для знаходження „строгого“ $co-k$ -плекса графа G обмеження (10) слід замінити на такі:

$$\sum_{j \in \Gamma(i)} x_i x_j = (k-1)x_i, \quad \forall i \in V.$$

До переваги лінійної булевої задачі над квадратичною можна віднести той факт, що для задачі (11)–(13) легко підраховувати верхні оцінки для $\rho_k(G)$ за допомогою релаксації обмеження (13). У ряді випадків, наприклад, коли $\rho_k(G)$ буде більше $n/3$, ці оцінки, як правило, можуть виявитися більш ефективними оцінками зверху для $\rho_k(G)$. У результаті для деяких спеціальних графів на базі методу гілок та границь можна реалізувати швидкі алгоритми для знаходження $\rho_k(G)$.

Знаходження верхніх оцінок для квадратичної задачі (7)–(9) є більш трудомістким, ніж для релаксованої задачі (11)–(13). Так, наприклад, якщо в такій якості використовувати лагранжеві двоїсті оцінки [4], [5], то знаходження таких оцінок за допомогою методів недиференційованої оптимізації вимагатиме більше часу, чим у випадку задач лінійного програмування. Однак, для ряду графів лагранжеві двоїсті оцінки можуть виявитися значно точнішими верхніми оцінками, ніж лінійні оцінки. Більш того, існує резерв для уточнення лагранжевих двоїстих оцінок задачі (7)–(9): це введення функціонально надлишкових обмежень [5].

Лінійні обмеження (12) можна використовувати для побудови функціонально надлишкових обмежень з метою покращити точність лагранжевих двоїстих оцінок у багатоекстремальних квадратичних задачах (7)–(9). Якщо скористатися схемою, що використовувалася Н. З. Шором для задачі про максимальну незалежну множину вершин графа [5, с. 250], то при цьому до задачі (7)–(9) додаються два види функціонально надлишкових обмежень. Обмеження першого виду одержано домноженням кожного з лінійних обмежень у (12) на ті змінні x_l , які не входять у це обмеження, тобто додаються $n(n-1)$ надлишкових обмежень вигляду

$$\sum_{j \in \bar{\Gamma}(i)} x_l x_j \leq (k-1)x_l x_i + \bar{d}_i(1-x_i)x_l \quad \forall i, l \in V, \quad i \neq l. \quad (14)$$

Функціонально надлишкові обмеження другого типу можна одержати з лінійних обмежень (12) домноженням на $1-x_l$, $l=1, 2, \dots$. Тут уже можна використовувати $i=l$, бо вони дають нові квадратичні обмеже-

ння у формі нерівностей. У результаті маємо n^2 обмежень

$$\sum_{j \in \Gamma(i)} x_l x_j - \sum_{j \in \overline{\Gamma}(i)} x_j \leq (k-1)x_l x_i + \overline{d}_i(1-x_i)x_l \quad \forall i, l \in V. \quad (15)$$

Зауважимо, що використання надлишкових обмежень (14), (15) значно збільшує розміри квадратичної задачі. Однак, якщо до попередньої задачі додавати тільки невелику кількість тих надлишкових обмежень, які уточнюють двоїсту оцінку для наступної задачі, то тоді загальна кількість обмежень у кінцевій квадратичній задачі буде невеликою. За такою ж схемою можна використовувати і інші види функціонально надлишкових обмежень для булевих задач, які розглядалися у роботах [6], [7].

Висновки

У роботі побудовано квадратичне формулювання оптимізаційної задачі знаходження максимального k -плекса для неорієнтованого графа. Спорідненість понять k -плекса та кліки дає підстави рекомендувати для її розв'язання підхід, пов'язаний з лагранжевими оцінками та використанням надлишкових обмежень. Цей підхід був запропонований Н. З. Шором та дав ряд важливих теоретичних результатів для задачі про максимальну незалежну множину вершин графа [5].

Роботу виконано за підтримки НАН України, проект 0117U000327.

Бібліографія

1. Кристофидес Н. Теория графов. Алгоритмический подход. — Мир: Москва, 1978. — 432 с.
2. Seidman S.B., Foster B.L. A graph theoretic generalization of the clique concept // J. of Math. Sociology. — 1978. — Vol. 6. — P. 139–154.
3. Balansundaram B., Butenko S., Hicks I. V. Clique Relaxations in Social Network Analysis: The Maximum k -plex Problem // Operations Research. — 2011. — Vol. 59, Number 1. — P. 133–142.
4. Нові мережево-орієнтовані методики для інформаційного аналізу великих масивів даних / П. І. Стецюк, М. Г. Журбенко, І. В. Сергієнко та інші // Звіт про науково-дослідну роботу М/163-2006,

- № держ. реєстрації 0106U010005. — Київ: Ін-т кібернетики ім. В.М. Глушкова НАН України, 2006. — 112 с.
5. Shor N. Z. *Nondifferentiable Optimization and Polynomial Problems*. — London/Boston/Dordrecht: Kluwer Academic Publishers, 1998. — 394 p.
 6. Стецюк П. И. О функционально избыточных ограничениях для булевых оптимизационных задач квадратичного типа // *Кибернетика и системный анализ*. — 2005. — № 6. — С. 168–172.
 7. Стецюк П. И. Новые модели квадратичного типа для задачи о максимальном взвешенном разрезе графа // *Кибернетика и системный анализ*. — 2006. — № 1. — С. 63–75.

Надійшла 28.04.2017

Властивості квадратичної задачі про максимальний k -плекс у неорієнтованому графі

П.І. Стецюк, В.І. Ляшко, Т.О. Бардадим

Наукові записки НаУКМА. – 2017.

Том 198. Комп'ютерні науки. – С. 8–13.

У статті розглянуто властивості верхніх оцінок для квадратичної задачі про максимальний k -плекс у неорієнтованому графі. Проаналізовано зв'язок квадратичної задачі для 1-плекса з відомим формулюванням квадратичної задачі для знаходження максимальної кліки графа. Наведено лагранжеві двоїсті оцінки для найпростіших квадратичних задач та показано, що їх можна покращити при додаванні функціонально надлишкових обмежень.

Вступ

Поняття k -плекса для неорієнтованого графа введено в [6] (k – деяке натуральне число). Якщо $k = 1$, то k -плекс збігається з клікою (повним підграфом) графа. При $k > 1$ k -плекс є ослабленням поняття кліки графа, і вимоги на включення вершини в k -плекс є слабшими, ніж вимоги на включення вершини в кліку. Ці поняття широко використовуються в соціології для виявлення та дослідження окремих підгруп населення, при кластеризації даних, для оптимізації інформаційних потоків у мережах тощо (див., напр., [2]– [5]). Слід зазначити, що зазвичай оптимізаційні задачі пошуку максимальних клік та k -плексів є NP-складними.

У статті буде проаналізовано зв'язок наведеного у [1] формулювання квадратичної задачі для 1-плекса з відомим формулюванням квадратичної задачі для знаходження максимальної кліки графа, буде розглянуто лагранжеві двоїсті оцінки для ряду квадратичних задач [7] та буде показано, що лагранжеві двоїсті оцінки можна покращити за допомогою додавання функціонально надлишкових обмежень, розглянутих у роботі [1]. Для обчислення лагранжевих двоїстих оцінок використано програму DSQTPr [8].

Означення та формулювання задачі

Нехай $G = (V, E)$ – неорієнтований граф із множиною вершин $V = \{1, \dots, n\}$ та множиною ребер E . Ребро графа G , що зв'язує вершини $i \in V$ та $j \in V$, умовимося позначати $(i, j) \in E$. Для графа G будемо використовувати також іншу форму його представлення: $G = (V, \Gamma)$, де $\Gamma = \{\Gamma(i), i = 1, \dots, n\}$, а $\Gamma(i)$ – кінцеві вершини тих дуг, у яких початковою вершиною є вершина i . Кількість ребер графа G в обох представленнях зв'язана співвідношенням: $|E| = \frac{1}{2} \sum_{i \in V} |\Gamma(i)|$. Комплементарний до G граф будемо позначати $\overline{G} = (V, \overline{E})$, і $\overline{\Gamma} = (V, \overline{\Gamma})$, де $(i, j) \in \overline{E}$ і $\overline{\Gamma} = \{\overline{\Gamma}(i), i = 1, \dots, n\}$.

Підмножину вершин S із V називають k -плексом графа G , якщо ступінь кожної вершини в індукованому підграфі $G[S]$ (підграфі, породженому підмножиною S) є не менше, ніж $|S| - k$. Тобто $S \subset V$ є k -плексом, якщо виконується така умова:

$$\deg_{G[S]}(i) = |\Gamma(i) \cap S| \geq |S| - k \quad \forall i \in S.$$

k -плекс є максимальним по включенню (maximal), якщо він не міститься в жодному іншому k -плексі. Найбільший із максимальних по включенню k -плексів називають максимальним (maximum), його розмір називають k -плексним числом графа G та позначають $\rho_k(G)$ [2]. Очевидно, що 1-плекс є клікою графа G , тому що ступінь кожної вершини в індукованому підграфі $G[S]$ є не менший, ніж $|S| - 1$, а це означає, що кожна з вершин у підграфі $G[S]$ зв'язана з усіма іншими вершинами, тобто підграф $G[S]$ є повним підграфом (клікою) графа G . У цьому випадку $\rho_1(G) = \omega(G)$, де $\omega(G)$ – клікове число графа G , що відповідає розміру максимальної кліки в графі G .

Для знаходження максимального k -плекса графа G у роботі [1] сформульовано оптимізаційну квадратичну задачу в такій формі:

$$\rho_k(G) = \max_x \sum_{i \in V} x_i \quad (1)$$

при обмеженнях:

$$\sum_{j \in \overline{\Gamma}(i)} x_i x_j \leq (k-1)x_i, \quad \forall i \in V, \quad (2)$$

$$x_i^2 - x_i = 0, \quad \forall i \in V. \quad (3)$$

Тут вершині $i \in V$ ($i = 1, 2, \dots$) відповідає булева змінна $x_i \in \{0, 1\}$ така, що $x_i = 1$, якщо $i \in S$, та $x_i = 0$, якщо $i \in V \setminus S$. Якщо у формулюванні задачі (1)–(3) квадратичні нерівності (2) замінити на квадратичні обмеження

$$\sum_{j \in \overline{\Gamma}(i)} x_i x_j = (k-1)x_i, \quad \forall i \in V,$$

то отримаємо формулювання квадратичної задачі для знаходження максимального "строого" k -плекса графа G , тобто підмножини його вершин, для яких ступінь вершини дорівнює $|S| - k$.

Максимальний 1-плекс та максимальна кліка

Оскільки для графа G 1-плекс збігається з клікою, природньо, що знаходження максимального 1-плекса графа G у задачі (1)–(3) відповідає знаходженню максимальної кліки графа G . Якщо $k = 1$, то права частина нерівностей (2) дорівнює нулю, і задача (1)–(3) переходить у таку квадратичну оптимізаційну задачу:

$$\rho_1(G) = \omega(G) = \max \sum_{i \in V} x_i \quad (4)$$

при обмеженнях:

$$\sum_{j \in \overline{\Gamma}(i)} x_i x_j \leq 0, \quad \forall i \in V, \quad (5)$$

$$x_i^2 - x_i = 0, \quad \forall i \in V. \quad (6)$$

Задачу (4)–(6) можна розглядати як одне з можливих квадратичних формулювань задачі про знаходження максимальної кліки графа G . Задачі (4)–(6) відповідає рівно n квадратичних обмежень нерівностей (5).

Більш відомим квадратичним формулюванням задачі для максимальної кліки графа G є квадратична оптимізаційна задача [7]:

$$\omega(G) = \max \sum_{i \in V} x_i \quad (7)$$

при обмеженнях:

$$x_i x_j = 0, \quad \forall (i, j) \in \overline{E}, \quad (8)$$

$$x_i^2 - x_i = 0, \quad \forall i \in V. \quad (9)$$

Задачу (7)–(9) можна інтерпретувати як очевидний наслідок задачі (4)–(6), тому що всі змінні x_i , $i = 1, \dots, n$ є невід’ємними. Дійсно, згідно квадратичних обмежень (5), сума невід’ємних добутків пар змінних повинна бути не більше нуля, а це еквівалентно (за умови невід’ємності) системі квадратичних рівностей, де кожен окремий добуток двох змінних, котрі входять в обмеження (5), дорівнює нулю. Це й становить зміст обмежень (8).

Яка з квадратичних моделей є кращою для знаходження максимальної кліки графа G ? Формально задача (4)–(6) має ту перевагу, що вона містить меншу кількість квадратичних обмежень, ніж задача (7)–(9). Однак вона програє в точності лагранжевій двоїстій оцінці.

Нехай $\psi_{\rho_1}^*(G)$ – оптимальна лагранжева оцінка для задачі (4)–(6), а $\psi_{\omega}^*(G)$ – оптимальна лагранжева оцінка для задачі (7)–(9). Кожна з них буде верхньою оцінкою для $\omega(G)$. Верхня оцінка $\psi_{\rho_1}^*(G)$ буде зазвичай гіршою за оцінку $\psi_{\omega}^*(G)$. Це можна пояснити тим, що в задачі (7)–(9) ми маємо більшу свободу у виборі множників Лагранжа, що відповідають обмеженням (8): кожному з ребер комплементарного графа \bar{G} буде відповідати свій множник Лагранжа. Для обмежень (5) таких множників Лагранжа усього n , тобто свій множник Лагранжа буде відповідати кожній з вершин графа G . Крім того, свобода на вибір множників Лагранжа для обмежень (5) обмежена ще й тим, що для кожного з цих множників потрібно врахувати його невід’ємність. Урахування невід’ємності обмежує можливості вибору множників Лагранжа в задачі (4)–(6) у порівнянні з випадком, коли обмеження (5) замінюються на

$$\sum_{j \in \Gamma(i)} x_i x_j = 0, \quad \forall i \in V,$$

(вони впливають із обмежень (8), якщо їх згрупувати за вершинами графа G). Зауважимо, що квадратичну задачу з цими обмеженнями замість обмежень (5) можна розглядати як задачу знаходження максимального "строогого" 1-плекса графа G . Природно, що лагранжева двоїста оцінка для квадратичної задачі, що відповідає максимальному "строогому" 1-плексу графа G , буде в багатьох випадках точнішою за оцінку $\psi_{\rho_1}^*(G)$.

Проілюструємо поведінку лагранжевих двоїстих оцінок $\psi_{\rho_1}^*(G)$ та $\psi_{\omega}^*(G)$ на прикладі двох графів G_1 і G_2 (рис. 1), що складаються усього з п’яти вершин. Тут ребра графів G_1 і G_2 позначено суцільною лінією,

а ребра комплементарних графів $\overline{G_1}$ і $\overline{G_2}$ – пунктирною. Для графа G_1 (рис. 1а)) системи обмежень для задач (4)–(6) та (7)–(9) будуть такими:

$$\begin{aligned} x_1 x_2 + x_1 x_5 &\leq 0, & x_1 x_2 &= 0, \\ x_1 x_2 + x_2 x_3 &\leq 0, & x_2 x_3 &= 0, \\ x_2 x_3 + x_3 x_4 &\leq 0, & x_3 x_4 &= 0, \\ x_3 x_4 + x_4 x_5 &\leq 0, & x_4 x_5 &= 0, \\ x_4 x_5 + x_1 x_5 &\leq 0, & x_1 x_5 &= 0. \end{aligned}$$

Лагранжеві двоїсті оцінки $\psi_{\rho_1}^*(G)$ та $\psi_{\omega}^*(G)$ однакові й дорівнюють $\sqrt{5}$. Вони є неточними оцінками зверху, тому що $\omega(G_1) = 2$. Приклад для графа G_1 демонструє, що лагранжеві двоїсті оцінки $\psi_{\rho_1}^*(G)$ та $\psi_{\omega}^*(G)$ можуть співпадати.

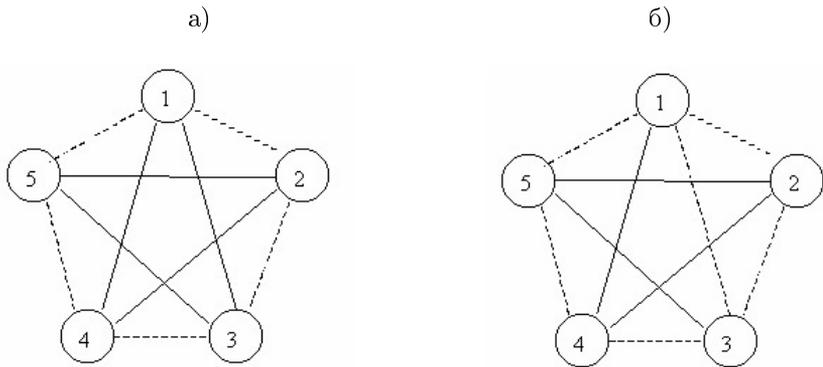


Рис. 1. Приклади графів G_1 та G_2

Це не виконується для графа G_2 на рис. 1б), що відрізняється від G_1 тим, що з нього вилучене ребро (1,3). Для графа G_2 системи обмежень задач (4)–(6) та (7)–(9) будуть такими:

$$\begin{aligned} x_1 x_2 + x_1 x_3 + x_1 x_5 &\leq 0, & x_1 x_2 &= 0, \\ x_1 x_2 + x_2 x_3 &\leq 0, & x_1 x_3 &= 0, \\ x_1 x_3 + x_2 x_3 + x_3 x_4 &\leq 0, & x_2 x_3 &= 0, \\ x_3 x_4 + x_4 x_5 &\leq 0, & x_3 x_4 &= 0, \\ x_4 x_5 + x_1 x_5 &\leq 0. & x_4 x_5 &= 0, \\ & & x_1 x_5 &= 0. \end{aligned}$$

Цього разу оцінки $\psi_{\rho_1}^*(G)$ та $\psi_{\omega}^*(G)$ різні. Зокрема, оцінка $\psi_{\omega}^*(G) = 2.0$ є точною верхньою оцінкою для $\omega(G_2)$. Оцінка $\psi_{\rho_1}^*(G) = 2.15222$ не є точною і потребує уточнення. Отже, приклад графа G_2 показує, що лагранжева двоїста оцінка для задачі (7)–(9) має перевагу над відповідною оцінкою задачі (4)–(6).

Зауважимо, що якщо для графа G_2 обмеження нерівності заміни-ти на обмеження у формі рівностей, що відповідає знаходженню максимального "строго" 1-плекса графа G_1 , то лагранжева двоїста оцінка стане рівною двом і буде точною верхньою оцінкою для $\omega(G_1)$.

Якщо потрібно знайти більш точну лагранжеву двоїсту оцінку для максимальної кліки графа G за допомогою однієї з двох розглянутих квадратичних задач, то краще користуватися квадратичною задачею у формі (7)–(9), ніж квадратичною задачею у формі (4)–(6), що для 1-плекса впливає з квадратичного формулювання (1)–(3). Аналогічна ситуація матиме місце і для квадратичної задачі [7], що пов'язана із знаходженням максимальної незалежної множини вершин графа.

Однак ситуація може змінитися, якщо мова йтиме не про знаходження оцінок $\psi_{\rho_1}^*(G)$ та $\psi_{\omega}^*(G)$, а про їх поліпшення за допомогою додавання функціонально надлишкових обмежень. Тоді квадратична задача (4)–(6) може мати переваги хоча б тому, що до неї можна додати більшу кількість функціонально надлишкових обмежень. Крім того, коли $k \geq 2$, то для знаходження максимального k -плекса навряд чи можна знайти альтернативу квадратичному формулюванню (1)–(3). Тому дослідження точності лагранжевих двоїстих оцінок для задачі (1)–(3), що підсилена за рахунок додавання функціонально надлишкових квадратичних обмежень, становить значний інтерес. У результаті може виявитися, що безпосереднє перенесення таких оцінок на окремий випадок $k = 1$ для ряду графів може призвести до досить хороших за точністю верхніх оцінок для $\omega(G)$.

Для уточнення верхніх оцінок можна використовувати квадратичні нерівності з двох сімейств функціонально надлишкових квадратичних обмежень [7]. Вони побудовані за схемою, яку використовував Н.З.Шор для задачі про максимальну незалежну множину вершин графа [7, с. 250]. За основу береться формулювання задачі про максимальний k -плекс, яке має вигляд:

$$\rho_k(G) = \max_x \sum_{i \in V} x_i \quad (10)$$

при обмеженнях:

$$\sum_{j \in \bar{\Gamma}(i)} x_j \leq (k-1)x_i + \bar{d}_i(1-x_i), \quad \forall i \in V, \quad (11)$$

$$x_i \in \{0, 1\}, \quad \forall i \in V, \quad (12)$$

де $\bar{d}_i = |\bar{\Gamma}(i)|$. Спочатку додаватимуться функціонально надлишкові обмеження, одержані домноженням кожного з лінійних обмежень у (11) на ті змінні x_l , які не входять у це обмеження, тобто додаються функціонально надлишкові обмеження першого типу із сімейства $n(n-1)$ обмежень вигляду

$$\sum_{j \in \bar{\Gamma}(i)} x_l x_j \leq (k-1)x_l x_i + \bar{d}_i(1-x_i)x_l, \quad \forall i, l \in V, \quad i \neq l. \quad (13)$$

Далі додаються функціонально надлишкові обмеження другого типу, одержані з лінійних обмежень (11) домноженням на $1-x_l$, $l = 1, 2, \dots$. Тут уже можна використовувати $i = l$, бо вони дають нові квадратичні обмеження у формі нерівностей. У результаті одержуємо n^2 обмежень

$$\sum_{j \in \bar{\Gamma}(i)} x_l x_j - \sum_{j \in \bar{\Gamma}(i)} x_j \leq (k-1)x_l x_i + \bar{d}_i(1-x_i)x_l, \quad \forall i, l \in V. \quad (14)$$

Для графів G_1 та G_2 з рис. 1 нижче буде показано, що за допомогою функціонально надлишкових обмежень із сімейств (13) і (14) лагранжеві двоїсті оцінки можна зробити точними верхніми оцінками для $\rho_2(G_1)$ та $\rho_2(G_2)$.

Уточнення лагранжевих двоїстих оцінок для максимального 2-плекса

Розглянемо задачу знаходження максимального 2-плекса для графа G_1 з рис. 1а). Вона формулюється у вигляді такої квадратичної задачі:

$$\rho_2(G_1) = \max_{x \in R^5} (x_1 + x_2 + x_3 + x_4 + x_5) \quad (15)$$

при обмеженнях:

$$\begin{cases} x_1 x_2 + x_1 x_5 \leq x_1, \\ x_1 x_2 + x_2 x_3 \leq x_2, \\ x_2 x_3 + x_3 x_4 \leq x_3, \\ x_3 x_4 + x_4 x_5 \leq x_4, \\ x_4 x_5 + x_1 x_5 \leq x_5. \end{cases} \quad (16)$$

$$x_i^2 - x_i = 0, \quad i = 1, 2, 3, 4, 5. \quad (17)$$

Тут $\rho_2(G_1) = 3$. У той же час лагранжева двоїста оцінка для задачі (15)–(17) дорівнює 3.618034. Вона є менш точною, ніж лінійна верхня оцінка, що виходить у результаті релаксації задачі булевого програмування (10)–(12) і дорівнює $10/3 = 3.33333$. Задачі лінійного програмування відповідають такі лінійні обмеження:

$$\begin{cases} x_1 + x_2 + x_5 \leq 2, \\ x_1 + x_2 + x_3 \leq 2, \\ x_2 + x_3 + x_4 \leq 2, \\ x_3 + x_4 + x_5 \leq 2, \\ x_1 + x_4 + x_5 \leq 2. \end{cases}$$

які впливають з обмежень (11).

Як лагранжева двоїста оцінка, так і лінійна оцінка є неточними верхніми оцінками. Покращити лагранжеву двоїсту оцінку, що відповідає задачі (15)–(17), можна, додаючи до квадратичної задачі (15)–(17) функціонально надлишкові обмеження із сімейства (13). Вони побудовані домноженням лінійних обмежень (14) на змінні x_i , $i = 1, \dots, n$. Послідовне додавання таких обмежень робить лагранжеву двоїсту оцінку більш точною. При послідовному додаванні функціонально надлишкових квадратичних обмежень із сімейства (13), вказаних у табл. 1, спостерігається монотонне зменшення оцінки ψ^* . У результаті додавання чотирьох додаткових обмежень верхня оцінка ψ^* стає точнішою за лінійну. А в результаті додавання всіх п'яти обмежень оцінка ψ^* стає точною верхньою оцінкою для $\rho_2(G_1)$.

Схожа, але дещо інша, ситуація має місце і для задачі знаходження максимального 2-плекса для графа G_2 з рис. 16). Вона формулюється у вигляді квадратичної задачі

$$\rho_2(G_2) = \max_{x \in R^5} (x_1 + x_2 + x_3 + x_4 + x_5) \quad (18)$$

Табл. 1. Покращення верхньої оцінки для задачі (15)–(17)

$+n$	Лінійні обмеження	$\times x_i$	Квадратичні обмеження	ψ^*
+1	$x_1 + x_2 + x_5 \leq 2$	$\times x_2$	$x_1x_2 + x_2x_5 \leq x_2$	3.56878
+2	$x_1 + x_2 + x_3 \leq 2$	$\times x_3$	$x_1x_3 + x_2x_3 \leq x_3$	3.52727
+3	$x_2 + x_3 + x_4 \leq 2$	$\times x_4$	$x_2x_4 + x_3x_4 \leq x_4$	3.42929
+4	$x_3 + x_4 + x_5 \leq 2$	$\times x_5$	$x_3x_5 + x_4x_5 \leq x_5$	3.21298
+5	$x_1 + x_4 + x_5 \leq 2$	$\times x_1$	$x_1x_4 + x_1x_5 \leq x_1$	3.00000

при обмеженнях:

$$\left\{ \begin{array}{l} x_1x_2 + x_1x_3 + x_1x_5 \leq x_1, \\ x_1x_2 + x_2x_3 \leq x_2, \\ x_1x_2 + x_2x_3 + x_3x_4 \leq x_3, \\ x_3x_4 + x_4x_5 \leq x_4, \\ x_4x_5 + x_1x_5 \leq x_5. \end{array} \right. \quad (19)$$

$$x_i^2 - x_i = 0, \quad i = 1, 2, 3, 4, 5, \quad (20)$$

і їй відповідає $\rho_2(G_2) = 3$. Лагранжева двоїста оцінка для задачі (18)–(20) дорівнює 3.37332 і є точнішою, ніж лінійна верхня оцінка, що дорівнює 3.4. Задачі лінійного програмування відповідають такі лінійні обмеження виду (11):

$$\left\{ \begin{array}{l} 2x_1 + x_2 + x_3 + x_5 \leq 3, \\ x_1 + x_2 + x_3 \leq 2, \\ x_1 + x_2 + 2x_3 + x_4 \leq 3, \\ x_3 + x_4 + x_5 \leq 2, \\ x_1 + x_4 + x_5 \leq 2. \end{array} \right. \quad (21)$$

Враховуючи, що лагранжева двоїста оцінка є точнішою за лінійну, може скластися враження, що точну оцінку ψ^* простіше одержати за допомогою додавання функціонально надлишкових обмежень (13), тобто тих, котрі побудовані з лінійних обмежень (21) домноженням на змінні x_i , $i = 1, \dots, 5$. Однак це не так, і максимальна за точністю лагранжева двоїста оцінка, яку можна досягти на цьому шляху, дорівнює 3.02316. Лише при додаванні функціонально надлишкових обмежень у вигляді (14) можна домогтися того, щоб лагранжева двоїста оцінка стала точною. Наприклад, після додавання до квадратичної задачі (18)–(20) функціонально надлишкового обмеження виду (13)

$$x_1x_1 + x_1x_4 + x_1x_5 - 2x_1 \leq 0,$$

яке одержано із п'ятого обмеження (21) домноженням на x_1 , оцінка ψ^* стає рівною 3.21576. Після додавання обмежень виду (14)

$$-x_1x_3 - x_1x_4 - x_1x_5 + 2x_1 + x_3 + x_4 + x_5 \leq 2$$

яке отримано домноженням четвертого обмеження з (21) на $(1 - x_1)$, оцінка ψ^* стає рівною 3.01571. А після додавання ще й обмежень виду (14)

$$-2x_1x_2 - x_2^2 - x_2x_3 - x_2x_5 + 2x_1 + 4x_2 + x_3 + x_5 \leq 3,$$

яке одержано домноженням першого обмеження з (21) на $(1 - x_2)$, оцінка ψ^* стає точною й дорівнює 3.0.

Висновки

Із прикладів квадратичних задач (15)–(17) та (18)–(20) для знаходження максимальних 2-плексів графів G_1 та G_2 видно, що функціонально надлишкові обмеження виду (13) та (14) відіграють значну роль для поліпшення точності лагранжевих двоїстих оцінок. Так, з їх допомогою лагранжеву двоїсту оцінку ψ^* можна зробити точною для $\rho_2(G_1)$ та $\rho_2(G_2)$. Однак, це не означає, що для будь-якої квадратичної задачі (1)–(3) за допомогою обмежень виду (13) та (14) можна домогтися того, щоб лагранжева двоїста оцінка була точною. Для підвищення точності лагранжевих двоїстих оцінок можна використовувати й інші способи побудови функціонально надлишкових обмежень [7].

Робота виконана при підтримці НАН України, проект 0117U000327.

Бібліографія

1. СТЕЦЮК П.І. Квадратична задача для максимального k -плекса в неорієнтованому графі / П.І. Стецюк П.І., Т.О. Бардадим, В.І. Ляшко // Журнал обчислювальної та прикладної математики. – 2017. – № 1 (124). – С. 71–78.
2. BALANSUNDARAM B. Clique Relaxations in Social Network Analysis: The Maximum k -plex Problem / B. Balansundaram, S. Butenko, I.V. Hicks // Operations Research. – 2011. – Vol. 59, no. 1. – P. 133–142.

3. GUO J. A more relaxed model for graph-based data clustering: s -plex cluster editing / J. Guo, C. Komusiewicz, R. Niedermeier, J. Uhlmann // SIAM J. Discrete Math. – 2010. – Vol. 24, No. 4. – P. 1662–1683.
4. McCLOSKEY B. Combinatorial algorithms for the maximum k -plex problem / B. McClosky, I.V. Hicks // J. of Combinatorial Optimization, 2012. – Vol. 23, No. 1. – P. 29–49.
5. PATTILLO J. On the maximum quasi-clique problem / J. Pattillo, A. Veremyev, S. Butenko, V. Boginski // Discrete Applied Mathematics, 2013. – Vol. 161, No. 1–2. – P. 244–257.
6. SEIDMAN S.B. A graph theoretic generalization of the clique concept / S.B. Seidman, B.L. Foster // J. of Math. Sociology. – 1978. – Vol. 6. – P. 139–154.
7. SHOR N.Z. Nondifferentiable Optimization and Polynomial Problems / N.Z. Shor. – London/Boston/Dordrecht: Kluwer Academic Publishers, 1998. – 394 p.
8. SHOR N.Z. Dual Solution of Quadratic-Type Problems by r -algorithm (subroutine DSQTP_r) / N.Z. Shor, P.I. Štetsyuk // Abstracts of the Second International Workshop "Recent Advances in Non-Differentiable Optimization" (October, 1-4, 2001, Kyiv, Ukraine). – Kyiv, 2001. – P. 36.

Надійшла 18.05.2017

Р А З Д Е Л VI

Специальные квадратичные задачи (часть II)

В раздел включены статьи:

- 6.1. Оптимальная нормированная структура спроса
и добавленной стоимости в продуктивной
модели Леонтьева (*Стецюк, Кошлай, 2010*) 400
- 6.2. Об одной экстремальной задаче для связи прямой и двойственной
моделей Леонтьева (*Стецюк, Кошлай, 2011*) 414
- 6.3. О спектральных свойствах матриц Леонтьева (*Стецюк, 2011*) . 423
- 6.4. Максимальное сингулярное число матрицы и его экономическая . .
интерпретация (*Стецюк, Эмменеггер, 2014*) 428
- 6.5. О глобальном минимуме целевой функции в задаче равновесной . .
упаковки кругов (*Стецюк, Романова, Шайтхауэр, 2014*) 438
- 6.6. Задача равновесной упаковки кругов (*Стецюк, Лиховид, 2015*) 444

Оптимальная нормированная структура спроса и добавленной стоимости в продуктивной модели Леонтьева

П.И. Стецюк, Л.Б. Кошлай

Кибернетика и систем. анализ. – 2010. – № 5. – С. 51–59.

Решена задача нахождения нормированных векторов спроса и добавленной стоимости, которые максимизируют доход в продуктивной модели Леонтьева. Показано, что если матрица Леонтьева продуктивна и неразложимая, то оптимальная нормированная структура спроса и добавленной стоимости определяется положительными компонентами собственных векторов, соответствующих максимальным собственным числам симметричных матриц. Приведены тестовые расчеты для семиотраслевой матрицы.

Введение

Структурная перестройка наряду с институциональными реформами является важнейшим направлением в процессе рыночной трансформации экономики. Удобным инструментарием для моделирования структурно-технологических изменений и определения пропорций между секторами экономики являются соотношения Леонтьева "затраты-выпуск" [1].

Для определения основных методов уменьшения производственных затрат в [2,3] предложена оптимизационная межотраслевая модель с переменными коэффициентами прямых затрат. Для уменьшения общего энергопотребления в [2] предложена линейная модель, которая также использует уравнения межотраслевого баланса. В [4, 5] применяется статическая двойственная модель Леонтьева для исследования связей между структурой труда и заработной платы, которые являются неотъемлемой частью структурных преобразований. Настоящая статья продолжает исследования в этом направлении. Здесь для статической модели Леонтьева исследуется соотношение между структурой спроса и добавленной стоимости, которые моделируются с помощью нормированных вектора спроса и вектора добавленной стоимости. Кроме того,

построено оптимальное соотношение между обеими структурами путем объединения прямой и двойственной моделей Леонтьева.

Структура статьи следующая. В разд. 1 рассмотрены основные балансовые соотношения в моделях Леонтьева. В разд. 2 приведены прямая и двойственная статические модели Леонтьева, проанализированы их основные свойства – продуктивность и прибыльность. Рассмотрен пример задачи линейного программирования на основе статической модели Леонтьева. В разделе 3 приведена задача нелинейного программирования, которая объединяет обе вышеуказанные модели Леонтьева и включает два квадратичных ограничения для нормировки вектора спроса и вектора добавленной стоимости. Доказано, что для продуктивной матрицы Леонтьева эта задача всегда имеет решение. Даны аналитические выражения для его нахождения через собственные векторы, которые соответствуют максимальным собственным числам некоторых симметричных матриц. Показано, что для продуктивной и неразложимой модели Леонтьева существует единственное решение, все компоненты которого положительны. Оно характеризует оптимальное соотношение между нормированными структурой спроса и структурой добавленной стоимости. Приведены теорема для нахождения этого оптимального соотношения и пример ее использования для семиотраслевой матрицы Леонтьева, построенной М.В. Михалевичем. В заключении обсуждаются возможные расширения нелинейной модели, рассмотренной в разд. 3.

1. Матрица Леонтьева и балансовые соотношения в моделях Леонтьева

Предположим, что в экономике существует n чистых отраслей (каждая отрасль производит один вид продукции и разные отрасли выпускают разные виды продукции). Пусть i, j – номера этих отраслей ($i, j = \overline{1, n}$). Обозначим a_{ij} величину прямых производственных затрат продукции отрасли i на изготовление единицы продукции отрасли j . Эта величина может быть выражена как в натуральном, так и в стоимостном выражении. Матрица $A = \{a_{ij}\}$ называется матрицей Леонтьева (матрицей коэффициентов прямых затрат или матрицей технологических коэффициентов). Матрица A несет информацию о сложившейся структуре межотраслевых связей, о существующей технологии общественного производства и т.д.

Пусть y_i и x_i – конечный и валовый продукты i -й отрасли соответственно. Эти величины связаны уравнением межотраслевого баланса:

$$x_i = \sum_{j=1}^n a_{ij} x_j + y_i, \quad i = \overline{1, n}. \quad (1)$$

В матричной форме система уравнений (1) имеет вид:

$$x = Ax + y \quad \text{или} \quad y = (I - A)x, \quad (2)$$

где $x = (x_1, \dots, x_n)^T$ – вектор валового продукта, $y = (y_1, \dots, y_n)^T$ – вектор конечного продукта, I – единичная матрица размерности $n \times n$. Здесь и далее T – символ транспонирования. Выражение Ax интерпретируется как затраты, в результате чего модель Леонтьева на основе (2) получила название затраты-выпуск.

Пусть p_i – цена единицы продукта отрасли i , c_i – добавленная стоимость (чистый доход от единицы выпуска) в отрасли i . Зависимость этих величин можно выразить через уравнение межотраслевого баланса для цен

$$p_i = \sum_{j=1}^n a_{ji} p_j + c_i, \quad i = \overline{1, n}.$$

В матричной форме система этих уравнений имеет вид

$$p = A^T p + c \quad \text{или} \quad c = (I - A^T)p, \quad (3)$$

где $p = (p_1, \dots, p_n)^T$ – вектор цен единичных продуктов отраслей, $c = (c_1, \dots, c_n)^T$ – вектор чистого дохода на единицу выпуска. Здесь выражение $A^T p$ интерпретируется как вектор суммы издержек на единицу выпуска.

Система (3 называется двойственной к системе (2). Обе системы связывает следующее соотношение:

$$p^T y = p^T (I - A)x = ((I - A^T)p)^T x = c^T x. \quad (4)$$

Соотношения (2), (3) и вытекающее из них равенство (4) положены в основу так называемых статических моделей Леонтьева.

2. Статические модели Леонтьева и их свойства

Статическая модель Леонтьева затраты–выпуск имеет форму

$$y = (I - A)x, \quad x \geq 0, \quad (5)$$

где A – известная матрица Леонтьева, y – известный вектор спроса, x – неизвестный вектор выпуска.

Наличие неотрицательного решения системы (5) при любом неотрицательном векторе спроса означает, что экономика согласно модели Леонтьева является продуктивной. Неотрицательная матрица A продуктивна, если существует обратная матрица $(I - A)^{-1}$, состоящая только из неотрицательных элементов. Пусть $\lambda_{max}(A)$ – максимальное собственное число матрицы A (число Фробениуса). Оно всегда положительно и не меньше модуля любого собственного числа матрицы A (см. теорему Фробениуса-Перрона для произвольных неотрицательных матриц [6]). Справедлива следующее утверждение.

Теорема 1 [6]. Матрица A продуктивна тогда и только тогда, когда $\lambda_{max}(A) < 1$.

Если матрица A – продуктивна, то при любом векторе $y \geq 0$ система (5) имеет решение, определяемое по формуле

$$x = (I - A)^{-1}y.$$

Матрицу $A^* = (I - A)^{-1}$ называют матрицей полных затрат. Ее можно представить в виде

$$A^* = I + A + A^2 + \dots$$

Теорема 1 означает необходимые и достаточные условия сходимости ряда $I + A + A^2 + \dots$ к матрице $A^* = (I - A)^{-1}$.

Статическая двойственная модель Леонтьева имеет вид

$$c = (I - A^T)p, \quad p \geq 0, \quad (6)$$

где A – известная матрица Леонтьева, c – известный вектор добавленной стоимости (чистый доход от единицы выпуска), p – неизвестный вектор цен. В основу модели (6) положена система (3).

Если система (6) при любом $c \geq 0$ имеет неотрицательное решение $p = (p_1, \dots, p_n)^T$, то двойственная модель Леонтьева называется прибыльной. Из теоремы 1 следует прибыльность модели (6). Свойство при-

быльности этой модели является двойственным к свойству продуктивности модели (5) в понимании, что выполнение одного из них обуславливает справедливость другого. Действительно, если неотрицательная матрица $A^* = (I - A)^{-1}$ существует, то отсюда следует существование неотрицательной матрицы $(I - A^T)^{-1} = ((I - A)^{-1})^T = (A^*)^T$.

Прямая (5) и двойственная (6) модели Леонтьева либо их аналоги в форме неравенств используются при решении ряда оптимизационных задач, применяемых при моделировании экономических процессов. Так, например, используя задачу линейного программирования (ЛП-задачу) можно определить при каком векторе выпуска x реализация конечного продукта y приведет к максимальному доходу с учетом наличного запаса $\nu = (\nu_1, \dots, \nu_m)^T$ первичных ресурсов. Получаем следующую ЛП-задачу:

$$\max_{x, y} p^T y \quad (7)$$

при ограничениях

$$x = Ax + y, \quad (8)$$

$$Bx \leq \nu, \quad (9)$$

$$x \geq 0, \quad (10)$$

где p – известный вектор цен, B – технологическая матрица размерности $m \times n$. Коэффициент b_{kj} задает количество k -го первичного ресурса ($k = 1, \dots, m$), необходимое для производства единицы продукции j -й отрасли ($j = 1, \dots, n$).

Центральными в модели (7)–(10) являются неравенства (9), определяющие количество наличного запаса первичных ресурсов. Именно эти неравенства ограничивают сверху целевую функцию (7). Однако ограничить ее можно и другим способом. Ниже рассмотрим одну из таких задач нелинейного программирования [7], которая базируется на объединении прямой (5) и двойственной (6) моделей Леонтьева с учетом связывающего их равенства $p^T y = c^T x$.

3. Экстремальная квадратичная задача с использованием обеих моделей Леонтьева и ее свойства

Пусть вектор $y \geq 0$ задает не конечный спрос, а лишь его структуру. Положим $\|y\| = 1$, где $\|\cdot\|$ – евклидова норма вектора. Аналогично и для

вектора добавленной стоимости $c \geq 0$ будем предполагать, что $\|c\| = 1$. Рассмотрим задачу нелинейного программирования в постановке

$$\max_{p,y} p^T y = \max_{x,c} c^T x \quad (11)$$

при ограничениях

$$y = (I - A)x, \quad x \geq 0, \quad y \geq 0, \quad (12)$$

$$c = (I - A^T)p, \quad p \geq 0, \quad c \geq 0, \quad (13)$$

$$\|y\| = 1, \quad \|c\| = 1, \quad (14)$$

где неизвестными являются компоненты векторов x, y, p, c . Задача (11)–(14) состоит в нахождении таких соотношений между структурой конечного спроса и структурой добавленной стоимости, чтобы максимума достигала такая величина, которая с точностью до множителя равнялась бы значению валового национального продукта.

Формально задача (11)–(14) является задачей нелинейного программирования. Нелинейны здесь целевая функция (11), выражающая связь между прямой и двойственной моделями Леонтьева, и ограничения (14), связанные с нормировками вектора конечного спроса и вектора добавленной стоимости. Если в ограничениях (14) вместо норм векторов y и c использовать квадраты этих норм, то задача (11)–(14) преобразуется в экстремальную квадратичную задачу с нелинейной целевой функцией и ограничениями, которые заданы линейными системами (12), (13) и двумя квадратичными равенствами: $\|y\|^2 = 1$ и $\|c\|^2 = 1$.

Если матрица A – продуктивна, то задача (11)–(14) всегда имеет решение: x^*, y^*, p^*, c^* . Справедлива следующая теорема.

Теорема 2. Пусть матрица Леонтьева A продуктивна и ей соответствует матрица полных затрат $A^* = (I - A)^{-1}$. Пусть $\hat{A} = (A^*)^T A^*$ и $\lambda_{max}(\hat{A}) = \lambda_{max}(\hat{A}^T)$ – максимальные собственные числа матриц \hat{A} и \hat{A}^T . Решением задачи (11)–(14) будут либо векторы

$$y^* = \xi(\hat{A}), \quad x^* = A^* y^*, \quad c^* = A^* y^* / \|A^* y^*\|, \quad p^* = (A^*)^T c^*, \quad (15)$$

где $\xi(\hat{A})$ – неотрицательный собственный вектор матрицы \hat{A} , соответствующий максимальному собственному числу $\lambda_{max}(\hat{A})$, либо векторы

$$c^* = \xi(\hat{A}^T), \quad p^* = (A^*)^T c^*, \quad y^* = (A^*)^T c^* / \|(A^*)^T c^*\|, \quad x^* = A^* y^*, \quad (16)$$

где $\xi(\hat{A}^T)$ – неотрицательный собственный вектор матрицы \hat{A}^T , соответствующий максимальному собственному числу $\lambda_{max}(\hat{A}^T)$. При этом

оптимальное значение целевой функции единственно и удовлетворяет следующим соотношениям:

$$(p^*)^T y^* = (c^*)^T x^* = (c^*)^T A^* y^* = \sqrt{\lambda_{max}(\hat{A})} = \sqrt{\lambda_{max}(\hat{A}^T)}. \quad (17)$$

Доказательство. Для продуктивной матрицы Леонтьева A (ей соответствует $\lambda_{max}(A) < 1$) имеем неотрицательную матрицу полных затрат $A^* = (I - A)^{-1}$, вследствие чего компоненты вектора $x = A^*y$ будут неотрицательными при любых $y \geq 0$, а компоненты вектора $p = (A^*)^T c$ будут неотрицательными при любых $c \geq 0$. С учетом этого задачу (11)–(14) для продуктивной матрицы A можно переписать в виде

$$\max_{y, c} c^T A^* y \quad (18)$$

при ограничениях

$$\sum_{i=1}^n y_i^2 = 1, \quad y \geq 0, \quad (19)$$

$$\sum_{i=1}^n c_i^2 = 1, \quad c \geq 0. \quad (20)$$

Здесь ограничения (14) переформулированы в виде двух квадратичных равенств. Решением задачи (18)–(20) есть либо решение (i), либо решение (ii):

(i) векторы $y^* = \xi(\hat{A})$ и $c^* = A^* y^* / \|A^* y^*\|$, где $\xi(\hat{A})$ – неотрицательный собственный вектор матрицы \hat{A} , соответствующий максимальному собственному числу $\lambda_{max}(\hat{A})$;

(ii) векторы $c^* = \xi(\hat{A}^T)$ и $y^* = (A^*)^T c^* / \|(A^*)^T c^*\|$, где $\xi(\hat{A}^T)$ – неотрицательный собственный вектор матрицы \hat{A}^T , соответствующий максимальному собственному числу $\lambda_{max}(\hat{A}^T)$.

Докажем сначала решение (i). Если в задаче (18)–(20) опустить требования на неотрицательность компонент векторов y и p , то тогда ее аналог можно записать в виде

$$\max_{c, y} c^T A^* y = \max_{\|y\|=1} \varphi(y), \quad (21)$$

где $\varphi(y)$ – решение подзадачи при фиксированном y

$$\varphi(y) = \max_c c^T A^* y \quad (22)$$

при ограничении

$$\sum_{i=1}^n c_i^2 = 1. \quad (23)$$

Пусть u – множитель Лагранжа, соответствующий ограничению (23). Функция Лагранжа для задачи (22), (23) имеет следующий вид:

$$L(c, u) = c^T A^* y + u(1 - \sum_{i=1}^n c_i^2).$$

Из условия $\frac{\partial L(c, u)}{\partial c} = 0$ находим

$$A^* y - 2uc(u) = 0,$$

откуда имеем

$$c(u) = \frac{1}{2u} A^* y. \quad (24)$$

В результате получаем

$$\psi(u) = L(c(u), u) = \frac{1}{4u} \|A^* y\|^2 + u.$$

Функция $\psi(u)$ при $u^* = \frac{1}{2} \|A^* y\|$ достигает минимума $\psi^* = \psi(u^*) = \|A^* y\|$. Из (24) определяем вектор:

$$c^*(y) = c(u^*) = A^* y / \|A^* y\|, \quad (25)$$

при котором $\varphi(y) = \|A^* y\|$. Следовательно,

$$\max_{\|y\|=1} \varphi(y) = \max_{\|y\|=1} \|A^* y\| = \max_{\|y\|=1} \sqrt{y^T (A^*)^T A^* y} = \max_{\|y\|=1} \sqrt{y^T \hat{A} y}. \quad (26)$$

С учетом того, что

$$\lambda_{max}(\hat{A}) = \max_{\|y\|=1} y^T \hat{A} y,$$

решением задачи (26) будет собственный вектор матрицы \hat{A} , который соответствует ее максимальному собственному числу $\lambda_{max}(\hat{A})$, т.е. $y^* = \xi(\hat{A})$. Положительность $\lambda_{max}(\hat{A})$ вытекает из определения матрицы \hat{A} . Следовательно, решением задачи (21)–(23) будут вектор $y^* = \xi(\hat{A})$ и вектор $c^* = A^* y^* / \|A^* y^*\|$, который вычислен согласно (25). Это завершает доказательство пункта (i).

Поскольку матрица \hat{A} – неотрицательная, то согласно теореме Фробениуса-Перрона для произвольных неотрицательных матриц [6] существует неотрицательный собственный вектор $\xi(\hat{A})$, соответствующий числу $\lambda_{max}(\hat{A})$. Это означает неотрицательность вектора y^* . Неотрицательность вектора $c^* = A^*y^*/\|A^*y^*\|$ следует из неотрицательности матрицы A^* и вектора y^* . Это дает соотношения для векторов y^* и c^* в (15). Соотношения для неотрицательных векторов x^* и p^* в (15) следуют из неотрицательности матрицы A^* и формул расчета для них $x^* = A^*y^*$ и $p^* = A^*c^*$, что завершает вывод всех соотношений в (15) на основе пункта (i).

Аналогично можно обосновать и формулы (16) на основе условия (ii). Здесь соответствующая задача имеет вид:

$$\max_{c, y} c^T A^* y = \max_{\|c\|=1} \varphi(c),$$

где $\varphi(c)$ – решение следующей подзадачи при фиксированном c :

$$\varphi(c) = \max_y c^T A^* y \quad (27)$$

при ограничении

$$\sum_{i=1}^n y_i^2 = 1. \quad (28)$$

Для задачи (27)–(28) имеем решение

$$y^*(c) = (A^*)^T c / \|(A^*)^T c\|, \quad (29)$$

при котором $\varphi(c) = \|(A^*)^T c\|$. Следовательно,

$$\max_{\|c\|=1} \varphi(c) = \max_{\|c\|=1} \sqrt{c^T A^* (A^*)^T c} = \max_{\|c\|=1} \sqrt{c^T \hat{A}^T c}. \quad (30)$$

Для задачи (30) решением c^* будет собственный вектор $\xi(\hat{A}^T)$ матрицы \hat{A}^T , соответствующий ее максимальному собственному числу $\lambda_{max}(\hat{A}^T)$. Учитывая, что матрица \hat{A}^T – неотрицательная, то для числа $\lambda_{max}(\hat{A}^T)$ такой неотрицательный собственный вектор существует согласно теореме Фробениуса-Перрона для произвольных неотрицательных матриц. Это гарантирует неотрицательность вектора c^* . Неотрицательность вектора $y^* = (A^*)^T c^* / \|(A^*)^T c^*\|$, полученного согласно (29), очевидна в силу неотрицательности матрицы \hat{A}^T и вектора $c^* = \xi(\hat{A}^T)$. Это задает соотношения для векторов y^* и c^* в (16). Соотношения для векторов x^*

и p^* в (16) следуют из неотрицательности матрицы A^* и формул расчета для них $x^* = A^*y^*$ и $p^* = A^*c^*$, что завершает вывод всех соотношений в (16) на основе пункта (ii).

Выполнение равенств в (17) следует из цепочки следующих равенств:

$$(p^*)^T y^* = ((A^*)^T c^*)^T y^* = (c^*)^T A^* y^*, \quad (c^*)^T x^* = (c^*)^T A^* y^*,$$

$$(c^*)^T A^* y^* = \max_{\|y\|=1} \varphi(y) = \max_{\|y\|=1} \sqrt{y^T \hat{A} y} = \sqrt{\lambda_{max}(\hat{A})},$$

$$(c^*)^T A^* y^* = \max_{\|c\|=1} \varphi(c) = \max_{\|c\|=1} \sqrt{c^T \hat{A}^T c} = \sqrt{\lambda_{max}(\hat{A}^T)}.$$

Теорема доказана.

Из теоремы 2 вытекают два следствия.

Следствие 1. Если матрица Леонтьева A продуктивна и максимальное собственное число матрицы \hat{A} – единственное, то задача (11)–(14) имеет единственное решение x^* , y^* , p^* и c^* , которое может быть вычислено либо согласно (15), либо согласно (16).

Приведенное далее следствие связано со свойством неразложимости неотрицательных $(n \times n)$ -матриц. Матрица A называется неразложимой, если одновременной перестановкой строк и столбцов ее нельзя привести к виду

$$A = \left\{ \begin{array}{cc} A_1 & A_2 \\ 0 & A_3 \end{array} \right\},$$

где A_1 и A_3 – квадратные подматрицы размеров $k \times k$ и $(n-k) \times (n-k)$. Свойство неразложимости матрицы A гарантирует неразложимость симметричной матрицы \hat{A} . Это означает, что ее нельзя представить в виде

$$\hat{A} = \left\{ \begin{array}{cc} \hat{A}_1 & 0 \\ 0 & \hat{A}_3 \end{array} \right\},$$

где \hat{A}_1 и \hat{A}_3 – квадратные подматрицы размеров $k \times k$ и $(n-k) \times (n-k)$.

Из теоремы Фробениуса-Перрона для неотрицательных и неразложимых матриц [6] следует, что числу $\lambda_{max}(\hat{A})$ соответствует единственный собственный вектор матрицы \hat{A} , все компоненты которого положительны. Поэтому если в качестве собственного вектора матрицы \hat{A} использовать именно этот вектор, то все остальные векторы из (15) также будут положительными в силу того, что в матрицах A^* и $(A^*)^T$ отсутствуют нулевые строки. Аналогичная ситуация имеет место и для матрицы \hat{A}^T , т.е. ее максимальному собственному числу соответствует

единственный собственный вектор, все компоненты которого положительны. Справедливо следующее утверждение.

Следствие 2. Если матрица Леонтьева A продуктивна и неразложима, то задаче (11)–(14) соответствует единственное решение, все компоненты которого положительны. Это решение имеет вид

$$y^* = \xi(\hat{A}), \quad x^* = A^* y^*, \quad c^* = \xi(\hat{A}^T), \quad p^* = (A^*)^T c^*,$$

где $\xi(\hat{A})$ и $\xi(\hat{A}^T)$ – положительные собственные векторы матриц \hat{A} и \hat{A}^T , соответствующие их максимальным собственным числам $\lambda_{max}(\hat{A})$ и $\lambda_{max}(\hat{A}^T)$.

Назовем это решение оптимальным соотношением между структурой спроса и структурой добавленной стоимости для матрицы Леонтьева A . Здесь оптимальную структуру спроса определяют компоненты собственного вектора $\xi(\hat{A})$, а оптимальную структуру добавленной стоимости – компоненты собственного вектора $\xi(\hat{A}^T)$. Оптимальное значение целевой функции в задаче (11)–(14) равно $\sqrt{\lambda_{max}(\hat{A})}$, или, что то же самое, $\sqrt{\lambda_{max}(\hat{A}^T)}$.

Однако наличие единственного решения в следствии 2 еще не означает, что это решение легко найти. Как правило, в случае высокой кратности максимального собственного числа нахождение такого вектора проблематично. Ниже приведена теорема, позволяющая упростить нахождение такого оптимального соотношения.

Теорема 3. Если матрица Леонтьева A продуктивна и неразложима, и значение $\lambda_{max}(\hat{A})$ – единственное, то тогда задача (11)–(14) имеет единственное решение, все компоненты которого положительны. Это решение определяется следующим образом:

$$y^* = \xi(\hat{A}), \quad c^* = A^* y^* / \|A^* y^*\|, \quad x^* = \sqrt{\lambda_{max}(\hat{A})} c^*, \quad p^* = \sqrt{\lambda_{max}(\hat{A})} y^*, \quad (31)$$

где $\xi(\hat{A})$ – неотрицательный собственный вектор матрицы \hat{A} , соответствующий максимальному собственному числу $\lambda_{max}(\hat{A})$. Решению (31) отвечает оптимальное значение целевой функции, равное $\sqrt{\lambda_{max}(\hat{A})}$.

На основе теоремы 3 можно построить алгоритм, который с помощью известных процедур нахождения собственных чисел и собственных векторов симметричных матриц позволяет определить оптимальное соотношение между структурой спроса и структурой добавленной стоимости. Реализация такого алгоритма предполагает то, что если единственному максимальному собственному числу соответствует собствен-

ный вектор, у которого все компоненты отрицательны, то следует взять их с обратным знаком.

Ниже приведем пример расчета оптимального соотношения между структурой спроса и структурой добавленной стоимости для семиотраслевой матрицы Леонтьева

$$A = \begin{pmatrix} 0.337 & 0.139 & 0.215 & 0.127 & 0.146 & 0.112 & 0.1960 \\ 0.023 & 0.251 & 0.179 & 0.089 & 0.019 & 0.131 & 0.0050 \\ 0.163 & 0.176 & 0.191 & 0.097 & 0.103 & 0.095 & 0.0870 \\ 0.012 & 0.009 & 0.157 & 0.031 & 0.029 & 0.026 & 0.0940 \\ 0.009 & 0.010 & 0.008 & 0.226 & 0.107 & 0.006 & 0.0071 \\ 0.153 & 0.121 & 0.099 & 0.031 & 0.025 & 0.019 & 0.0330 \\ 0.161 & 0.193 & 0.103 & 0.101 & 0.095 & 0.087 & 0.0910 \end{pmatrix}, \quad (32)$$

построенной М.В. Михалевичем на основе межотраслевого баланса Украины (2007) для моделирования структурно-технологических изменений в энергетических отраслях.

Матрица A является продуктивной ($\lambda_{max}(A) = 0.75374$) и неразложимой (не содержит нулевых компонент). Ей соответствует матрица

$$\hat{A} = \begin{pmatrix} 5.5705 & 3.5710 & 4.1980 & 2.5066 & 1.9989 & 2.3038 & 2.6338 \\ 3.5710 & 4.7218 & 3.8463 & 2.1869 & 1.6037 & 2.1307 & 2.2138 \\ 4.1980 & 3.8463 & 5.2490 & 2.5991 & 1.9358 & 2.2573 & 2.4627 \\ 2.5066 & 2.1869 & 2.5991 & 2.5403 & 1.4014 & 1.3074 & 1.5987 \\ 1.9989 & 1.6037 & 1.9358 & 1.4014 & 2.0990 & 0.9927 & 1.2061 \\ 2.3039 & 2.1307 & 2.2573 & 1.3074 & 0.9927 & 2.1507 & 1.3295 \\ 2.6338 & 2.2138 & 2.4627 & 1.5987 & 1.2061 & 1.3295 & 2.5240 \end{pmatrix},$$

для которой максимальное собственное число $\lambda_{max}(\hat{A})$ равно 18.105 и является единственным.

Согласно теореме 3 задача (11)–(14) с матрицей A имеет единственное решение, все компоненты которого положительны. Таким решением являются вектор

$$y^* = (0.5017, 0.4451, 0.4965, 0.3001, 0.2325, 0.2660, 0.2980)^T$$

для оптимальной нормированной структуры спроса, и вектор

$$c^* = (0.6258, 0.3469, 0.4599, 0.2061, 0.1325, 0.2731, 0.3766)^T$$

для оптимальной нормированной структуры добавленной стоимости. Соответствующие этой оптимальной структуре компоненты векторов x^* и p^* вычисляются умножением на $4.255 = \sqrt{18.105}$ (оптимальное значение целевой функции) компонент векторов c^* и y^* .

Направления дальнейших исследований

Дальнейшим развитием исследований может быть добавление к модели (11)–(14) ограничений в виде линейных неравенств на переменные x и p . Так, например, если добавить к модели неравенства (9), то для переменных y ограничения приобретают вид $BA^*y \leq \nu$. Подобные линейные неравенства могут определять требования к неизвестным компонентам вектора p , которые для переменных s будут выражены с помощью матрицы $(A^*)^T$. При наличии таких ограничений собственные векторы будут решением только тогда, когда отсутствуют активные ресурсные ограничения вида (9) либо их „ценовые“ аналоги для переменных p . Иначе это приводит к семейству оптимизационных задач, которые можно интерпретировать как задачи нахождения некоторых собственных векторов симметричных матриц с неявно заданными ограничениями (требованиями) на компоненты собственных векторов.

К такому же типу оптимизационных задач приводит и следующее обобщение задачи (11)–(14). Например, необходимо найти оптимальную структуру векторов y и s при дополнительном требовании, чтобы спрос на продукцию отрасли j был в два раза больше, чем спрос на продукцию отрасли i . Аналогичные условия могут характеризовать и соотношение добавленной стоимости в заданных отраслях. Дополнение задачи (11)–(14) подобными требованиями позволит более глубоко исследовать связи между отдельными секторами экономики.

Литература

1. ЛЕОНТЬЕВ В.В. Избранные произведения. Т. 1–3. – М.: "Экономика", 2006 – 2008.
2. МИХАЛЕВИЧ М.В., СЕРГИЕНКО И.В. Моделирование переходной экономики. Модели, методы, информационные технологии. – Киев: Наук. думка, 2005. – 670 с.
3. СЕРГИЕНКО И.В., МИХАЛЕВИЧ М.В., СТЕЦЮК П.И., КОШЛАЙ Л.Б. Модели и информационные технологии для поддержки принятия решений при проведении структурно-технологических преобразований // Кибернетика и систем. анализ. – 2009. – № 2. – С. 26–49.

4. БОРТИС Г. Институции, поведение и экономическая теория. Вклад в классико-кейнсианскую политическую экономию. – Киев: Изд. дом “Києво-Могилянська академія”, 2009. – 598 с.
5. BORTIS H. Keynes and the Classics: Notes on the Monetary Theory of Production, in: Modern Theories of Money. The Nature and Role of Money in Capitalist Economies, Rochon L.-P. and Sergio Rossi (eds). – Edward Elgar: UK, USA, 2003. – P. 411–474.
6. АШМАНОВ С.А. Введение в математическую экономику. – М.: Наука, 1984. – 296 с.
7. СТЕЦЮК П.И., КОШЛАЙ Л.Б., ПИЛИПОВСКИЙ А.В. О задаче оптимального соотношения между спросом и добавленной стоимостью в моделях Леонтьева // Теорія оптимальних рішень. – 2010. – № 9. – С. 136-143.

Поступила 01.06.2010

Об одной экстремальной задаче для связи прямой и двойственной моделей Леонтьева

П.И. Стецюк, Л.Б. Кошлай

*Спектральные и эволюционные задачи. – 2011. – Т. 21, № 2. –
С. 164–169.*

Рассматривается квадратичная экстремальная задача, объединяющая прямую и двойственную модели Леонтьева. Показано, что если матрица Леонтьева продуктивна и неразложима, то задача имеет единственное решение. Разработан алгоритм нахождения этого решения, даны тестовые расчеты для 7-отраслевой технологической матрицы.

Введение

Предложенный Леонтьевым алгебраический метод "затраты-выпуск" (в зарубежной литературе input – output) позволяет одновременно учитывать затраты на производство продукции в заданной номенклатуре и распределение продукции в той же номенклатуре. Он сводится к системе линейных уравнений, параметрами которых являются коэффициенты затрат на производство продукции [1].

Матрица Леонтьева (другие ее названия – матрица прямых затрат или технологическая матрица) есть неотрицательная $n \times n$ -матрица $A = \{a_{ij}\}$. Коэффициент a_{ij} задает величину прямых производственных затрат продукции отрасли i на изготовление единицы продукции отрасли j . Здесь i, j – номера отраслей ($i, j = \overline{1, n}$) и считается, что отрасли производят различные виды продукции.

Пусть A – известная матрица Леонтьева, I – единичная $n \times n$ -матрица. Прямой моделью Леонтьева будем называть статическую модель „затраты–выпуск“

$$y = (I - A)x, \quad x \geq 0, \quad (1)$$

где y – известный вектор спроса, x – неизвестный вектор выпуска. Двойственной моделью Леонтьева будем называть статическую ценовую модель

$$c = (I - A^T)p, \quad p \geq 0, \quad (2)$$

где c – известный вектор добавленной стоимости (чистый доход от единицы выпуска), p – неизвестный вектор цен.

В работе покажем, что если прямую и двойственную модели Леонтьева объединить в единую математическую модель, то такую модель можно использовать для исследования взаимосвязей между структурами спроса и добавленной стоимости экономической системы.

1. Экстремальная квадратичная задача для прямой и двойственной моделей Леонтьева

Условимся считать, что вектор $y \geq 0$ ($\|y\| = 1$) задает нормированную структуру конечного спроса в прямой модели Леонтьева (1), а вектор $c \geq 0$ ($\|c\| = 1$) задает нормированную структуру добавленной стоимости в двойственной модели Леонтьева (2). Здесь $\|\cdot\|$ – евклидова норма вектора.

Рассмотрим следующую задачу нелинейного программирования

$$f^* = \max_{y \in R^n, p \in R^n} p^T y \quad \equiv \quad \max_{x \in R^n, c \in R^n} c^T x \quad (3)$$

при ограничениях

$$y = (I - A)x, \quad x \geq 0, \quad y \geq 0, \quad (4)$$

$$c = (I - A^T)p, \quad p \geq 0, \quad c \geq 0, \quad (5)$$

$$\|y\|^2 = 1, \quad \|c\|^2 = 1. \quad (6)$$

где матрица A – известная матрица Леонтьева, а неизвестными являются компоненты n -мерных векторов x , y , p , c .

Объясним содержательный смысл целевой функции (3). Здесь условие эквивалентности „ \equiv “ означает использование либо целевой функции $p^T y$, либо $c^T x$. Это следует из справедливости соотношения

$$p^T y = p^T (I - A)x = ((I - A^T)p)^T x = c^T x, \quad (7)$$

которое связывает равенства из ограничений (4) для прямой модели Леонтьева с равенствами из ограничений (5) для двойственной модели Леонтьева. Заметим, что соотношение (7) означает, что валовой национальный продукт совпадает с национальным доходом.

Следовательно, в задаче (3)–(6) требуется найти такие нормированные векторы конечного спроса (вектор y^*) и добавленной стоимости (вектор c^*), чтобы с точностью до некоторого постоянного множителя максимума достигал валовой национальный продукт (он же равен национальному доходу). Оптимальным значениям векторов y^* и c^* будут соответствовать такие значения векторов x^* и p^* , которые с точностью до постоянных множителей будут определять оптимальный выпуск в прямой модели Леонтьева и оптимальные цены в двойственной модели Леонтьева.

В общем случае нахождение векторов y^* и c^* , x^* и p^* требует привлечения численных методов оптимизации, так как это связано с решением квадратичной экстремальной задачи (3)–(6). Здесь целевая функция (3) есть билинейной функцией либо от переменных p и y , либо от переменных c и x , а ограничения (6) содержат два квадратичных равенства. Однако оказывается, что в важных для экономических приложений случаях задача (3)–(6) может быть решена аналитически в терминах собственных чисел и собственных векторов симметричных матриц.

2. Аналитическое решение задачи (3)–(6) для продуктивной матрицы Леонтьева

Матрица Леонтьева A продуктивна тогда и только тогда, когда $\lambda_{max}(A) < 1$ (теорема 1.5, [2]). Здесь $\lambda_{max}(A)$ – максимальное собственное число матрицы A (число Фробениуса). Для произвольной неотрицательной матрицы A число Фробениуса всегда положительно и не меньше модуля любого собственного числа матрицы A (теорема 1.2, [2]). Продуктивность матрицы Леонтьева A эквивалентна тому, что существует матрица $B = (I - A)^{-1}$, состоящая только из неотрицательных элементов. Матрицу B называют матрицей полных затрат.

Если матрица A – продуктивна, то прямая модель Леонтьева называется продуктивной, что означает наличие решения у системы (1) при любом неотрицательном векторе спроса. Это решение определяется через матрицу B по формуле

$$x = (I - A)^{-1}y = By, \quad \text{где } x \geq 0 \text{ для } \forall y \geq 0. \quad (8)$$

Если матрица A – продуктивна, то двойственная модель Леонтьева называется прибыльной, что означает наличие решения у системы (2) при любом неотрицательном векторе добавленной стоимости. Это решение

определяется через матрицу B по формуле

$$p = (I - A^T)^{-1}c = B^T c, \quad \text{где } p \geq 0 \text{ для } \forall c \geq 0. \quad (9)$$

Свойство прибыльности модели (2) следует из свойства продуктивности модели (1). Действительно, если $B = (I - A)^{-1}$ – неотрицательная матрица, то неотрицательной будет и матрица $B^T = ((I - A)^{-1})^T = (I - A^T)^{-1}$.

С учетом соотношений (8) и (9) задачу (3)–(6) для продуктивной матрицы A можно переписать с помощью неотрицательной матрицы B в следующем виде

$$f^* = (c^*)^T B y^* = \max_{y \geq 0, c \geq 0} c^T B y \quad (10)$$

при ограничениях

$$\sum_{i=1}^n y_i^2 = 1, \quad \sum_{i=1}^n c_i^2 = 1. \quad (11)$$

Пусть $\xi(B^T B)$ – неотрицательный нормированный собственный вектор матрицы $B^T B$, соответствующий ее максимальному собственному числу $\lambda_{max}(B^T B)$, а $\xi(BB^T)$ – неотрицательный нормированный собственный вектор матрицы BB^T , соответствующий ее максимальному собственному числу $\lambda_{max}(BB^T)$. Поскольку матрицы $B^T B$ и BB^T – неотрицательные, то векторы $\xi(B^T B)$ и $\xi(BB^T)$ (не обязательно единственные) существуют согласно теореме Фробениуса-Перрона для произвольных неотрицательных матриц (теорема 1.2, [2]).

В [3] показано, что решением задачи (10)–(11) есть либо векторы $y^* = \xi(B^T B)$ и $c^* = B y^* / \|B y^*\|$, либо векторы $c^* = \xi(BB^T)$ и $y^* = B^T c^* / \|B^T c^*\|$. При этом векторы y^* и c^* не обязательно определяются единственным образом, но оптимальное значение целевой функции единственно и удовлетворяет соотношению:

$$f^* = (c^*)^T B y^* = \sqrt{\lambda_{max}(B^T B)} = \sqrt{\lambda_{max}(BB^T)}. \quad (12)$$

Если матрица A – продуктивна, то задача (3)–(6) всегда имеет решение (x^*, y^*, p^*, c^*) . Здесь векторы y^* и c^* определяются как оптимальное решение задачи (10)–(11), а векторы x^* и p^* получаются в результате подстановки y^* и c^* в соотношения (8) и (9) соответственно. Аналитический способ решения задачи (3)–(6) обеспечивает следующее утверждение.

Теорема 1. Пусть матрица Леонтьева A продуктивна и ей соответствует матрица полных затрат $B = (I - A)^{-1}$. Решением задачи (3)–(6) будут как векторы

$$y^* = \xi(B^T B), \quad x^* = B y^*, \quad c^* = B y^* / \|B y^*\|, \quad p^* = B^T c^*, \quad (13)$$

где $\xi(B^T B)$ – неотрицательный нормированный собственный вектор матрицы $B^T B$, соответствующий ее максимальному собственному числу $\lambda_{\max}(B^T B)$, так и векторы

$$c^* = \xi(B B^T), \quad p^* = B^T c^*, \quad y^* = B^T c^* / \|B^T c^*\|, \quad x^* = B y^*, \quad (14)$$

где $\xi(B B^T)$ – неотрицательный нормированный собственный вектор матрицы $B B^T$, соответствующий ее максимальному собственному числу $\lambda_{\max}(B B^T)$.

При этом оптимальное значение целевой функции единственно и удовлетворяет следующим соотношениям:

$$f^* = (p^*)^T y^* = (c^*)^T x^* = \sqrt{\lambda_{\max}(B^T B)} = \sqrt{\lambda_{\max}(B B^T)}. \quad (15)$$

Из теоремы 1 вытекает следствие о единственности решения задачи (3)–(6).

Следствие 1. Если матрица Леонтьева A продуктивна и матрице $B^T B$ соответствует единственное максимальное собственное число $\lambda_{\max}(B^T B)$, то задача (3)–(6) имеет единственное решение x^* , y^* , p^* и c^* , которое может быть вычислено либо согласно (13), либо согласно (14).

Следствие 1 позволяет найти единственное решение задачи (3)–(6) с помощью компьютерных программ нахождения собственных чисел и собственных векторов симметричных матриц. Для этого достаточно лишь учесть, что если единственному максимальному собственному числу соответствует собственный вектор, у которого все компоненты неположительны, то следует взять их с обратным знаком.

3. Единственность решения задачи (3)–(6) для продуктивной и неразложимой матрицы Леонтьева

Матрица A называется неразложимой, если одновременной перестановкой строк и столбцов ее нельзя привести к виду

$$A = \left\{ \begin{array}{cc} A_1 & A_2 \\ 0 & A_3 \end{array} \right\},$$

где A_1 и A_3 – квадратные подматрицы размеров $k \times k$ и $(n-k) \times (n-k)$.

Если матрица A – продуктивна и неразложима, то ей соответствует неразложимая матрица полных затрат B . Действительно, если $\lambda_{max}(A) < 1$, то неотрицательная матрица B представима в виде

$$B = (I - A)^{-1} = I + A + A^2 + \dots = \sum_{i=0}^{\infty} A^i.$$

Отсюда легко видеть, что если матрица A неразложима, то $I + A$ также неразложима, чего достаточно для неразложимости матрицы B .

Из неразложимости матрицы B следует неразложимость симметричных матриц $B^T B$ и $B B^T$. Это означает, что одновременной перестановкой строк и столбцов их нельзя привести к виду

$$\hat{B} = \left\{ \begin{array}{cc} \hat{B}_1 & 0 \\ 0 & \hat{B}_2 \end{array} \right\},$$

где \hat{B}_1 и \hat{B}_2 – квадратные подматрицы размеров $k \times k$ и $(n-k) \times (n-k)$.

Из теоремы Фробениуса-Перрона для неотрицательных и неразложимых матриц (теорема 1.1, [2]) следует, что числу $\lambda_{max}(B^T B)$ соответствует единственный нормированный собственный вектор матрицы $B^T B$, все компоненты которого положительны. Поэтому если в качестве собственного вектора матрицы $B^T B$ использовать именно этот вектор, то все остальные векторы из (13) также будут положительными в силу того, что в матрицах B и B^T отсутствуют нулевые строки. Аналогичная ситуация имеет место и для матрицы $B B^T$, т.е. ее максимальному собственному числу соответствует единственный нормированный собственный вектор, все компоненты которого положительны. Справедливо следующее утверждение.

Теорема 2. Пусть матрица Леонтьева A продуктивна и неразложима, ей соответствует матрица полных затрат $B = (I - A)^{-1}$. Задача (3)–(6) имеет единственное решение, все компоненты которого положительны. Это решение имеет вид

$$y^* = \xi(B^T B), \quad x^* = B y^*, \quad c^* = \xi(B B^T), \quad p^* = B^T c^*,$$

где $\xi(B^T B)$ и $\xi(B B^T)$ – положительные нормированные собственные векторы матриц $B^T B$ и $B B^T$, соответствующие их максимальным собственным числам $\lambda_{max}(B^T B)$ и $\lambda_{max}(B B^T)$.

Вектор y^* из теоремы 2 назовем оптимальной структурой спроса, а вектор c^* – оптимальной структурой добавленной стоимости для продуктивной и неразложимой матрицы Леонтьева A . Первую определяют положительные компоненты собственного вектора $\xi(B^T B)$, а вторую – компоненты собственного вектора $\xi(B B^T)$.

4. Численный пример для семиотраслевой матрицы Леонтьева

Рассмотрим матрицу

$$A = \begin{pmatrix} 0.337 & 0.139 & 0.215 & 0.127 & 0.146 & 0.112 & 0.1960 \\ 0.023 & 0.251 & 0.179 & 0.089 & 0.019 & 0.131 & 0.0050 \\ 0.163 & 0.176 & 0.191 & 0.097 & 0.103 & 0.095 & 0.0870 \\ 0.012 & 0.009 & 0.157 & 0.031 & 0.029 & 0.026 & 0.0940 \\ 0.009 & 0.010 & 0.008 & 0.226 & 0.107 & 0.006 & 0.0071 \\ 0.153 & 0.121 & 0.099 & 0.031 & 0.025 & 0.019 & 0.0330 \\ 0.161 & 0.193 & 0.103 & 0.101 & 0.095 & 0.087 & 0.0910 \end{pmatrix},$$

построенную М.В. Михалевичем на основе межотраслевого баланса Украины 2007 года для моделирования структурно-технологических преобразований в энергетических отраслях [4].

Матрица A является продуктивной (число Фробениуса $\lambda_{max}(A) = 0.75374$ меньше единицы) и неразложимой (не содержит нулевых ком-

понтент). Ей соответствует матрица полных затрат

$$B = \begin{pmatrix} 2.1022 & 0.9089 & 1.0487 & 0.6874 & 0.5911 & 0.5428 & 0.6541 \\ 0.3319 & 1.6425 & 0.5823 & 0.3285 & 0.1962 & 0.3399 & 0.1842 \\ 0.6504 & 0.7036 & 1.7547 & 0.4682 & 0.3899 & 0.3864 & 0.3775 \\ 0.2072 & 0.2201 & 0.3777 & 1.1778 & 0.1467 & 0.1404 & 0.2101 \\ 0.0910 & 0.0979 & 0.1362 & 0.3180 & 1.1727 & 0.0591 & 0.0774 \\ 0.4633 & 0.4472 & 0.4483 & 0.2549 & 0.2020 & 1.2026 & 0.2169 \\ 0.5934 & 0.6669 & 0.6073 & 0.4330 & 0.3487 & 0.3490 & 1.3500 \end{pmatrix}.$$

Симметричные матрицы $B^T B$ и $B B^T$ имеют следующий вид:

$$B^T B = \begin{pmatrix} 5.5705 & 3.5710 & 4.1980 & 2.5066 & 1.9989 & 2.3039 & 2.6338 \\ 3.5710 & 4.7218 & 3.8463 & 2.1869 & 1.6037 & 2.1307 & 2.2138 \\ 4.1980 & 3.8463 & 5.2490 & 2.5991 & 1.9358 & 2.2573 & 2.4627 \\ 2.5066 & 2.1869 & 2.5991 & 2.5403 & 1.4014 & 1.3074 & 1.5987 \\ 1.9989 & 1.6037 & 1.9358 & 1.4014 & 2.0990 & 0.9927 & 1.2061 \\ 2.3039 & 2.1307 & 2.2573 & 1.3074 & 0.9927 & 2.1507 & 1.3294 \\ 2.6338 & 2.2138 & 2.4627 & 1.5987 & 1.2061 & 1.3294 & 2.5240 \end{pmatrix}$$

и

$$B B^T = \begin{pmatrix} 7.8896 & 3.4481 & 4.8559 & 2.1417 & 1.4177 & 2.9397 & 4.0668 \\ 3.4481 & 3.4428 & 2.8244 & 1.1523 & 0.6391 & 1.7214 & 2.2239 \\ 4.8559 & 2.8244 & 4.6600 & 1.6945 & 1.0253 & 2.1472 & 2.9041 \\ 2.1417 & 1.1523 & 1.6945 & 1.7066 & 0.6630 & 0.9079 & 1.3929 \\ 1.4177 & 0.6391 & 1.0253 & 0.6630 & 1.5223 & 0.5528 & 0.8738 \\ 2.9397 & 1.7214 & 2.1472 & 0.9079 & 0.5528 & 2.2147 & 1.7387 \\ 4.0668 & 2.2239 & 2.9041 & 1.3929 & 0.8738 & 1.7387 & 3.4192 \end{pmatrix}.$$

Матрице $B^T B$ соответствует единственное максимальное собственное число $\lambda_{max}(B^T B)$, равное 18.10503.

Согласно следствию 1 задача (3)–(6) с матрицей A имеет единственное решение, а согласно теореме 2 все компоненты этого решения положительны. Вектор $\xi(B^T B)$ определяет оптимальную структуру спроса

$$y^* = (0.5017, 0.4451, 0.4965, 0.3001, 0.2325, 0.2660, 0.2980)^T,$$

а вектор $\xi(B B^T)$ – оптимальную структуру добавленной стоимости

$$c^* = (0.6258, 0.3469, 0.4599, 0.2061, 0.1325, 0.2731, 0.3766)^T.$$

Выводы

На основе прямой и двойственной статических моделей Леонтьева построена квадратичная экстремальная задача нахождения нормированных структур конечного спроса и добавленной стоимости с целью максимизации национального дохода. Показано, что для продуктивной матрицы Леонтьева задача всегда имеет решение и даны аналитические выражения для его нахождения через собственные векторы, соответствующие максимальным собственным числам некоторых симметричных матриц. Для продуктивной и неразложимой модели Леонтьева показано существование единственного решения, все компоненты которого положительны. Полученные результаты могут быть использованы для исследования связи между структурами спроса и добавленной стоимости в межотраслевых балансовых моделях.

Литература

1. ЛЕОНТЬЕВ В.В. Избранные произведения: в 3 т. – М.: Экономика. – 2006 – 2007.
2. АШМАНОВ С.А. Введение в математическую экономику – М.: Наука. – 1984.
3. СТЕЦЮК П.И., КОШЛАЙ Л.Б. Оптимальная нормированная структура спроса и добавленной стоимости в продуктивной модели Леонтьева // Кибернетика и систем. анализ. – 2010. – № 5. – С. 51–59.
4. СЕРГИЕНКО И.В., МИХАЛЕВИЧ М.В., СТЕЦЮК П.И., КОШЛАЙ Л.Б. Модели и информационные технологии для поддержки принятия решений при проведении структурно-технологических преобразований // Кибернетика и систем. анализ. – 2009. – № 2. – С. 26–49.

О спектральных свойствах матриц Леонтьева

П.И. Стецюк

Статистика. Моделирование. Оптимизация: сборник трудов Всероссийской конференции (Челябинск, 28 ноября – 3 декабря 2011 г.). – Челябинск: Издательский центр ЮУрГУ, 2011. – С. 173–178.

В заметке описано назначение octave-функций `frobenius` и `step2010` и приведены их коды на языке Octave [1]. Вычислены спектральные характеристики матриц прямых и полных затрат (матриц Леонтьева A и $B = (I - A)^{-1}$, соответственно) для 22-отраслевого баланса России за 2001–2003 годы [2]. Обе octave-функции могут применяться для вычислений в Матлабе, так как используют функцию `eig(·)`¹, октавовский вариант которой совпадает с матлабовским.

1. Octave-функция `frobenius(A)` вычисляет число Фробениуса λ_A (равно максимальному собственному числу матрицы A) и соответствующий числу λ_A вектор Фробениуса x_A , такой что

$$Ax_A = \lambda_A x_A \quad \text{и} \quad \sum_{i=1}^n (x_A)_i = 1. \quad (1)$$

Вектор x_A называют правым вектором Фробениуса. Вычисление λ_A и x_A для матрицы A реализует следующий код на языке Octave.

```
function [lambda_A x_A] = frobenius(A)
[N M] = size(A);
if N ~= M
    error('matrix A is not square: N= %i M= %i.', N, M);
    stop
end
[V D] = eig(A);
[lambda_A, ind] = max(ones(1,M)*D);
x_A = V(:,ind)/sum(V(:,ind));
endfunction
```

¹Функция `[R, D] = eig(A)` вычисляет диагональную матрицу D собственных значений и матрицу R правых собственных векторов, удовлетворяющих соотношению $A * R = R * D$. Эти векторы нормированы так, что норма каждого из них равна единице.

Вызов octave-функции $[\lambda_A \ x_A] = \text{frobenius}(A')$ соответствует вычислению числа λ_A и левого вектора Фробениуса p_A , такого что

$$A^T p_A = \lambda_A p_A \quad \text{и} \quad \sum_{i=1}^n (p_A)_i = 1. \quad (2)$$

Если матрица A – продуктивна ($\lambda_A < 1$) [3], то векторы x_A и p_A есть правым и левым векторами Фробениуса для матрицы $B = (I - A)^{-1}$, т.е.

$$B x_A = \lambda_B x_A, \quad B^T p_A = \lambda_B p_A, \quad \text{где} \quad \lambda_B = \frac{1}{1 - \lambda_A}.$$

Ниже приведены числа λ_A , λ_B (первая строка) и векторы x_A (первая колонка), p_A (вторая колонка) для 22-отраслевых матриц прямых затрат (таблицы 8.1) из системы таблиц „затраты-выпуск“ России за 2001-2003 годы [2].

	2003		2002		2001	
	0.52430	2.10216	0.51903	2.07913	0.51922	2.07995
1	0.08677	0.03672	0.08652	0.03540	0.07306	0.03325
2	0.09156	0.03025	0.10261	0.03153	0.11968	0.03254
3	0.01662	0.06154	0.01861	0.05221	0.01537	0.04220
4	0.00013	0.02310	0.00015	0.02131	0.00012	0.02217
5	0.07252	0.05720	0.05527	0.05901	0.06010	0.06179
6	0.08793	0.05321	0.11136	0.07786	0.08260	0.05777
7	0.05800	0.06666	0.05701	0.06491	0.06390	0.06476
8	0.08982	0.06553	0.08589	0.06544	0.08665	0.06563
9	0.02733	0.05635	0.02501	0.05104	0.02615	0.05161
10	0.01575	0.05004	0.01568	0.04752	0.01624	0.04752
11	0.02200	0.09134	0.02135	0.08795	0.02788	0.10528
12	0.01975	0.05907	0.01894	0.05703	0.02122	0.06075
13	0.01163	0.05864	0.01135	0.06539	0.01207	0.05800
14	0.02758	0.04210	0.02808	0.04064	0.03306	0.04070
15	0.02350	0.03088	0.02253	0.03322	0.02639	0.03397
16	0.09712	0.03299	0.09600	0.02897	0.09988	0.02749
17	0.18485	0.02012	0.17905	0.01764	0.17616	0.01799
18	0.01333	0.03209	0.01167	0.02976	0.00895	0.03119
19	0.01301	0.03162	0.01523	0.03238	0.01481	0.03313
20	0.00134	0.02870	0.00107	0.02763	0.00073	0.03387
21	0.02147	0.04249	0.02051	0.04181	0.02195	0.04296
22	0.01797	0.02935	0.01613	0.03135	0.01302	0.03542

Здесь по номерам идут следующие продукты: (1) Электро- и теплоэнергия; (2) Продукты нефтегазовой промышленности; (3) Уголь; (4) Горючие сланцы и торф; (5) Черные металлы; (6) Цветные металлы; (7) Продукты химической и нефтехимической промышленности; (8) Машины и оборудование, продукты металлообработки; (9) Продукты лесной, деревообрабатывающей и целлюлозно-бумажной промышленности; (10) Строительные материалы (включая продукты стекольной и фарфоро-фаянсовой промышленности); (11) Продукты легкой промышленности; (12) Продукты пищевой промышленности; (13) Прочие промышленные продукты; (14) Продукция строительства; (15) Сельхозпродукты, услуги по обслуживанию сельского хозяйства и продукты лесного хозяйства; (16) Услуги транспорта и связи; (17) Торгово-посреднические услуги (включая услуги общественного питания); (18) Продукты прочих видов деятельности; (19) Услуги жилищно-коммунального хозяйства и непроизводственных видов бытового обслуживания населения; (20) Услуги здравоохранения, физической культуры и социального обеспечения, образования, культуры и искусства; (21) Услуги науки и научного обслуживания, геологии и разведки недр, геодезической и гидрометеорологической служб; (22) Услуги финансового посредничества, страхования, управления и общественных объединений.

2. Octave-функция `step2010(B)` вычисляет σ_B – максимальное сингулярное число неотрицательной матрицы B (арифметическое значение квадратного корня максимального собственного числа матрицы $B^T B$ или, что то же самое, матрицы BB^T) и два неотрицательных вектора y_B и c_B таких, что

$$B^T B y_B = \sigma_B^2 y_B \text{ и } \sum_{i=1}^n (y_B)_i = 1, \quad BB^T c_B = \sigma_B^2 c_B \text{ и } \sum_{i=1}^n (c_B)_i = 1. \quad (3)$$

Вычисление числа σ_B , векторов y_B и c_B для матрицы B (может быть и прямоугольной) реализует следующий код на языке Octave.

```
function [sigma_B y_B c_B] = step2010(B)
[N M] = size(B); [V D] = eig(B'*B);
[lambda, ind] = max(ones(1,M)*D);
y_B = V(:,ind)/sum(V(:,ind));
[V D] = eig(B*B');
[lambda, ind] = max(ones(1,N)*D);
c_B = V(:,ind)/sum(V(:,ind));
sigma_B=sqrt(lambda);
endfunction
```

Если B – обратная матрица Леонтьева (матрица полных затрат) и кратность σ_B равна единице, то векторы y_B и c_B реализуют оптимальные нормированные структуры конечного выпуска $y^* = y_B / \|y_B\|$ и добавленной стоимости $c^* = c_B / \|c_B\|$ [4]. Это означает, что

$$\sigma_B = (c^*)^T B y^* = f_B^* = \max_{c \in R^n, y \in R^n} c^T B y, \quad \text{где } \|y\| = \|c\| = 1, \quad (4)$$

и никакие другие нормированные векторы y и c не могут давать значение больше, чем f_B^* .

Если матрица A – продуктивна и неразложима (каждая отрасль использует, хотя бы косвенно, продукцию всех отраслей), то векторы y_B и c_B определяются однозначно и их компоненты положительны. Это имеет место для 22-отраслевых матриц прямых затрат A за 2001–2003 годы и означает, что для соответствующих им матриц полных затрат $B = (I - A)^{-1}$ сингулярное число σ_B имеет кратность равную единице, а векторы y_B и c_B определяются единственным образом. Результаты этих расчетов даны ниже: σ_B – в первой строке, а векторы y_B и c_B – соответственно в первой и второй колонках.

	2003		2002		2001	
	2.43968		2.38831		2.38698	
1	0.06006	0.07425	0.06040	0.07412	0.05505	0.06558
2	0.05001	0.06567	0.05593	0.07255	0.06252	0.08190
3	0.04698	0.03376	0.04133	0.03154	0.03442	0.02579
4	0.02196	0.01031	0.01954	0.00934	0.01776	0.00846
5	0.06278	0.06449	0.06010	0.05683	0.06351	0.06071
6	0.05872	0.06720	0.07487	0.08319	0.05924	0.06476
7	0.06886	0.06685	0.06763	0.06593	0.06860	0.06993
8	0.06867	0.07590	0.06683	0.07371	0.06755	0.07466
9	0.05115	0.04319	0.04800	0.04090	0.04844	0.04124
10	0.04689	0.03217	0.04484	0.03135	0.04519	0.03122
11	0.05634	0.04771	0.05540	0.04730	0.06263	0.05518
12	0.04457	0.03455	0.04325	0.03414	0.04594	0.03734
13	0.04418	0.02773	0.04644	0.02910	0.04271	0.02820
14	0.03922	0.03187	0.03818	0.03147	0.03977	0.03408
15	0.03165	0.03371	0.03329	0.03496	0.03516	0.03839
16	0.05208	0.07150	0.04993	0.06932	0.05010	0.07064
17	0.05992	0.12189	0.05778	0.11630	0.05694	0.11265
18	0.02376	0.01734	0.02106	0.01530	0.02169	0.01433
19	0.03137	0.02170	0.03325	0.02426	0.03347	0.02434
20	0.02139	0.01051	0.02096	0.01030	0.02499	0.01193

21	0.03343	0.02688	0.03386	0.02761	0.03496	0.02855
22	0.02601	0.02080	0.02713	0.02049	0.02937	0.02013

Значение σ_B для нормированных векторов y^* и c^* (см. (4)) не менее чем на 10% больше, чем значение $f_B = c_A^T B y_A$, которое реализуется на нормированных векторах $y_A = x_A / \|x_A\|$ и $c_A = p_A / \|p_A\|$. Так, например, для 2003 года имеем $\sigma_B = 2.4397$, $f_B = 2.1274$, для 2002 года – $\sigma_B = 2.3883$, $f_B = 2.0768$ и для 2001 года – $\sigma_B = 2.3870$, $f_B = 2.0301$.

Работа выполнена при частичной финансовой поддержке SNSF (Швейцария), проект №IZ73ZO_127962 „Analysis of Institutional and Technological Changes in Market and Transition Economies on the Background of the Present Financial Crisis“

Литература

1. Octave [Электронный ресурс]: <http://www.octave.org/>. – Режим доступа: свободный.
2. Статистическая информация [Электронный ресурс]: Система таблиц „затраты-выпуск“ России за 2001-2003 годы. Федеральная служба государственной статистики (Росстат). http://www.gks.ru/wps/wcm/connect/rosstat/rosstatsite/main/publishing/catalog/statisticCollections/doc_1135086739625 – Режим доступа: свободный.
3. АШМАНОВ, С.А. Введение в математическую экономику / С.А.Ашманов. – М.: Наука, 1984. – 296 с.
4. СТЕЦЮК, П.И. Оптимальная нормированная структура спроса и добавленной стоимости в продуктивной модели Леонтьева / П.И.Стецюк, Л.Б.Кошлай // Кибернетика и системный анализ, №5, 2010. – С. 51–59.

Максимальное сингулярное число матрицы и его экономическая интерпретация

П.И. Стецюк, Ж.Ф. Эмменеггер

Кибернетика и системный анализ. – 2014. – № 3. – С. 51–57.

Показано, что максимальное сингулярное число матрицы и соответствующие ему сингулярные векторы являются оптимальным решением специальной квадратичной экстремальной задачи. Рассматривается экономическая интерпретация оптимального решения для линейной модели производства и продуктивной модели Леонтьева, показана связь оптимального решения с числом и векторами Фробениуса. Приведено сравнение чисел Фробениуса и максимальных сингулярных чисел для обратной матрицы Леонтьева в 15-ти отраслевом балансе Украины за 2003–2009 годы.

Введение

Устойчивость экономических, транспортных, энергетических и других систем, как правило, характеризуется величиной максимального собственного числа некоторой матрицы. Компоненты собственных векторов, отвечающих этому числу, связаны со значениями параметров системы в состоянии устойчивого равновесия. Для экономических систем эту роль выполняют число и векторы Фробениуса квадратных матриц с неотрицательными коэффициентами [1, 2].

Технологические процессы часто описываются с помощью прямоугольных $m \times n$ -матриц A , в которых количество столбцов n (технологии) и количество строк m (выпуск) не совпадают. Для таких процессов число и векторы Фробениуса напрямую не применимы, так как не имеют аналогов для прямоугольных матриц. Оказывается, что в этом случае они связаны с максимальным сингулярным числом прямоугольной матрицы A , квадрат которого можно интерпретировать как число Фробениуса для квадратной матрицы, которая получена умножением прямоугольной матрицы на транспонированную. Роль векторов Фробениуса здесь выполняют левый и правый сингулярные векторы – собственные векторы $m \times m$ -матрицы AA^T и $n \times n$ -матрицы $A^T A$, которые соответствуют квадрату максимального сингулярного числа.

Свойства максимального сингулярного числа σ_A и его роль в экономических моделях является предметом исследования в настоящей статье. В первом разделе рассмотрим число σ_A и его свойства для вещественных матриц с использованием квадратичной экстремальной задачи, в которой оптимальное значение целевой функции равно σ_A [3]. Во втором разделе для неотрицательных матриц дадим экономическую интерпретацию числа σ_A и отвечающих ему сингулярных векторов. Приведем сравнение числа σ_A с числом Фробениуса для обратной матрицы Леонтьева в агрегированном 15-отраслевом балансе Украины за 2003–2009 годы.

1. Число σ_A и квадратичная экстремальная задача

Пусть A – вещественная $m \times n$ -матрица и $\sigma_A > 0$ – ее максимальное сингулярное число, $\sigma_A = \sqrt{\lambda_{max}(A_1)} = \sqrt{\lambda_{max}(A_2)}$, где $\lambda_{max}(A_1)$ и $\lambda_{max}(A_2)$ – максимальные собственные числа симметричных $m \times m$ -матрицы $A_1 = AA^T$ и $n \times n$ -матрицы $A_2 = A^T A$, соответственно. Левый и правый сингулярные векторы числа σ_A равны собственным векторам матриц A_1 и A_2 , соответствующим максимальному собственному числу $\lambda_{max}(A_1) = \lambda_{max}(A_2) = \sigma_A^2$

Лемма 1. Число σ_A равно оптимальному значению целевой функции в квадратичной экстремальной задаче

$$\sigma_A = (u^*)^T A x^* = \max_{x \in R^n, u \in R^m} u^T A x \quad (1)$$

при ограничениях

$$\sum_{i=1}^m u_i^2 = 1, \quad \sum_{i=1}^n x_i^2 = 1. \quad (2)$$

Оптимальным решением (u^*, x^*) в задаче (1),(2) являются либо векторы

$$u^* = \xi(A_1), \quad x^* = A^T u^* / \|A^T u^*\|, \quad (3)$$

либо векторы

$$x^* = \xi(A_2), \quad u^* = A x^* / \|A x^*\|, \quad (4)$$

где $\xi(A_1)$ и $\xi(A_2)$ – собственные векторы матриц A_1 и A_2 , соответствующие максимальным собственным числам $\lambda_{max}(A_1) = \lambda_{max}(A_2)$.

Доказательство. Задачу (1),(2) можно записать в виде

$$\sigma_A = \max_{\|u\|=\|x\|=1} u^T A x = \max_{\|u\|=1} \varphi(u),$$

где $\varphi(u)$ – решение подзадачи:

$$\varphi(u) = \max_x u^T A x \quad (5)$$

при ограничении

$$\sum_{i=1}^n x_i^2 = 1. \quad (6)$$

Пусть v – множитель Лагранжа, соответствующий ограничению (6). Функция Лагранжа для позадачи (5)–(6) имеет вид:

$$L(x, v) = u^T A x + v(1 - \sum_{i=1}^n x_i^2).$$

Из условия $\frac{\partial L(x, v)}{\partial x} = 0$ находим $A^T u - 2vx(v) = 0$, откуда имеем

$$x(v) = \frac{1}{2v} A^T u. \quad (7)$$

В результате получаем

$$\psi(v) = L(x(v), v) = \frac{1}{4v} \|A^T u\|^2 + v.$$

Функция $\psi(v)$ при $v^* = \frac{1}{2} \|A^T u\|$ достигает минимума $\psi^* = \psi(v^*) = \|A^T u\|$. Из (7) определяем оптимальное решение подзадачи (5),(6):

$$x^*(u) = x(v^*) = A^T u / \|A^T u\|, \quad (8)$$

на основании которого $\varphi(u) = \|A^T u\|$. Следовательно,

$$\max_{\|u\|=1} \varphi(u) = \max_{\|u\|=1} \|A^T u\| = \max_{\|u\|=1} \sqrt{u^T A A^T u} = \max_{\|u\|=1} \sqrt{u^T A_1 u}. \quad (9)$$

С учетом того, что $\lambda_{max}(A_1) = \max_{\|u\|=1} u^T A_1 u$, получаем $\max_{\|u\|=1} \varphi(u) = \sqrt{\lambda_{max}(A_1)} = \sigma_A$, что доказывает равенство оптимального значения целевой функции в задаче (1),(2) и числа σ_A .

Решением задачи (9) будет собственный вектор матрицы A_1 , который соответствует ее максимальному собственному числу $\lambda_{max}(A_1)$.

Отсюда с учетом (8) имеем $u^* = \xi(A_1)$ и $x^* = A^T u^* / \|A^T u^*\|$, что доказывает для задачи (1),(2) оптимальность решения (u^*, x^*) по формуле (3).

Аналогично можно обосновать и формулу (4). Здесь соответствующая задача имеет вид:

$$\sigma_A = \max_{\|x\|=\|u\|=1} u^T A x = \max_{\|x\|=1} \varphi(x),$$

где $\varphi(x)$ – решение подзадачи

$$\varphi(x) = \max_u u^T A x \quad (10)$$

при ограничении

$$\sum_{i=1}^m u_i^2 = 1. \quad (11)$$

Для подзадачи (10),(11) имеем решение

$$u^*(x) = A x / \|A x\|, \quad (12)$$

при котором $\varphi(x) = \|A x\|$. Следовательно,

$$\max_{\|x\|=1} \varphi(x) = \max_{\|x\|=1} \sqrt{x^T A^T A x} = \max_{\|x\|=1} \sqrt{x^T A_2^T x} = \sqrt{\lambda_{max}(A_2)} = \sigma_A. \quad (13)$$

Для задачи (13) оптимальным решением x^* будет $\xi(A_2)$ – собственный вектор матрицы A_2 , соответствующий ее максимальному собственному числу $\lambda_{max}(A_2)$. С учетом (12) имеем $x^* = \xi(A_2)$ и $u^* = A x^* / \|A x^*\|$, что доказывает оптимальность решения (u^*, x^*) по формуле (4) для задачи (1),(2). Это завершает доказательство леммы.

Если кратность числа σ_A больше единицы, то число σ_A – единственное оптимальное значение целевой функции, но ему соответствует бесконечно много оптимальных решений в задаче (1)–(2). Лемма 1 отражает взаимосвязь между входящими в оптимальное решение векторами из множеств собственных векторов матриц A_1 и A_2 , отвечающих их максимальным собственным числам $\lambda_{max}(A_1) = \lambda_{max}(A_2)$. С помощью леммы 1 можно найти одно (но произвольное) оптимальное решение из множества оптимальных решений, которое зависит от выбора векторов в формулах (3) или (4).

Если используется формула (3), то оптимальное решение (u^*, x^*) состоит из вектора u^* – некоторого произвольного собственного вектора матрицы A_1 (им может быть как любой вектор из базисной системы

собственных векторов, так и их произвольная нормированная линейная комбинация с неотрицательными коэффициентами) и вектора x^* – собственного вектора матрицы A_2 , но не произвольного, а вычисленного по формуле (3). По существу, для собственного вектора матрицы $A_1 = AA^T$ выполняется равенство

$$\lambda_{max}(A_1)\xi(A_1) = \lambda_{max}(A_1)AA^T\xi(A_1). \quad (14)$$

Если это равенство умножить справа на матрицу A^T , то получаем

$$\lambda_{max}(A_1)A^T\xi(A_1) = \lambda_{max}(A_1)A^T A (A^T\xi(A_1)). \quad (15)$$

Из (15) легко увидеть, что вектор $\xi(A_2) = A^T\xi(A_1)/\|A^T\xi(A_1)\|$ является собственным вектором матрицы $A_2 = A^T A$, так как при $\lambda_{max}(A_1) = \lambda_{max}(A_2)$ для него выполняется равенство

$$\lambda_{max}(A_2)\xi(A_2) = \lambda_{max}(A_2)A^T A\xi(A_2) \quad (16)$$

Если для вычисления сингулярного вектора использовать формулу (4), то получим ситуацию аналогичную, но при этом выбирается любой вектор из базисной системы собственных векторов для матрицы A_2 или произвольная их нормированная линейная комбинация с неотрицательными коэффициентами. Здесь оптимальное решение (u^* , x^*) состоит из вектора x^* – некоторого произвольного собственного вектора матрицы A_2 и вектора u^* – собственного вектора матрицы A_1 , но не произвольного, а вычисленного по формуле (4). Действительно, если равенство (16) умножить справа на матрицу A , то получим равенство

$$\lambda_{max}(A_2)A\xi(A_2) = \lambda_{max}(A_2)AA^T (A\xi(A_2)). \quad (17)$$

Учитывая, что $\lambda_{max}(A_2) = \lambda_{max}(A_1)$, из (17) следует справедливость равенства (14) для вектора $\xi(A_1) = A\xi(A_2)/\|A\xi(A_2)\|$. Следовательно, вектор $\xi(A_1) = A\xi(A_2)/\|A\xi(A_2)\|$ является собственным вектором матрицы A_1 , соответствующим ее максимальному собственному числу $\lambda_{max}(A_1)$.

Из вышеизложенного вытекает справедливость следующей леммы.

Лемма 2. Если кратность σ_A равна единице, то числу σ_A в задаче (1),(2) соответствует единственное оптимальное решение (u^* , x^*), компоненты которого

$$u^* = \xi(A_1) \quad \text{и} \quad x^* = \xi(A_2) \quad (18)$$

равны собственным векторам матриц A_1 и A_2 , отвечающим их максимальным собственным числам $\lambda_{max}(A_1) = \lambda_{max}(A_2)$.

Этот случай имеет место при анализе экономических моделей, где коэффициенты матрицы являются неотрицательными. Аналогичная ситуация и для случая, когда кратность числа σ_A больше единицы, если коэффициенты в матрице A возмутить так, чтобы возмущенные матрицы AA^T и $A^T A$ имели единственное максимальное собственное число.

2. Экономическая интерпретация

В экономических приложениях оперируют с неотрицательными и неразложимыми матрицами. Неотрицательной называют матрицу, все компоненты которой неотрицательны. Квадратную $n \times n$ -матрицу A называют неразложимой, если одновременной перестановкой строк и столбцов ее нельзя привести к виду

$$A = \left\{ \begin{array}{cc} \mathcal{A}_1 & \mathcal{A}_2 \\ 0 & \mathcal{A}_3 \end{array} \right\},$$

где \mathcal{A}_1 – квадратная матрица. В макроэкономике неразложимость матрицы A означает, что каждая отрасль использует (хотя бы косвенно) продукты всех других отраслей.

Число Фробениуса λ_A равно максимальному собственному числу $n \times n$ -матрицы A с неотрицательными коэффициентами. Правый вектор Фробениуса равен вектору x_A , который удовлетворяет условиям

$$Ax_A = \lambda_A x_A \quad \text{и} \quad \sum_{i=1}^n (x_A)_i = 1. \quad (19)$$

Левый вектор Фробениуса равен вектору p_A , такому что

$$A^T p_A = \lambda_A p_A \quad \text{и} \quad \sum_{i=1}^n (p_A)_i = 1. \quad (19')$$

В соотношениях (19) и (19') векторы Фробениуса отмасштабированы таким образом, чтобы сумма их компонент равнялась единице. Это связано с удобством интерпретации каждой из компонент этих векторов, например, как доли (в процентном соотношении) вклада той или иной технологии в конечный выпуск продукции. Для симметричной матрицы правый и левый векторы Фробениуса совпадают. Такой вектор назовем вектором Фробениуса. Нормированный вектор Фробениуса является собственным вектором симметричной матрицы A , соответствующим ее максимальному собственному числу λ_A .

Для неотрицательных и неразложимых матриц теорема Фробениуса-Перрона [1, теорема 1.1, с. 29] гарантирует существование правого (левого) вектора Фробениуса, все компоненты которого положительны. Из изложенного выше следует справедливость такой леммы.

Лемма 3. Если неотрицательная $m \times n$ -матрица A не содержит нулевых строк и столбцов, а минимальная по размеру из матриц $A_1 = AA^T$ и $A_2 = A^T A$ является неразложимой, то число σ_A в задаче (1)–(2) достигается в единственной точке (u^*, x^*) , все компоненты которой положительны. При этом вектор u^* равен нормированному вектору Фробениуса для матрицы A_1 , а вектор x^* равен нормированному вектору Фробениуса для матрицы A_2 .

Содержательный смысл оптимального решения задачи (1),(2) рассмотрим на примере линейной системы $y = Ax$ со входом $x \in R^n$ и выходом $y \in R^m$, которая описывает процесс производства m видов продукции n предприятиями (технологиями). Здесь A – неотрицательная технологическая $m \times n$ -матрица, в которой коэффициент a_{ij} означает количество i -го продукта произведенного по j -ой технологии, $i = 1, \dots, m$, $j = 1, \dots, n$. Величина $a_{ij} > 0$, если i -й продукт производится по j -ой технологии и $a_{ij} = 0$, если не производится.

Задачу (1),(2) для указанного процесса можно записать в таком виде: найти

$$\sigma_A = (u^*)^T y^* = \max_{y \in R^m, u \in R^m} u^T y \quad (20)$$

при ограничениях

$$y = Ax, \quad x \in R^n, \quad (21)$$

$$\sum_{i=1}^m u_i^2 = 1, \quad \sum_{i=1}^n x_i^2 = 1. \quad (22)$$

Задача (20)–(22) является более прозрачной для интерпретации содержательного смысла максимального сингулярного числа σ_A и соответствующих ему сингулярных векторов, чем задача (1),(2). Легко видеть, что если компоненты вектора $u \in R^m$ интерпретировать как цены на единицу производимых продуктов, то получаем оптимальное соотношение между нормированным вектором использования технологий и нормированным вектором цен. Оптимальное соотношение реализует максимум суммарной цены на производимые продукты, и этот максимум равен максимальному сингулярному числу матрицы A . Если выполнены условия леммы 3, то оптимальное соотношение однозначно определяет оптимальные цены, равные положительному нормированному вектору

Фробениуса матрицы $A_1 = AA^T$, и оптимальное использование технологий, равное положительному нормированному вектору Фробениуса матрицы $A_2 = A^T A$.

К интересной двойственной интерпретации числа σ_A и векторов Фробениуса легко придти, если задачу (1),(2) записать в таком виде: найти

$$\sigma_A = (p^*)^T x^* = \max_{p \in R^n, x \in R^n} p^T x \quad (23)$$

при ограничениях

$$p = A^T u, \quad u \in R^m, \quad (24)$$

$$\sum_{i=1}^m u_i^2 = 1, \quad \sum_{i=1}^n x_i^2 = 1. \quad (25)$$

Задача (23)–(25) записана с помощью матрицы A^T и является не менее прозрачной, чем задача (20)–(22). Здесь вектор p определяет стоимости использования технологий, где цена на производство единицы каждого отдельного продукта одинакова. Легко видеть, что оптимальное значение целевой функции (23) отражает максимальную стоимость использования технологий и реализуется на таких же нормированном векторе использования технологий и нормированном векторе цен, как и в задаче (20)–(22).

Указанная интерпретация оптимального решения задачи (1)–(2) справедлива и если $n = m$, т.е. количество технологий равно количеству продуктов. Для квадратных матриц леммы 1 и 3 имеют больше приложений, так как в качестве неизвестных в задаче (1),(2) можно рассматривать не только используемые технологии, но и выпуск продукции. Проиллюстрируем это на примере модели Леонтьева $y = (I - A)x$, где x – валовый выпуск, y – конечный выпуск, A – матрица коэффициентов прямых затрат, I – единичная матрица. Здесь матрица Леонтьева $(I - A)$ не является неотрицательной, но для продуктивной модели Леонтьева (ей соответствует $\lambda_A < 1$), неотрицательной является обратная матрица Леонтьева $B = (I - A)^{-1}$. Отсюда имеем, что $x = By$, и леммы 1 и 3 можно задействовать для конечного выпуска y и добавленной стоимости w , которая связана с двойственной (ценовой) моделью Леонтьева $w = (I - A^T)p$, где w – вектор норм добавленной стоимости, а p – вектор цен.

Для продуктивной модели Леонтьева квадратичная экстремальная задача имеет вид: найти

$$\sigma_B = (w^*)^T B y^* = \max_{v \in R^n, y \in R^n} w^T B y \quad (26)$$

Таблица 1. Сравнение чисел λ_B и σ_B (Украина, 15 отраслей)

Год	λ_A	λ_B	σ_B	$\frac{(\sigma_B - \lambda_B)}{\lambda_B}$
2003	0.58641	2.41787	2.914	0.205
2004	0.58476	2.40825	2.937	0.220
2005	0.59611	2.47591	3.107	0.255
2006	0.58495	2.40936	2.980	0.237
2007	0.57231	2.33812	2.865	0.225
2008	0.56623	2.30535	2.884	0.251
2009	0.56958	2.32332	2.866	0.234

при ограничениях

$$\sum_{i=1}^m w_i^2 = 1, \quad \sum_{i=1}^n y_i^2 = 1. \quad (27)$$

Если переобозначить $y \rightarrow x$, $w \rightarrow u$, то задача (26),(27) совпадает с задачей (1),(2) и для нее применимы леммы 1 и 3.

В результате получим нормированные векторы конечного продукта и добавленной стоимости, которые отвечают максимизации числа σ_B – величине пропорциональной национальному доходу [4]. При этом число σ_B всегда будет лучше, чем число Фробениуса λ_B . На основании таблицы 1 приведем сравнение чисел σ_B и λ_B для агрегированного 15-ти отраслевого баланса Украины за 2003–2009 годы [5]. Здесь λ_A – числа Фробениуса технологических матриц A : $\lambda_B = 1/(1 - \lambda_A)$. Из последнего столбца таблицы видим, что число σ_B превышает число Фробениуса λ_B не меньше, чем на 20%.

Заключение

В настоящей статье показано, что сингулярное число матрицы и сингулярные векторы для экономических моделей позволяют построить некоторые оптимальные соотношения, которые связывают ценовой и ресурсный факторы. Сингулярные векторы тесно связаны с векторами Фробениуса, они расширяют область действия последних на класс неотрицательных прямоугольных матриц.

Однако интерпретация оптимальных соотношений связана с нормированными векторами, которые не играют большой роли в экономи-

ческих приложениях. Чтобы заменить нормированные векторы на линейные выпуклые комбинации компонент, существует большой резерв по модификации квадратичной экстремальной задачи (1),(2). Так, например, масштабирование переменных u и x позволяет рассматривать векторы, которые легко приблизить к выпуклым комбинациям. Такие задачи незначительно сложнее, чем задача (1),(2). Их оптимальные решения будут определяться сингулярным числом некоторой модифицированной технологической матрицы, которая получена умножением технологической матрицы слева и справа на диагональные матрицы.

Литература

1. АШМАНОВ С.А. Введение в математическую экономику. – М.: Наука, 1984. – 296 с.
2. ПОНОМАРЕНКО О.І., ПЕРЕСТЮК М.О., БУРИМ В. М. Сучасний економічний аналіз. Ч. 2. Макроекономіка. – Київ: Вища школа, 2004. – 208 с.
3. СТЕЦЮК П.И. Квадратичная задача для максимального сингулярного числа // Праці Міжнародної наукової конференції "Питання оптимізації обчислень"(ПОО-XL), присвяченої 90-річчю від дня народження академіка В.М. Глушкова, Україна, Крим, Велика Ялта, смт. Кацивелі, 30 вересня – 4 жовтня 2013 р. – Київ: Ін-т кібернетики ім. В.М. Глушкова НАН України, 2013. – С. 255.
4. СТЕЦЮК П.И., КОШЛАЙ Л.Б. Оптимальная нормированная структура спроса и добавленной стоимости в продуктивной модели Леонтьева // Кибернетика и системный анализ. – 2010. – № 5. – С. 51–59.
5. СТЕЦЮК П.И., БОНДАРЕНКО А.В. О спектральных свойствах модели Леонтьева // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2011. – № 10. – С. 84–90.

Поступила 03.01.2014

О глобальном минимуме целевой функции в задаче равновесной упаковки кругов

П.И. Стецюк, Т.Е. Романова, Г. Шайтхауэр
Доповіді НАН України. – 2014. – № 6. – С. 53–57.

Рассматривается задача равновесной упаковки семейства кругов в круге минимального радиуса в виде многоэкстремальной задачи нелинейного программирования. С помощью негладких штрафов задача сводится к задаче безусловной минимизации негладкой функции. Предлагается алгоритм поиска локальных экстремумов негладкой функции и алгоритм уточнения оценки снизу для значения глобального минимума целевой функции, которые базируются на использовании методов оптимизации негладких функций с применением модификации r -алгоритма Шора. Приводятся результаты тестовых экспериментов.

Задача равновесной упаковки неодинаковых кругов в круг наименьшего радиуса возникает в задачах плотной упаковки параллельных одинаковых по высоте круговых цилиндров в цилиндрический контейнер при ограничениях на динамическое равновесие системы [1, 2]. Динамическое равновесие определяется требованием, чтобы центр тяжести системы круговых цилиндров находился в центре кругового контейнера.

Математическая модель задачи равновесной упаковки неравных кругов может быть сформулирована в виде различных многоэкстремальных задач математического программирования [3]. Одна из этих формулировок является предметом исследования данной работы. Для нее мы опишем алгоритм нахождения локальных экстремумов и алгоритм уточнения оценки снизу для значения глобального минимума целевой функции, которые базируются на использовании методов оптимизации негладких функций.

Математическая модель. Имеется семейство кругов S_i с радиусами r_i и весами w_i , $i = 1, \dots, m$. Полагаем, что центр тяжести круга S_i находится в его центре. Равновесной упаковкой семейства кругов S_i , $i = 1, \dots, m$, в круг S назовем такую их упаковку, чтобы радиус круга S был минимальным и центр тяжести семейства кругов S_i , $i = 1, \dots, m$, совпадал с центром круга S .

Не ограничивая общности будем считать, что центр круга S находится в начале неподвижной системы координат. Пусть (x_i, y_i) – неизвестный центр круга S_i , r – неизвестный радиус круга S . Обозначим известные величины $\lambda_i = w_i / \sum_{i=1}^m w_i$, $i = 1, \dots, m$, и очевидную нижнюю границу на искомый радиус $r_{low} = \max_{i=1, \dots, m} r_i$. Тогда равновесной упаковке семейства кругов S_i , $i = 1, \dots, m$, соответствует многоэкстремальная задача нелинейного программирования:

$$r^* = \min_{x, y, r} r \quad (1)$$

при ограничениях

$$x_i^2 + y_i^2 \leq (r - r_i)^2, \quad i = 1, \dots, m, \quad (2)$$

$$(x_i - x_j)^2 + (y_i - y_j)^2 \geq (r_i + r_j)^2, \quad 1 \leq i < j \leq m, \quad (3)$$

$$\sum_{i=1}^m \lambda_i x_i = 0, \quad \sum_{i=1}^m \lambda_i y_i = 0, \quad (4)$$

$$r \geq r_{low}, \quad (5)$$

где $x = (x_1, \dots, x_m)$, $y = (y_1, \dots, y_m)$. Здесь целевая функция (1) является линейной. Ограничения (2) гарантируют, что $S_i \subset S$, а ограничения (3) описывают условие $\text{int } S_i \cap \text{int } S_j = \emptyset$, $1 \leq i < j \leq m$, где $\text{int}(\cdot)$ означает внутренность множества (\cdot) . Ограничения (4) означают, что центр тяжести семейства кругов S_i , $i = 1, \dots, m$, находится в центре круга S . Ограничение (5) обеспечивает то, что значение радиуса круга S не уходит к минус бесконечности, чему формально не препятствует правая часть ограничения (2).

В работе [3] приведены еще две формулировки этой задачи. Первая является задачей обратно-выпуклого программирования, а вторая – задачей минимизации функции максимума из выпуклых функций при ограничениях (3) и (4). Во второй формулировке переменная r не используется, и ее оптимальное значение r^* определяется из минимального значения негладкой целевой функции. Обе формулировки свободны от ограничения (5), так как неотрицательность r учитывается за счет формулировки ограничения (2) в виде $\sqrt{x_i^2 + y_i^2} \leq r - r_i$, $i = 1, \dots, m$.

Алгоритм поиска наилучшего решения. С помощью негладких штрафов задача (1)–(5) сводится к задаче безусловной минимизации негладкой функции

$$\min_{r, x, y} \{f(r, x, y) = r + \Phi_F(r, x, y)\} \quad (6)$$

где штрафная функция $\Phi_P(r, x, y)$ имеет вид

$$\Phi_P(r, x, y) = P_1 F_1(r, x, y) + P_2 F_2(x, y) + P_3 \max\{0, -r + r_{low}\} \quad (7)$$

Здесь P_k – положительные штрафные коэффициенты, $k = 1, 2, 3$, а функции $F_1(r, x, y)$ и $F_2(x, y)$ определяются так:

$$F_1(r, x, y) = \sum_{i=1}^m \max\{0, x_i^2 + y_i^2 - (r - r_i)^2\} + \sum_{i=1}^m \sum_{j=i+1}^m \max\{0, -(x_i - x_j)^2 - (y_i - y_j)^2 + (r_i + r_j)^2\}, \quad (8)$$

$$F_2(x, y) = \max\{0, -\sum_{i=1}^m \lambda_i x_i - \Delta x, \sum_{i=1}^m \lambda_i x_i - \Delta x\} + \max\{0, -\sum_{i=1}^m \lambda_i y_i - \Delta y, \sum_{i=1}^m \lambda_i y_i - \Delta y\}, \quad (9)$$

где Δx и Δy – заданные допуски на отклонения координат центра тяжести семейства кругов от начала координат. Использование в (7) штрафных коэффициентов P_k , $k = 1, 2, 3$, позволяет учесть точность выполнения ограничений (2)–(5). Коэффициент P_1 , согласно (8), отвечает за ограничения (2),(3), коэффициент P_2 , согласно (9), – за ограничения (4), а коэффициент P_3 – за ограничение (5).

Алгоритм поиска наилучшего решения задачи (1)–(5) состоит в следующем. Для заданного набора стартовых точек осуществляется поиск локальных минимумов в задаче (6) с помощью модификации r -алгоритма [4]. Наилучший из локальных минимумов функции $f(r, x, y)$, для которой штрафная функция $\Phi_P(r, x, y)$ близка к нулю, принимается за решение задачи (1)–(5). Ему соответствует значение целевой функции r_{up} – наилучшее значение радиуса r круга S . Стартовые точки генерируются случайным образом в круге заданного радиуса, который последовательно уточняется по мере нахождения лучшего локального минимума. Отметим, что данный алгоритм можно использовать и в случае, когда не требуется учитывать ограничения на центр тяжести. Для этого достаточно положить равным нулю штрафной коэффициент P_2 .

Програмная реализация алгоритма выполнена на некоммерческом языке GNU Octave [5]. Программа либо находит один из локальных минимумов в задаче (1)–(5), либо сообщает о невозможности найти допустимую точку для системы ограничений (2)–(5). Ядром программы яв-

ляется octave-функция galgb5, которая реализует r -алгоритм с постоянной величиной коэффициента растяжения пространства и адаптивной регулировкой шага в направлении нормированного антисубградиента. Эта регулировка направлена на увеличение точности поиска минимума функции по направлению в процессе счета и при этом гарантирует, что среднее (по итерациям) число шагов не превышает двух-трех.

Двойственная оценка ψ^* . Эта оценка аппроксимирует снизу минимальное значение целевой функции в квадратичных невыпуклых задачах, и ее значение со сколь угодно большой точностью может быть найдено с помощью методов минимизации негладких выпуклых функций [6]. Модель (1)–(5) можно преобразовать к виду квадратичной экстремальной задачи

$$f^* = (r^*)^2 = \min_{r,x,y} r^2, \quad (10)$$

при ограничениях

$$x_i^2 + y_i^2 - r^2 + 2r_i r - r_i^2 \leq 0, \quad i = 1, \dots, m, \quad (11)$$

$$-x_i^2 + 2x_i x_j - x_j^2 - y_i^2 + 2y_i y_j - y_j^2 + (r_i + r_j)^2 \leq 0, \quad 1 \leq i < j \leq m, \quad (12)$$

$$\sum_{i=1}^m \sum_{j=1}^m \lambda_i \lambda_j x_i x_j = 0, \quad \sum_{i=1}^m \sum_{j=1}^m \lambda_i \lambda_j y_i y_j = 0, \quad (13)$$

$$r^2 - (r_{low} + r_{up})r + r_{low} r_{up} \leq 0. \quad (14)$$

Здесь ограничения (11) и (12) соответствуют другой записи ограничений (2) и (3), два ограничения в (13) являются возведенными в квадрат линейными равенствами из (4). Квадратичное неравенство (14) следует из соотношения $r_{low} \leq r \leq r_{up}$.

Наличие квадратичного неравенства (14) обеспечивает нетривиальную, т.е. не равную $-\infty$, двойственную оценку ψ^* , которая будет оценкой снизу для f^* в задаче (10)–(14). Более того, оно гарантирует, что оценка ψ^* будет всегда не меньше, чем квадрат r_{low} . Поэтому, если алгоритм поиска наилучшего решения позволяет уточнять величину r_{up} , то двойственная оценка ψ^* позволяет уточнять величину r_{low} . На самом деле, из свойства оценки $\psi^* \leq f^* \leq (r^*)^2$ следует $r^* \geq \sqrt{\psi^*} \geq r_{low}$, откуда понятно, что если величина $\sqrt{\psi^*}$ больше, чем r_{low} , то она может быть использована для уточнения нижней оценки r_{low} . Это означает, что для рассматриваемого класса задач вида (1)–(5) оценка ψ^* может использоваться для доказательства того, что найден глобальный минимум.

Рассмотрим тестовый пример: $m = 5$, $r_1 = 0, 1$, $r_2 = 0, 2$, $r_3 = 0, 3$, $r_4 = 0, 5$, $r_5 = 0, 8$, $w_1 = 0, 0785$, $w_2 = 0, 314$, $w_3 = 0, 7065$, $w_4 = 1, 9625$, $w_5 = 5, 024$. Полагая $P_1 = P_2 = P_3 = 10$ и $\Delta x = \Delta y = 0, 0001$ алгоритм уже для 4-й стартовой точки находит точку локального минимума (r_{up}, x_{up}, y_{up}) , где

$$r_{up} = 1, 316108,$$

$$x_{up} = (-0, 474894, -1, 115151, 0, 025054, -0, 615244, 0, 314084),$$

$$y_{up} = (1, 119551, -0, 046204, 1, 015799, 0, 536197, -0, 372840).$$

Значение функции цели не улучшилось для 100 сгенерированных стартовых точек. При генерации стартовых точек использовался датчик случайных чисел с равномерным распределением внутри единичного куба. В точке (r_{up}, x_{up}, y_{up}) ограничения (12) выполняются с некоторым запасом до 10^{-7} , а ограничения (11) нарушаются не более, чем на величину 10^{-7} . С учетом квадратичности этих ограничений это означает, что условие касания кругов выполнено с точностью до $\sqrt{10^{-7}} = 0, 0003$, т.е. немного хуже, чем $\Delta x = \Delta y = 0, 0001$.

Естественен вопрос: может ли найденный наилучший локальный минимум претендовать на то, чтобы быть глобальным минимумом? Оказывается, что может, и доказать это можно с помощью двойственной оценки ψ^* . Так, например, пусть $r_{low} = 0, 8$ (максимальный из пяти радиусов) и $r_{up} = 1, 35$ (немного больше, чем $1, 316108$, поскольку ограничения точно не выполняются). Тогда полученная оценка $\psi^* = 1, 7309$. Отсюда получаем, что $r^* \geq \sqrt{\psi^*} \geq 1, 3156$ и есть больше, чем $1, 3 = r_4 + r_5$ — граница снизу. Следовательно, улучшить значение целевой функции $f^* = 1, 316108$ меньше, чем на величину $0, 0005$, нельзя. А это и есть доказательство того, что найденный локальный минимум является глобальным с наперед заданной точностью.

В заключение отметим, что двойственную оценку ψ^* можно уточнять за счет добавления в задаче (10)–(14) функционально избыточных квадратичных ограничений [6], которые являются нетривиальными следствиями условий задачи. Так, например, подобно ограничению (14) можно построить и ограничения для некоторых переменных из $x_i, y_i, i = 1, \dots, m$, используя диапазоны их изменения.

Работа поддержана совместным грантом НТЦУ и НАН Украины (проект № 5710).

Литература

1. КОВАЛЕНКО А.А., ПАНКРАТОВ А.В., РОМАНОВА Т.Е., СТЕЦЮК П.И. Упаковка круговых цилиндров в цилиндрический контейнер с учетом специальных ограничений поведения системы // Журнал обчислювальної та прикладної математики. – 2013. – № 1(111). – С. 126–134.
2. STOYAN YU., ROMANOVA T. Mathematical models of placement optimisation: two- and three-dimensional problems and applications // Modeling and Optimization in Space Engineering / G. Fasano, J.D. Pinter, eds. – New York: Springer, 2012. – P. 363–388.
3. НЕНАХОВ Э.И., РОМАНОВА Т.Е., СТЕЦЮК П.И. Равновесная упаковка кругов в круг минимального радиуса // Теорія оптимальних рішень. – Київ: Ін-т кібернетики ім. В.М. Глушкова НАН України, 2013. – С. 143–153.
4. 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.
5. Octave [Электронный ресурс]: <http://www.octave.org>.
6. SHOR N.Z. Nondifferentiable optimization and polynomial problems. – Dordrecht: Kluwer, 1998. – 394 p.

Поступила в редакцию 26.12.2013

Задача равновесной упаковки кругов

П.И. Стецюк, А.П. Лиховид
Раздел 1 отчета о НИР¹, С. 11–31.

Задача равновесной упаковки неодинаковых кругов в круг наименьшего радиуса возникает в задачах плотной упаковки параллельных одинаковых по высоте круговых цилиндров в цилиндрический контейнер при ограничениях на динамическое равновесие системы [1, 2]. Равновесная упаковка означает выполнение требования, что центр тяжести системы круговых цилиндров находится в центре кругового контейнера. На рисунке 1 проиллюстрировано оптимальное размещение 49-ти круговых цилиндров одинаковой плотности в круговой контейнер радиуса R из работы [2].

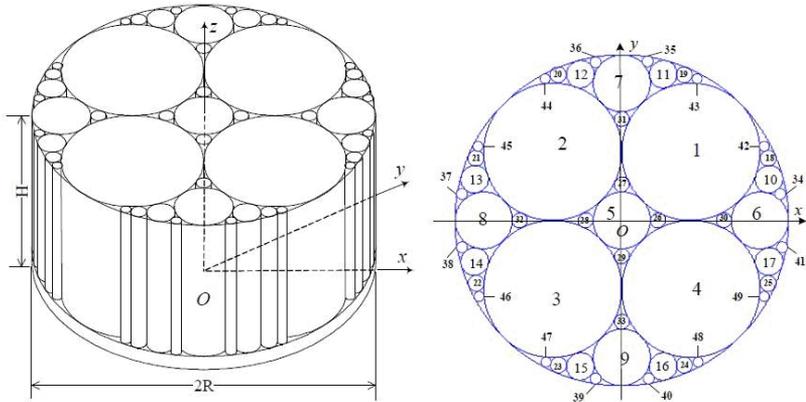


Рис. 1. Упаковка 49-ти круговых цилиндров в круговой контейнер

Здесь все круговые цилиндры имеют одинаковую высоту H (рисунок 1, слева), а в их основании лежат круги, порядковые номера которых

¹Разработать последовательные и параллельные алгоритмические и программные средства для решения прикладных экстремальных задач методами негладкой оптимизации /П.И. Стецюк, А.П. Лиховид и другие // Заключительный отчет о научно-исследовательской работе № гос. регистрации 0113U003146. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2015. – 168 с.

приведены на рисунке 1, справа. Круги 1-4 имеют радиус $(\sqrt{2} - 1) R$, круги 5-9 – радиус $(3 - 2\sqrt{2}) R$, круги 10-17 – радиус $((\sqrt{2} - 1) / 5) R$, круги 18-25 – радиус $((10\sqrt{2} - 7) / 151) R$, круги 26-33 имеют радиус $((8\sqrt{2} - 11) / 7) R$, круги 34-41 – радиус $((\sqrt{2} - 1) / 13) R$, а круги 42-49 имеют радиус $((17\sqrt{2} - 9) / 497) R$.

Задача равновесной упаковки кругов рассматривается в различных формулировках во многих публикациях, см. например [3–6]. Задача является многоэкстремальной и методы ее решения в основном используют эвристические подходы в сочетании с градиентными процедурами и методами поиска локального экстремума. В статье [3] представлен алгоритм (basin filling algorithm), который объединяет улучшенный алгоритм, в англоязычной литературе носящий название “energy landscape raving method”, градиентный метод и эвристический механизм обновления конфигурации. Быстрый эвристический алгоритм с хорошими полученными результатами и временем решения представлен в работе [4]. В статье [5] предлагаются конструктивные эвристические процедуры. Эти процедуры сначала оптимизируют порядок размещения объектов, а затем оптимизируют их позиционирование. Далее описывается алгоритм оптимизации, так называемый алгоритм «колонии муравьев», который использует эти эвристические процедуры. В статье представлены вычислительные результаты, которые показывают, что этот алгоритм для больших примеров работает эффективнее, чем существующие методы. В работе [6] авторы предлагают технику позиционирования, основанную на порядке размещения объектов с использованием генетического алгоритма для поиска возможного размещения объектов.

Ниже опишем математическое и программное обеспечение для задачи равновесной упаковки кругов, основанное на решении задачи нелинейного программирования. Математическая модель равновесной упаковки неравных кругов описана в разделе 1. В разделе 2 изложен последовательный алгоритм поиска наилучшего решения и его программная реализация, а в разделе 3 описан параллельный алгоритм и его программная реализация. В разделе 4 описана квадратичная экстремальная задача и соответствующая ей двойственная оценка ψ^* . В разделе 5 обсуждаются направления дальнейшего усовершенствования оптимизационных моделей, алгоритмов нахождения приближенных решений и двойственных оценок для соответствующих им квадратичных задач.

1. Формулировка оптимизационной задачи

Имеется семейство кругов S_i с радиусами r_i и весами w_i , $i = 1, \dots, m$. Полагаем, что центр тяжести круга S_i находится в его центре. Равновесной упаковкой семейства кругов S_i , $i = 1, \dots, m$, в круг S назовем такую их упаковку, чтобы радиус круга S был минимальным и центр тяжести семейства кругов S_i , $i = 1, \dots, m$, совпадал с центром внешнего круга S .

Не ограничивая общности будем считать, что центр круга S находится в начале системы координат. Пусть (x_i, y_i) – неизвестный центр круга S_i , r – неизвестный радиус круга S . Обозначим известные величины $\lambda_i = w_i / \sum_{j=1}^m w_j$, $i = 1, \dots, m$, и очевидную нижнюю границу на искомый радиус $r_{low} = \max_{i=1, \dots, m} r_i$. Тогда равновесной упаковке семейства кругов S_i , $i = 1, \dots, m$, соответствует многоэкстремальная задача нелинейного программирования [7, 8]

$$r^* = \min_{r, x, y} r \quad (1)$$

при ограничениях

$$x_i^2 + y_i^2 \leq (r - r_i)^2, \quad i = 1, \dots, m, \quad (2)$$

$$(x_i - x_j)^2 + (y_i - y_j)^2 \geq (r_i + r_j)^2, \quad 1 \leq i < j \leq m, \quad (3)$$

$$\sum_{i=1}^m \lambda_i x_i = 0, \quad \sum_{i=1}^m \lambda_i y_i = 0, \quad (4)$$

$$r \geq r_{low}, \quad (5)$$

где $x = (x_1, \dots, x_m)$, $y = (y_1, \dots, y_m)$.

Целевая функция (1) является линейной функцией и связана с минимизацией r – радиуса круга S . Квадратичные ограничения (2) означают, что каждый круг S_i , $i = 1, \dots, m$ содержится внутри круга S . Квадратичные ограничения (3) гарантируют, что никакие два круга из семейства S_i , $i = 1, \dots, m$ не пересекаются (не имеют общих внутренних точек). Линейные ограничения (4) означают, что центр тяжести семейства кругов S_i , $i = 1, \dots, m$, локализован в начале координат. Ограничение (5) обеспечивает положительность значения r , которое требуется для корректности ограничений (2).

Две альтернативные формулировки задачи равновесной упаковки неравных кругов приведены в статье [7]. Они свободны от ограничения (5), так как связаны с заменой ограничений (2) на следующие ограничения

$$\sqrt{x_i^2 + y_i^2} \leq r - r_i, \quad i = 1, \dots, m, \quad (2')$$

которые автоматически учитывают положительность переменной r . Первая формулировка является задачей обратно-выпуклого программирования. Она связана с минимизацией целевой функции (1) при ограничениях (2'), (3) и (4). Вторая формулировка связана с минимизацией негладкой выпуклой функции и состоит в нахождении

$$r^* = \min_{x,y} \max_{i=1,\dots,m} \left\{ \sqrt{x_i^2 + y_i^2} + r_i \right\}, \quad (1')$$

при ограничениях (3) и (4). Здесь переменная r не используется. Ее оптимальное значение r^* определяется из минимального значения негладкой целевой функции.

2. Последовательный алгоритм поиска наилучшего решения

С помощью негладких штрафов задача (1)–(5) сводится к задаче безусловной минимизации негладкой функции

$$\min_{r,x,y} \{f(r, x, y) = r + \Phi_P(r, x, y)\}, \quad (6)$$

где штрафная функция $\Phi_P(r, x, y)$ имеет вид

$$\Phi_P(r, x, y) = P_1 F_1(r, x, y) + P_2 F_2(x, y) + P_3 \max\{0, -r + r_{low}\}. \quad (7)$$

Здесь $P = \{P_1, P_2, P_3\}$, где $P_k, k = 1, 2, 3$ – положительные штрафные коэффициенты, а функции $F_1(r, x, y)$ и $F_2(x, y)$ определяются так:

$$F_1(r, x, y) = \sum_{i=1}^m \max\{0, x_i^2 + y_i^2 - (r - r_i)^2\} + \\ + \sum_{i=1}^m \sum_{j=i+1}^m \max\{0, -(x_i - x_j)^2 - (y_i - y_j)^2 + (r_i + r_j)^2\}, \quad (8)$$

$$F_2(x, y) = \max\left\{0, \sum_{i=1}^m \lambda_i x_i - \Delta x, - \sum_{i=1}^m \lambda_i x_i - \Delta x\right\} + \\ + \max\left\{0, \sum_{i=1}^m \lambda_i y_i - \Delta y, - \sum_{i=1}^m \lambda_i y_i - \Delta y\right\}, \quad (9)$$

где Δx и Δy – заданные допуски на отклонения координат центра тяжести семейства кругов от начала координат.

Нахождение локального минимума задачи (1)–(5) можно заменить на поиск локального минимума задачи (6)–(9), которая является задачей безусловной минимизации многоэкстремальной негладкой функции $f(r, x, y)$. Если при некоторых положительных значениях штрафных коэффициентов $P = \{P_1, P_2, P_3\}$ локальному минимуму функции $f(r, x, y)$ соответствует равное нулю значение штрафной функции $\Phi_P(r, x, y)$, то он будет локальным минимумом задачи (1)–(5). Выбор штрафных коэффициентов P_1, P_2 и P_3 позволяет учитывать точность выполнения ограничений (2)–(5). Коэффициент P_1 отвечает за "суммарное нарушение" квадратичных ограничений (2)–(3), коэффициент P_2 – за "суммарное нарушение" линейных ограничений (4), а коэффициент P_3 – за "нарушение" ограничения (5).

Последовательный алгоритм поиска наилучшего решения задачи (1)–(5) базируется на методе мультистарта со случайным выбором стартовых точек и модификации r -алгоритма Шора [7] для поиска локальных минимумов функции $f(r, x, y)$. Алгоритм не требует, чтобы стартовые точки были допустимыми для задачи (1)–(5). Его можно использовать и при отсутствии ограничений на центр тяжести системы кругов. Для этого достаточно положить $P_2 = 0$, что равносильно тому, что из задачи (1)–(5) исключаются ограничения (4).

Последовательный алгоритм состоит в следующем. Пусть n_{test} – количество стартовых точек, которые будут генерироваться с помощью датчика равномерного распределения в круге радиуса r_{up} (наилучшая верхняя оценка радиуса внешнего круга). В начале процесса наилучшая верхняя оценка радиуса устанавливается равной $r_{up} = \sum_{i=1}^m r_i$ и последовательно уточняется для каждой очередной стартовой точки по мере нахождения локального минимума с меньшим значением целевой функции. Поиск локального минимума функции $f(r, x, y)$ осуществляется с помощью модификации r -алгоритма с постоянным коэффициентом растяжения пространства и адаптивной регулировкой шага [9], стр. 384–385. Наилучший локальный минимум функции $f(r, x, y)$, в котором ре-

лизуется близкое к нулю значение штрафной функции $\Phi_P(r, x, y)$, принимаем за решение задачи (1)–(5). Ему соответствует наилучшее значение радиуса r_{up} и соответствующие ему координаты центров размещаемых кругов (x^{up}, y^{up}) .

Программная реализация последовательного алгоритма поиска наилучшего решения задачи (1)–(5) выполнена в виде octave-функции **bcp(rw, pen, dxy, ntest)**. Здесь **rw** – данные о задаче, **pen** – штрафные параметры, **dxy** – допуски на отклонения координат центра тяжести семейства кругов от начала координат, **ntest** – количество стартовых точек. Ядром функции **bcp** служит octave-функция **ralgb5** ([9], стр. 384–385), для которой вычисление обобщенного градиента функции $\Phi_P(r, x, y)$ реализовано octave-функцией **fgbcp**. Параметры r -алгоритма присваиваются в теле функции **bcp** и могут изменяться пользователем. Программа подробно описана в авторском свидетельстве [10], где приведены код программы, пример протокола ее работы и интерфейс пользователя для работы с программой. Для запуска программы **bcp** необходим Open Source-пакет для математических вычислений **GNU Octave** [11]. Программа использует версии Octave 3.0.0. и выше. Для нее не требуется никаких специальных конфигураций компьютера.

С помощью последовательного алгоритма более чем на десять единиц уменьшен радиус внешнего круга для равновесного размещения 40 кругов из работы [3]. Радиусы r_i и веса w_i , кругов $i = 1, 2, \dots, 40$ для этого примера даны в таблице 1 (первый и четвертый столбцы). В работе [3] найден радиус внешнего круга $r = 725.0435$ и приведено соответствующее ему размещение кругов. Вычислительное время составило 71.87 секунд на персональном компьютере с процессором Intel Core 2 Duo, 1.6 GHz (язык Java). Последовательный алгоритм нашел радиус $r_{up} = 714.563$ [12], улучшив радиус внешнего круга из работы [3] на 1.47%. Вычислительное время составило 396.10 секунд на персональном компьютере Pentium, 3 GHz (GNU Octave, версия 3.6.4). Координаты размещаемых кругов (x_i^{up}, y_i^{up}) , $i = 1, 2, \dots, 40$, которые соответствуют радиусу внешнего круга $r_{up} = 714.563$, даны в таблице 1 (второй и третий, пятый и шестой столбцы).

Следует заметить, что для этого примера ограничения (2) и ограничения (3) выполняются с точностью 10^{-3} в точке локального минимума (r^{up}, x^{up}, y^{up}) , $x^{up} = (x_1^{up}, \dots, x_{40}^{up})$, $y^{up} = (y_1^{up}, \dots, y_{40}^{up})$. Это обеспечивается благодаря поиску решения (x_i^{up}, y_i^{up}) , $i = 1, \dots, m$ с точностью 10^{-7} для r -алгоритма и последующим округлением этого решения до семи значащих десятичных цифр. Оба условия на центр тяжести системы кругов из ограничения (4) выполняются с точностью 10^{-7} .

Таблица 1. Координаты размещаемых 40 кругов для $r_{up} = 714.563$

(i, r_i, w_i)	x_i^{up}	y_i^{up}	(i, r_i, w_i)	x_i^{up}	y_i^{up}
(1,106,11)	148.3161	-197.173	(21,108,11)	10.53656	606.4715
(2,112,12)	76.15339	-423.691	(22,86,7)	339.8889	-184.372
(3,98,9)	-47.6005	-254.03	(23,93,8)	586.5976	125.8166
(4,105,11)	245.8454	-10.0658	(24,100,10)	523.16	322.3535
(5,93,8)	-589.996	-195.564	(25,102,10)	-308.141	268.5048
(6,103,10)	-412.389	-112.667	(26,106,11)	163.8511	184.3511
(7,82,6)	616.7154	-46.5723	(27,111,12)	-59.5939	-600.614
(8,93,8)	441.9598	-37.3261	(28,107,11)	-427.302	96.80256
(9,117,13)	-351.65	483.1392	(29,109,11)	-163.423	-82.4655
(10,81,6)	-119.181	-418.094	(30,91,8)	266.9723	-354.43
(11,89,7)	-625.331	-17.0275	(31,111,12)	533.5077	-220.714
(12,92,8)	-32.6334	208.8015	(32,91,8)	403.3459	471.1001
(13,109,11)	-511.8	323.6775	(33,101,10)	449.1051	-415.189
(14,104,10)	-147.785	397.8176	(34,91,8)	-190.201	115.7337
(15,115,13)	-261.838	-283.688	(35,108,11)	28.69188	18.43552
(16,110,12)	223.8659	561.5875	(36,114,12)	382.6064	160.9822
(17,114,12)	-474.903	-367.618	(37,118,13)	74.1686	389.6145
(18,89,7)	276.0447	343.8428	(38,85,7)	-609.857	156.283
(19,82,6)	132.5742	-618.514	(39,87,7)	-184.352	599.8748
(20,120,14)	-294.416	-516.55	(40,98,9)	295.1779	-541.313

Последовательный алгоритм позволил уменьшить радиусы внешнего круга для серии примеров с $m = 10, 15, 20, 25, 30, 35, 40, 45, 50, 55$ из работы [4]. Значения r_{up} и их сравнение в радиусами внешнего круга, найденными в работах [4, 5, 6], приведены в таблице 2. Радиусы и веса соответствующих этим примерам коллекций кругов можно найти в работе [4]. Для каждого из примеров радиус внешнего круга r_{up} , найденный последовательным алгоритмом (второй столбец) есть меньше, чем наилучшие радиусы внешнего круга из работ [4, 5, 6] (третий, пятый и седьмой столбцы). Самые лучшие результаты из работ [4, 5, 6] последовательный алгоритм улучшил от 0.42% в примере с $m = 25$ до 2.46% в примере с $m = 55$ (четвертый, шестой и восьмой столбцы).

Для примера $m = 50$ кругов последовательный алгоритм нашел $r_{up} = 122.30$, вычислительное время составило 446.56 секунд на персональном компьютере Pentium, 3 GHz (GNU Octave, версия 3.6.4); алгоритм из [4] нашел $r = 125.62$, вычислительное время составило 12 секунд (Pentium, 3.40 GHz, язык MATLAB); алгоритм из [5] нашел $r = 125.14$, вычислительное время составило 533.328 секунд (Intel 1.83GHz/512Mb, язык C, компилятор MS VC++); алгоритм из [6] нашел $r = 124.91$, вычислительное время составило 47716 секунд (Intel Celeron, 1.5GHz). Для примера $m = 55$ кругов последовательный алгоритм нашел $r_{up} = 134.19$, вычислительное время составило 468.88 секунд; алгоритм из [4] нашел $r = 139.17$, вычислительное время составило 14 секунд; алгоритм из [5] нашел $r = 137.49$, вычислительное время составило 703.265 секунд; алгоритм из [6] нашел $r = 137.52$, вычислительное время составило 64363 секунд.

Результаты вычислительного эксперимента показывают, что последовательный алгоритм за приемлемое время позволяет найти достаточно хороший локальный минимум для задачи равновесной упаковки нескольких десятков кругов.

3. Реализация параллельного алгоритма

Параллельный алгоритм отличается от последовательного тем, что он запускает множественный поиск локальных решений с помощью модификации r -алгоритма в отличие от того, что последовательный алгоритм запускает всего один поиск. Реализация параллельного алгоритма использует процедуру "Master-Slave" на $(k+1)$ процессорах. Один из них выбирается "ведущим" (Master), а остальные k - "ведомыми" (Slave). Эта процедура апробирована в статье [13].

Таблица 2. Сравнение r_{up} для серии примеров $m = 10 \div 55$ [4, 5, 6, 12]

m	r_{up}	радиус из [4]	улучшение [4]	радиус из [5]	улучшение [5]	радиус из [6]	улучшение [6]
10	59.34	61.21	3.15%	60.61	2.14%	60.41	1.80%
15	67.13	69.32	3.26%	67.58	0.67%	67.49	0.54%
20	81.50	85.08	4.39%	82.22	0.88%	82.63	1.39%
25	82.77	84.57	2.17%	83.12	0.42%	83.74	1.17%
30	97.70	101.23	3.61%	99.40	1.74%	99.78	2.13%
35	100.60	104.99	4.36%	102.29	1.68%	103.56	2.94%
40	113.34	119.11	5.09%	114.26	0.81%	116.17	2.50%
45	117.08	120.85	3.22%	119.66	2.20%	119.02	1.66%
50	122.30	125.62	2.71%	125.14	2.32%	124.91	2.13%
55	134.19	139.17	3.71%	137.49	2.46%	137.52	2.48%

В начале вычислений в Master-процессоре случайным образом генерируются k стартовых точек в круге с радиусом $r_{up} = \sum_{i=1}^m r_i$ и пересылаются в Slave-процессоры. Slave-процессор занимается поиском локального минимума функции $f(r, x, y)$ для полученной им стартовой точки. Как только r -алгоритм заканчивает работу на каком-либо Slave-процессоре, то результат поиска передается в Master-процессор. Если при этом найден локальный минимум задачи (1)–(5), то значение радиуса внешнего круга сравнивается с наилучшим из найденных к этому моменту значением r_{up} . Если радиус меньше r_{up} , то он становится новым значением r_{up} и запоминаются соответствующие ему значения координат центров размещаемых кругов (x^{up}, y^{up}) . Затем Master-процессор генерирует новую стартовую точку, которая передается для очередного поиска локального минимума в тот Slave-процессор, для которого r -алгоритм закончил работу. Процесс завершается, если превышено заданное количество стартовых точек или заказанное время.

Программная реализация параллельного алгоритма выполнена на языке программирования C++ в среде параллельного программирования MPI. Она оформлена как программа под названием “**A parallel algorithm for a balanced circular packing problem**” и подробно описана в авторском свидетельстве [14], где приведены код программы, пример протокола ее работы и интерфейс пользователя для работы с программой. В качестве генератора псевдослучайных чисел программа использует стандартную функцию **rand** из библиотеки C++. Поиск локальных минимумов осуществляется модулем **ralgb5**, который соответствует octave-коду модификации r -алгоритма из [9], стр. 384–385. Программа предназначена для работы на кластере со средой MPI под управлением операционной системы Linux. Она может работать как на одном процессоре, так и на многих параллельных процессорах.

На кластере СКИТ-3 Института кибернетики им. В.М. Глушкова [15] проведено исследование ускорения и эффективности параллельного алгоритма для задачи равновесной упаковки 40 кругов [13]. Количество генерируемых точек для метода мультистарта равнялось 50, а количество процессоров кластера варьировалось от 1 до 24. Результаты представлены в таблице 3, где N – количество процессоров, t – время решения задачи в секундах, S_i – ускорение параллельного алгоритма, E_i – эффективность параллельного алгоритма.

Параллельное ускорение вычисляется по формуле $S_i = t_0/t_i$, $i = 1, \dots, n$, где t_0 – время работы последовательного алгоритма, а t_i – время решения задачи параллельным алгоритмом с использованием i процессоров. Эффективность параллельного алгоритма вычисляется по фор-

Таблица 3. Время решения на СКИТ-3, ускорение и эффективность

N	t	S_i	E_i	N	T	S_i	E_i
1	26.51	1.00	1.00	8	4.92	5.38	0.67
2	26.71	0.99	0.50	12	3.20	8.27	0.69
3	15.05	1.76	0.59	16	3.37	7.87	0.49
4	10.01	2.65	0.66	20	2.32	11.45	0.57
5	7.94	3.36	0.67	24	2.83	9.37	0.39

муле $E_i = S_i/i$. Из таблицы 3 видим, что для четырех процессоров время решения задачи уменьшается в 2,65 раза, для восьми – в 5,38 раза, шестнадцати – в 7,87 раза, а для двадцати четырех – в 9,37 раза.

На рисунке 2 приведена диаграмма, которая отражает зависимость времени решения задачи от количества используемых процессоров. Из нее легко видеть, что начиная с 12 процессоров дальнейшее добавление процессоров не оказывает существенного влияния на время решения задачи. Это связано с ростом потерь на обмены информацией при увеличении степени параллелизма. В целом ускорение и эффективность параллельного алгоритма соответствуют стандартным характеристикам для процедуры “Master-Slave”, когда время решения каждой отдельной подзадачи на Slave-процессоре сравнительно небольшое.

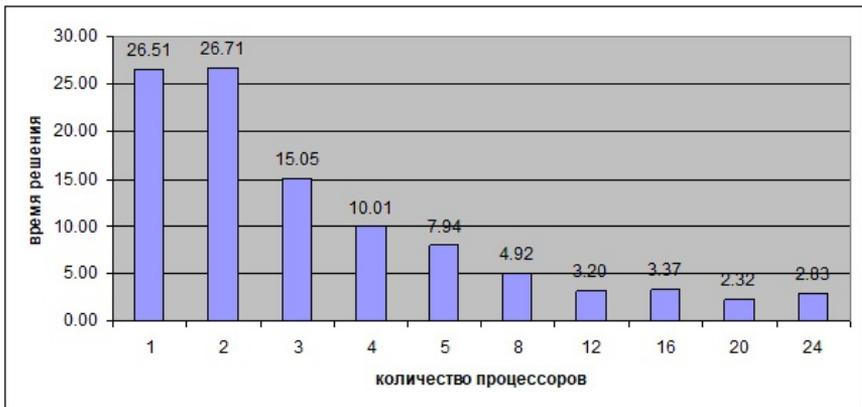


Рис. 2. Диаграмма (время решения/количество процессоров СКИТ-3)

Параллельная программа позволила улучшить найденный последовательным алгоритмом радиус внешнего круга $r_{up} = 714.563$ для рав-

новесной упаковки 40 кругов. Два близких значения $r_{up} = 711.9252$ и $r_{up} = 711.9522$ (уменьшены более чем на две единицы) получены при использовании 24-х процессоров СКИТ-3 и очень большого количества стартовых точек (равнялось 5.000.000). Координаты центров размещаемых кругов, округленные до четырех значащих цифр после точки, приведены в таблице 4 (столбцы 4 и 5 соответствуют $r_{up} = 711.9252$, а столбцы 6 и 7 – $r_{up} = 711.9522$).

Для удобства в таблице 4 приведены также радиусы r_i и веса w_i кругов $i = 1, 2, \dots, 40$ (второй и третий столбцы), они такие же как и в таблице 1. Время выполнения программы в первом случае составило 104762 секунд (приблизительно 29 часов 10 минут), а во втором – 120910 секунд (приблизительно 33 часа 30 минут). Двум найденным локальным минимумам соответствуют существенно различные размещения кругов (для первого минимума они приведены на рисунке 3, а для второго – на рисунке 4). Различные траектории поиска наилучшего радиуса внешнего круга объясняются тем, что параллельная программа существенно зависит от конфигурации загружаемых процессоров.

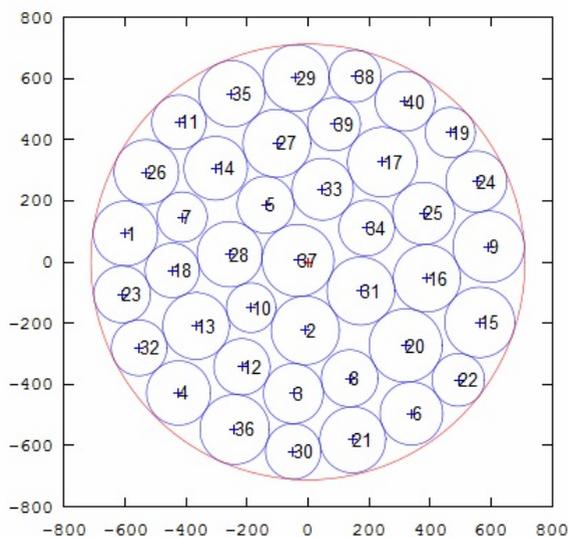


Рис. 3. Размещение 40 кругов для $r_{up} = 711.9252$

Таблица 4. Размещение 40 кругов для $r_{up} = 711.9252$ и $r_{up} = 711.9522$

i	r_i	w_i	x_i^{up}	y_i^{up}	x_i^{up}	y_i^{up}
1	106	11	-598.7338	93.0767	50.7008	-211.0646
2	112	12	-8.6709	-223.7210	417.2515	-103.7310
3	98	9	-47.9483	-430.0152	-605.3276	102.5470
4	105	11	-424.4947	-428.1792	130.1298	176.7494
5	93	8	-139.2399	187.0966	-394.1751	477.2084
6	103	10	339.3850	-495.2939	-58.3472	606.1505
7	82	6	-412.7108	146.7259	205.7851	-104.7966
8	93	8	136.4842	-380.3600	283.7769	51.8631
9	117	13	592.8649	49.4694	-354.9349	-164.0976
10	81	6	-186.7412	-149.2900	461.7161	84.0772
11	89	7	-423.4067	456.9055	592.3418	192.8748
12	92	8	-216.7105	-342.7265	-146.5073	-193.3742
13	109	11	-367.0216	-209.2815	359.4074	484.1257
14	104	10	-302.8163	306.2155	355.3129	235.4158
15	115	13	563.4406	-197.1153	-586.6447	-110.4535
16	110	12	390.4387	-53.2565	-84.1099	194.8138
17	114	12	243.8727	326.4675	156.6935	577.0563
18	89	7	-445.4699	-27.4853	-434.8079	25.7874
19	82	6	466.6679	423.1156	-396.0844	-465.1835
20	120	14	320.5037	-273.0947	-510.8747	299.0226
21	108	11	146.7110	-581.3039	-315.0662	182.2192
22	86	7	493.1765	-385.4340	-244.9335	576.0415
23	93	8	-609.8479	-105.6127	-11.9597	-600.0933
24	100	10	552.6614	262.7126	-251.3826	-354.7961
25	102	10	379.6380	158.4682	284.1286	-271.2848
26	106	11	-530.0235	293.6332	-514.8889	-319.4801
27	111	12	-102.4866	387.7584	-212.4027	-562.1643
28	107	11	-256.6874	25.2136	14.7713	-1.1169
29	109	11	-38.9945	601.6629	9.5224	405.3080
30	91	8	-48.6834	-619.0138	510.6637	353.2765
31	111	12	173.4128	-94.9789	508.6244	-320.0698
32	91	8	-553.7740	-280.8603	203.9259	358.3262
33	101	10	48.0307	237.7494	-58.6326	-411.7913
34	91	8	192.4937	111.2803	172.0081	-596.6531
35	108	11	-249.7244	549.8758	-200.2283	-0.7241
36	114	12	-240.7383	-547.3204	-211.8952	378.7892
37	118	13	-32.5944	5.0314	346.8533	-482.1535
38	85	7	154.9182	607.4830	608.4048	-151.3693
39	87	7	85.4279	450.1455	624.6365	19.8631
40	98	9	318.2680	524.9854	140.3620	-410.3213

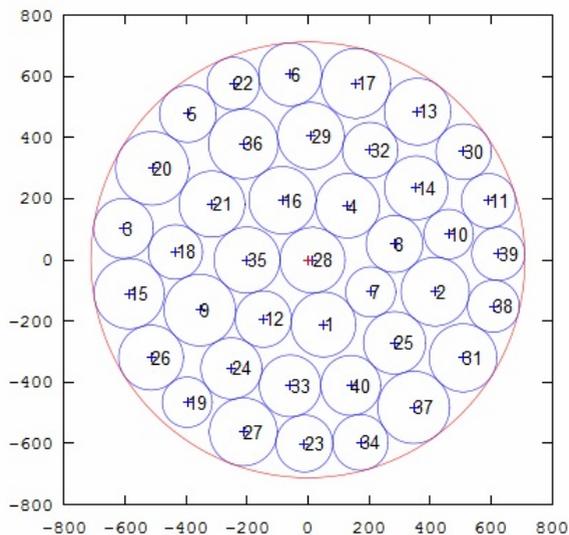


Рис. 4. Размещение 40 кругов для $r_{up} = 711.9522$

4. Двойственная оценка ψ^*

Двойственная оценка предложена Н.З.Шором (см. [16], с. 220–221) и позволяет найти нижнюю границу для значения целевой функции в квадратичных невыпуклых задачах. Значение оценки ψ^* можно вычислить с произвольной заданной точностью с помощью методов минимизации негладких выпуклых функций. Задачу (1)–(5) можно преобразовать в следующую квадратичную экстремальную задачу

$$f^* = (r^*)^2 = \min_{r, x, y} r^2 \quad (10)$$

при ограничениях

$$x_i^2 + y_i^2 - r^2 + 2r_i r - r_i^2 \leq 0, \quad i = 1, \dots, m, \quad (11)$$

$$-x_i^2 + 2x_i x_j - x_j^2 - y_i^2 + 2y_i y_j - y_j^2 + (r_i + r_j)^2 \leq 0, \quad 1 \leq i < j \leq m, \quad (12)$$

$$\sum_{i=1}^m \sum_{j=1}^m \lambda_i \lambda_j x_i x_j = 0, \quad \sum_{i=1}^m \sum_{j=1}^m \lambda_i \lambda_j y_i y_j = 0, \quad (13)$$

$$r^2 - (r_{low} + r_{up})r + r_{low}r_{up} \leq 0. \quad (14)$$

Здесь квадратичные неравенства (11) и (12) соответствуют ограничениям (2) и (3), два квадратичных равенства в (13) соответствуют возведенным в квадрат равенствам из (4). Квадратичное неравенство (14) следует из условия $r_{low} \leq r \leq r_{up}$ и получено умножением линейных неравенств $r_{low} - r \leq 0$ и $r_{up} - r \geq 0$.

Неравенство (14) обеспечивает нетривиальную, т.е. не равную $-\infty$, двойственную оценку ψ^* , значение которой всегда не меньше, чем $(r_{low})^2$. Это следует из того, что, если из задачи (10)–(14) убрать ограничения (11), (12) и (13), то получаем одномерную задачу выпуклого программирования:

$$f^* = (r^*)^2 = \min_r r^2 \quad (15)$$

при ограничении

$$r^2 - (r_{low} + r_{up})r + r_{low}r_{up} \leq 0, \quad (16)$$

оптимальное решение которой $r^* = r_{low}$ и $f^* = r_{low}^2$. Для задачи (15)–(16) оценка $\psi^* = (r_{low})^2$ является точной. Ее значение не может быть уменьшено, если к задаче (15)–(16) прибавить ограничения (11), (12) и (13).

Если последовательный и параллельный алгоритмы позволяют улучшить r_{up} – границу сверху для радиуса внешнего круга, то двойственная оценка ψ^* может позволить улучшить значение r_{low} – границу снизу для радиуса внешнего круга. Действительно, соотношение $r^* \geq \sqrt{\psi^*} \geq r_{low}$ следует из соотношения $\psi^* \leq f^* \leq (r^*)^2$. Очевидно, что если значение $\sqrt{\psi^*}$ больше, чем r_{low} , то оно может быть использовано для улучшения r_{low} . Более того, если удастся показать, что $r_{up} = \sqrt{\psi^*}$, то это означает, что найден глобальный минимум задачи (1)–(5).

Ниже рассмотрим свойства оценки ψ^* для упаковки пяти неравных кругов и упаковки семи кругов, где шесть кругов имеют одинаковый радиус.

Пример 1 [1, 17]. Пусть $m = 5$, $r_1 = 0.1$, $r_2 = 0.2$, $r_3 = 0.3$, $r_4 = 0.5$, $r_5 = 0.8$, $w_1 = 0.0785$, $w_2 = 0.314$, $w_3 = 0.7065$, $w_4 = 1.9625$,

$w_5 = 5.024$. Задача равновесной упаковки этих кругов имеет два локальных минимума, значения целевой функции в которых различаются в четвертом десятичном знаке. Первому локальному минимуму соответствует $r_{up} = 1.3175$ и размещение кругов, которое приведено на рисунке 5 (слева). Второму локальному минимуму соответствует $r_{up} = 1.3161$, что меньше, чем предыдущее значение, на 0,0014. Ему соответствует размещение кругов, которое приведено на рисунке 5 (справа). Оба локальных минимума найдены последовательным алгоритмом при $P_1 = P_2 = P_3 = 10$ и $\Delta x = \Delta y = 0.0001$. Первый получен уже для второй стартовой точки, а второй – для пятнадцатой стартовой точки.

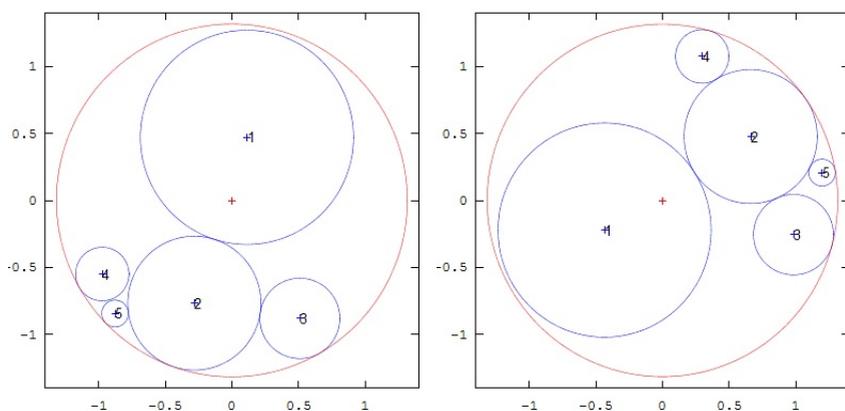


Рис. 5. Размещение 5 кругов: $r_{up} = 1.3175$ (слева), $r_{up} = 1.3161$ (справа)

Поскольку уменьшить значение $r_{up} = 1.3161$ не удавалось и для тысячи стартовых точек, то возник вопрос: не является ли второй локальный минимум глобальным. Ответ на него дан в работе [17], где показано, что радиус внешнего круга $r_{up} = 1.3161$ реализуется в точке глобального минимума целевой функции в задаче (1)–(5). В самом деле, если $r_{low} = 0.8$ и $r_{up} = 1.35$ (чуть больше, чем 1.3161, так как ограничения (11), (12) выполнены с точностью до 0.0001) то получим $\psi^* = 1.7309$ (вычислено с помощью программы DSQTPR [18]). Это означает, что $r^* \geq \sqrt{\psi^*} \geq 1.3156$. Таким образом, значение целевой функции r^* невозможно уменьшить больше, чем на 0,0005, по сравнению с $r_{up} = 1.3161$. Следовательно, с точностью до 0.001 второй локальный минимум для задачи равновесной упаковки 5 кругов является глобальным минимумом, т.е. $r^* = r_{up} = 1.3161$.

Пример 2 [12]. Пусть $m = 7$, $r_i = 0.5$, $i = 1, \dots, 6$, $r_7 = 1$, $w_i = 1$, $i = 1, \dots, 6$, $w_7 = 4$. Если мы положим $P_2 = 0$, тогда последовательный алгоритм находит локальный минимум для задачи упаковки кругов без ограничений на сбалансированность. Если выбрать $P_1 = P_3 = 10$, то алгоритм уже для третьей стартовой точки находит глобальный минимум $r_{up} = r^* = 1.9012$. На рисунке 6 (слева) показано расположение кругов, которое соответствует этому глобальному минимуму. С точностью до поворота решение совпадает с точкой минимума, приведенной для этого примера в работе [19].

При $P_1 = P_2 = P_3 = 10$ и $\Delta x = \Delta y = 0.0001$ последовательный алгоритм нашел $r_{up} = 1.9289$ для двадцать пятой стартовой точки. Ему соответствует равновесная упаковка кругов, которая приведена на рисунке 6 (справа). Значение $r_{up} = 1.9289$ оказалось неулучшаемо последовательным алгоритмом.

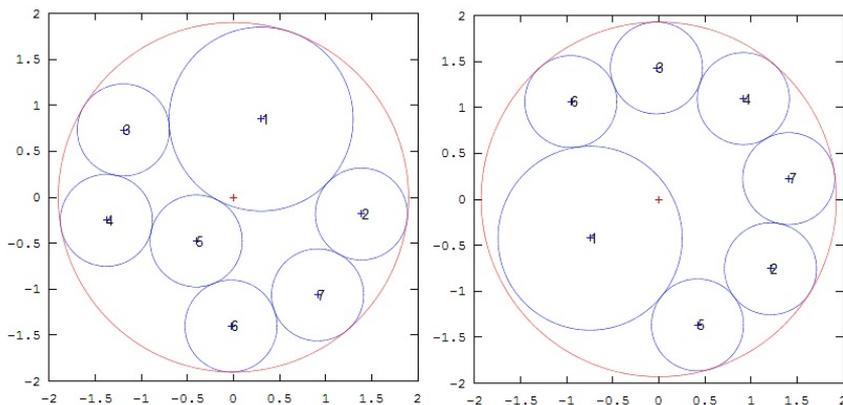


Рис. 6. Наилучшие размещения 7 кругов, справа – равновесное

Оценка ψ^* не является точной для f^* как для случая без учета сбалансированности, так и для случая с учетом сбалансированности кругов. Это легко видеть из таблицы 5. Для случая без учета сбалансированности использование оценки ψ^* позволяет последовательно улучшить r_{low} до значения 1.5394 (см. колонки 2–5). Для случая с учетом сбалансированности мы можем получить $r_{low} = 1.5558$ (см. колонки 6–9 первой и второй строки). Для случая с учетом сбалансированности значения оценки ψ^* для $r_{up} = 2.0, 2.2, 2.5$ (см. строки 4–6 таблицы 5) завышены по отношению к $r_{up} = 1.9289$. Когда $r_{up} = 2.2$, оценка ψ^* по-

Таблица 5. Оценка ψ^* для 7 кругов: случай без учета сбалансированности (колонки 2–5) и случай с учетом сбалансированности (колонки 6–9)

i	r_{low}	r_{up}	ψ^*	$\sqrt{\psi^*}$	r_{low}	r_{up}	ψ^*	$\sqrt{\psi^*}$
1	1.0000	1.9013	2.1357	1.4614	1.0000	1.9290	2.4299	1.5588
2	1.4614	1.9013	2.3440	1.5310	1.5558	1.9290	2.4299	1.5588
3	1.5310	1.9013	2.3670	1.5385	—	—	—	—
4	1.5385	1.9013	2.3693	1.5393	1.0000	2.0000	2.3846	1.5442
5	1.5393	1.9013	2.3696	1.5394	1.0000	2.2000	2.2799	1.5099
6	1.5394	1.9013	2.3696	1.5394	1.0000	2.5000	2.1666	1.4719

звolyет нам устанавливать $r_{low} = 1.5099$, что больше, чем тривиальная оценка $r_{low} = 1.5 = r_1 + r_7$. При $r_{up} = 2.5$ оценка ψ^* дает $r_{low} = 1.4719$, что меньше, чем 1.5.

Заметим, что двойственная оценка ψ^* можно улучшить путем последовательного добавления к задаче (10)–(14) функционально избыточных квадратичных ограничений, которые являются нетривиальными следствиями условий задачи равновесной упаковки кругов. Например, аналогично (14), мы можем построить дополнительные квадратичные неравенства для некоторых переменных $x_i, y_i, i = 1, \dots, m$, используя диапазоны их изменений.

5. Выводы

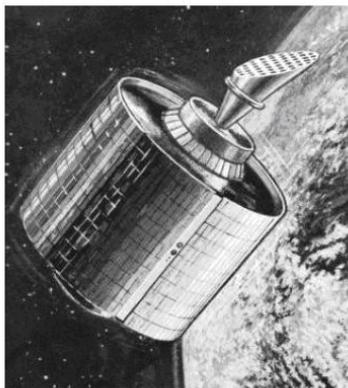
Представлено математическое и программное обеспечение для задачи равновесной упаковки кругов, которая имеет приложения при оптимальной упаковке одинаковых по высоте круговых контейнеров в круговой контейнер минимального радиуса. Разработанные последовательный и параллельный алгоритмы могут быть использованы для равновесной упаковки сорока–пятидесяти контейнеров. Для них решения могут быть найдены за несколько часов на современных компьютерах.

Последовательный и параллельный алгоритмы можно усовершенствовать за счет более рационального выбора стартовых точек. Здесь могут помочь метаэвристики Стояна Ю.Г. [20, 21], Гуляницкого Л.Ф. [22], Шилов В.П. [23] и др. Если стартовые точки будут допустимыми, то актуальным является поиск ближайшего локального минимума штрафной функции. Его можно обеспечить, если модификацию r -алгоритма с регулировкой шага для выпуклых функций заменить на модифика-

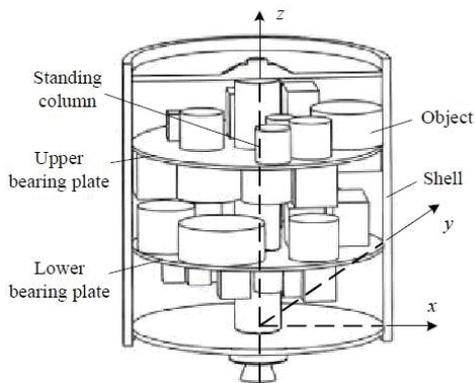
цию с регулировкой шага для почти-дифференцируемых функций.

Параллельный алгоритм можно значительно ускорить, если метод локального спуска снабдить критерием останова, который позволяет прервать процесс минимизации, если его продолжение не приведет к улучшению рекордного значения функции. Такую возможность предоставляют субградиентные методы с преобразованием пространства, которые используют аппроксимацию (сверху) множества экстремумов монотонно уменьшающимися по объему эллипсоидами [9]. Параллельные алгоритмы на их основе являются перспективными для создания эффективного программного обеспечения для решения многоэкстремальных задач оптимизации.

Все сказанное будет справедливо и для более сложных задач упаковки круговых цилиндров [24, 25, 26]. Так, например, если ввести еще координату z , то легко описывать равновесные модели для размещения круговых цилиндров, которые должны крепиться к разным полкам (см. рисунок 7). Если заменить критерий "минимального радиуса" на критерий близости центра тяжести к заданной точке в евклидовой или манхэттенской метрике, то это даст возможность приблизить описание модели к практике.



a. the international communication satellite



b. the simplified satellite module

Рис. 7. Упрощенный спутниковый модуль [2]

Самостоятельный интерес представляют двойственные оценки для квадратичных экстремальных задач. С их помощью можно обосновать оптимальность решения в многоэкстремальной задаче равновесной упа-

ковки кругов. Учитывая, что задачи об упаковке кругов тесно связаны с графовыми задачами о помехоустойчивых кодах, или, другими словами, кодами, корректирующими ошибки при передаче информации, то для оценок можно привлекать числа Ловаса и оценки Шора для числа устойчивости графа.

Литература

1. КОВАЛЕНКО А.А., ПАНКРАТОВ А.В., РОМАНОВА Т.Е., СТЕЦЮК П.И. Упаковка круговых цилиндров в цилиндрический контейнер с учетом специальных ограничений поведения системы // Журнал обчислювальної та прикладної математики. – 2013. – № 1 (111). – С. 126–134.
2. CHE C., WANG Y., TENG H. Test problems for quasi-satellite packing: Cylinders packing with behavior constraints and all the optimal solutions known. // Optimization Online. – 2008. – 11 p. http://www.optimization-online.org/DB_HTML/2008/09/2093.html
3. JINGFA L., GANG L. Basin filling algorithm for the circular packing problem with equilibrium behavioral constraints // SCIENCE CHINA Information Sciences. – 2010. – 53 (5). – P. 885–895.
4. OLIVEIRA W.A., MORETTI A.C., SALLES NETO L.L. A heuristic for the nonidentical circle packing problem // Anais do CNMAC. – 2010. – 3. P. 626–632.
5. XU Y.-C., DONG F.-M., LIU Y., XIAO R.-B., AMOS M. Ant Colony Algorithm for the Weighted Item Layout Optimization Problem // ArXiv preprint. – 2010. – 1001. – P. 4099.
6. XU Y.-C., XIAO R.-B., AMOS M. A Novel Algorithm for the Layout Optimization Problem // Proceedings of the 2007 IEEE Congress on Evolutionary Computation (CEC07). IEEE Press. – 2007. P. 3938–3942.
7. НЕНАХОВ Э.И., РОМАНОВА Т.Е., СТЕЦЮК П.И. Равновесная упаковка кругов в круг минимального радиуса // Теорія оптимальних рішень. – Київ: Ін-т кібернетики ім. В.М. Глушкова НАН України, 2013. – С. 143–153.
8. ROMANOVA T., STETSYUK P., KOVALENKO A. Optimal balanced circular packing problem // Proceedings of abstracts of reports presented

- at the V International Conference on Optimization Methods and Applications "Optimization and applications" (OPTIMA-2014) held in Petrovac, Montenegro, September 28 – October 4, 2014. – P. 158–159.
9. СТЕЦЮК П.И. Методы эллипсоидов и r -алгоритмы. – Эврика: Кишинэу, 2014. – 488 с.
 10. СТЕЦЮК П.І., РОМАНОВА Т.Є., КОВАЛЕНКО Г.А. Комп'ютерна програма "Рівноважна упаковка кругів" // Свідоцтво про реєстрацію авторського права на твір № 56609. Україна. Міністерство освіти і науки. Державний департамент інтелектуальної власності. – Дата реєстрації 23.09.2014.
 11. Octave [Электронный ресурс]: <http://www.octave.org> – Режим доступа: свободный.
 12. STETSYUK P., ROMANOVA T., SCHEITHAUER G. On the global minimum in a balanced circular packing problem // Optimization Letters. – 2016, № 10, P. 1347–1360.
 13. ЛИХОВИД А.П. О реализации параллельного алгоритма для решения задач равновесной упаковки // Теорія оптимальних рішень. – Київ: Ін-т кібернетики ім. В.М. Глушкова НАН України, 2015. – С. 154–159.
 14. СТЕЦЮК П.І., ЛИХОВИД О.П. Комп'ютерна програма "A parallel algorithm for a balanced circular packing problem" // Свідоцтво про реєстрацію авторського права на твір № 62184. Україна. Міністерство освіти і науки. Державний департамент інтелектуальної власності. – Дата реєстрації 20.10.2015.
 15. Кластерный комплекс Института кибернетики. Кластерный комплекс СКИТ. <https://icybcluster.org.ua/>
 16. SHOR N.Z. Nondifferentiable Optimization and Polynomial Problems. – Boston/Dordrecht/London, Kluwer Academic Publishers, 1998. – 412 p.
 17. СТЕЦЮК П.И., РОМАНОВА Т.Е., ШАЙТХАУЭР Г. О глобальном минимуме целевой функции в задаче равновесной упаковки кругов // Доповіді НАН України. – 2014. – № 6. – С. 53–57.
 18. SHOR N.Z., STETSYUK P.I. Dual Solution of Quadratic-Type Problems by r -algorithm (subroutine DSQTP r) // Abstracts of Second International Workshop "Recent Advances in Non-Differentiable Optimization" (October 1–4, 2001, Kyiv, Ukraine). – P. 36.

19. ПШЕНИЧНЫЙ Б.Н., СОБОЛЕНКО Л.А. Метод линеаризации для обратно-выпуклого программирования // Кибернетика и системный анализ. – 1995. – № 6. – С. 86–97.
20. FASANO G., PINTER J.D., EDS. Modeling and Optimization in Space Engineering. Springer Optimization and Its Applications. – Springer. – New York. – 2012. – 404 p.
21. STOYAN YU., ROMANOVA T. Mathematical Models of Placement Optimisation: Two- and Three-Dimensional // In book "Modeling and Optimization in Space Engineering. Springer Optimization and Its Applications" / G. Fasano, J.D. Pinter, eds. – Springer. – New York. – 2012. – P. 363–388.
22. Гуляницкий Л.Ф., Сиренко С.И. Метаэвристический метод комбинаторной оптимизации ОМК-Н // Проблемы управления и информатики. – 2010. – № 4. – С. 31–42.
23. СЕРГИЕНКО И.В., ШИЛО В.П. Задачи дискретной оптимизации: проблемы, методы решения, исследования. К.: Наук. думка, 2003. – 262 с.
24. КОВАЛЕНКО А.А., ПАНКРАТОВ А.В., РОМАНОВА Т.Е., СТЕЦЮК П.И. Балансная компоновка цилиндрических объектов: математические модели и методы решения // Радиоэлектроника и информатика. – 2014. – № 4. – С. 70–79.
25. STOYAN YU., STETSYUK P., ROMANOVA T. Optimal Balanced Packing Using Phi-Function Technique // In S. Butenko, E. L. Pasiliao, and V. Shylo (Editors), Examining Robustness and Vulnerability of Networked Systems, P. 251–271. – IOS Press, 2014. <http://ebooks.iospress.nl/volumearticle/36433>
26. КОВАЛЕНКО А.А., РОМАНОВА Т.Е., СТЕЦЮК П.И. Задача балансной компоновки 3D-объектов: математическая модель и методы решения // Кибернетика и системный анализ. – 2015. – № 4. – С. 71–81.

Список научных трудов (Стецюк, 1982–2018)

1982–1989 годы

1. Юн Г.Н., Андрусенко С.К., Стецюк П.И., Дрюченков В.С. Опыт применения экономико-математического моделирования при обосновании уровня лимитных цен на новые воздушные суда ГА // В сб. "Проблемы исследования перспектив гражданской авиации" – М: ГосНИИ ГА, 1982.
2. Нурминский Е.А., Андрусенко С.К., Стецюк П.И. О новом полиномиальном алгоритме линейного программирования // Кибернетика. – 1985. – № 4. – С. 118–120.
3. Андрусенко С.К., Ивличев В.П., Стецюк П.И. и др. „ПЕ-ГАС“ – проблемно-ориентированная система перспективного планирования // IV Всесоюзная конференция „Диалог человек–ЭВМ“. Тезисы докладов. – Киев, 1985.
4. Нурминский Е.А., Стецюк П.И. Вопросы реализации полиномиального алгоритма ЛП // Девятый всесоюзный симпозиум "Системы программного обеспечения решения задач оптимального планирования" (г. Минск, 23 февраля – 3 марта 1986 г.). Краткие тезисы докладов. – М.: ЦЭМИ АН СССР, 1986. – С. 158–159.
5. Андрусенко С.К., Нурминский Е.А., Стецюк П.И. Численные эксперименты с новым классом алгоритмов в линейном программировании // Журнал вычислительной математики и математической физики. – 1987. – Т. 27, № 3. – С. 349–356.
6. Стецюк П.И. Простые границы в полиномиальном алгоритме ЛП // Системы программного обеспечения решения задач оптимального планирования. 10-й Всесоюзн. симп. (г. Нарва-Йыэссуу). Краткие тез. докл. – М: ЦЭМИ АН СССР, 1988.
7. Стецюк П.И. Об одном подходе к учету простых границ в полиномиальном алгоритме ЛП // Методы решения экстремальных задач и смежные вопросы. – Киев: Ин-т кибернетики им. В.М.Глушкова АН УССР, 1989. – С. 60–65.

8. Ивличев В.П., Лефтеров А.В., Стецюк П.И. Об опыте создания системы поддержки и принятия решений "АНАЛИЗ" // Диалоговая оптимизация планово-управленческих решений и проблемы внедрения ее в практику. Тезисы докладов научно-практического семинара, 3–5 октября 1989 г. – Киев: Общество "Знание" Украинской ССР, 1989. – С. 32–34.

1990–1994 годы

9. ХАВАТОВ Р.Ш., ЖУРБЕНКО Н.Г., МИЩУК С.А., СТЕЦЮК П.И. Методика оптимизации состава МТП // Оптимизация машинно-тракторного парка: Сб. науч. тр. – М.: МСХА, 1990. – С. 10–17.
10. СТЕЦЮК П.И. Об одном алгоритме ЛП // Краткие тезисы докладов. Одиннадцатая всесоюзная школа "Системы программного обеспечения решения задач оптимального планирования" (г. Кострома, 21–29 мая 1990 г.). – Москва, 1990. – С. 77–78.
11. ЖУРБЕНКО Н.Г., СТЕЦЮК П.И. Агрегирование и декомпозиция для одного класса динамических распределительных задач // Исследования методов решения экстремальных задач. – Киев: Ин-т кибернетики им. В.М. Глушкова АН УССР, 1990. – С. 4–9.
12. МИХАЛЕВИЧ В.С., СЕРГИЕНКО И.В., ШОР Н.З., др., СТЕЦЮК П.И. Пакет прикладных программ для решения на ЕС ЭВМ в диалоговом режиме задач нелинейной и дискретной оптимизации (ППП ДИСНЕЛ) // АН УССР, Ин-т кибернетики им. В.М. Глушкова. – Киев, 1990. – Рук. деп. в ГосФАП СССР, ВНИЦентр, Инв. № 029.10 019092, 1991. – 449 с.
13. СТЕЦЮК П.И. Об одном методе решения блочной задачи линейного программирования // Моделирование и оптимизация. – Киев: Ин-т кибернетики им. В.М. Глушкова АН УССР, 1991. – С. 35–40.
14. ЖУРБЕНКО Н.Г., СТЕЦЮК П.И., ДОЛГОВ В.В., ТИХОНОВ О.А. Об одной задаче оптимизации многослойных структур в плоских защитах // Методы решения задач нелинейного и дискретного программирования. – Киев: Ин-т кибернетики им. В.М. Глушкова АН УССР, 1991. – С. 36–41.
15. ЖУРБЕНКО Н.Г., САИМБЕТОВ Д.Х., СТЕЦЮК П.И., ТОРЕБАЕВ О. Задача перспективного планирования структуры машинно-тракторного парка // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 1992. – С. 58–63.

16. ЛАПТИН Ю.П., МЕДВЕДЕВ В.А., ЖУРБЕНКО Н.Г., СТЕЦЮК П.И., КУЗЬМИН А.И. Задача оптимизации многоступенчатых конвективных пароперегревателей // Информационные технологии в задачах управления сложными системами. – Киев: Ин-т кибернетики им. В.М. Глушкова АН Украины, 1992. – С. 39–47.
17. ЖУРБЕНКО Н.Г., ЛАПТИН Ю.П., МЕДВЕДЕВ В.А., СТЕЦЮК П.И., КУЗЬМИН А.И. Об одной задаче оптимизации теплоэнергетических установок // Теория и вычисл. проблемы оптимизации. – Киев: Ин-т кибернетики им. В.М. Глушкова АН Украины, 1993. – С. 36–39.
18. ШОР Н.З., СТЕЦЮК П.И. Кусочно-вогнутая задача рюкзачного типа // Методы исследования экстремальных задач. – Киев: Ин-т кибернетики им. В.М. Глушкова АН Украины, 1994. – С. 21–28.

1995 год

19. СТЕЦЮК П.И. К вопросу сходимости r -алгоритмов // Кибернетика и систем. анализ. – 1995. – № 6. – С. 173–177.
20. СТЕЦЮК П.И. К обоснованию сходимости алгоритмов с растяжением пространства // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 1995. – С. 4–8.
21. СТЕЦЮК П.И. Об одной схеме методов отсечений – Киев, 1995. – 34с. – (Препр. / НАН Украины. Ин-т кибернетики им. В.М.Глушкова: 95-25).

1996 год

22. СТЕЦЮК П.И. r -алгоритмы и эллипсоиды // Кибернетика и систем. анализ. – 1996. – № 1. – С. 113-134.
23. СТЕЦЮК П.И. Классические фейеровские методы с преобразованием пространства // Методы решения экстремальных задач. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 1996. – С. 3–9.
24. СТЕЦЮК П.И. Квазиньютоновские методы и r -алгоритмы – Киев, 1996. – 21 с. – (Препр. / НАН Украины. Ин-т кибернетики им. В.М. Глушкова: 96-10).

25. СТЕЦЮК П.І. Субградієнтні методи з перетворенням простору для мінімізації негладких опуклих функцій // Автореферат дисертації на здобуття наукового ступеня кандидата фізико-математичних наук. Ін-т кібернетики ім. В.М. Глушкова НАН України. – Київ, 1996. – 16 с.
26. СТЕЦЮК П.И. Субградиентные методы с преобразованием пространства для минимизации негладких выпуклых функций // Рукопись диссертации на соискание научной степени кандидата физико-математических наук. – Киев, 1996. – 135 с.
27. СТЕЦЮК П.И. Метод центров тяжести простых тел // Кибернетика и систем. анализ. – 1996. – № 5. – С. 117–138.

1997 год

28. СТЕЦЮК П.И. Ортогонализующие линейные операторы в выпуклом программировании (Часть I) // Кибернетика и системный анализ. – 1997. – № 3. – С. 97–119.
29. СТЕЦЮК П.И. Ортогонализующие линейные операторы в выпуклом программировании (Часть II) // Кибернетика и системный анализ. – 1997. – № 5. – С. 111–124.
30. СТЕЦЮК П.И. Об одном способе ускорения метода ортогонального субградиентного спуска // Оптимизация и ее приложения. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 1997. – С. 42–49.
31. ШОР Н.З., СТЕЦЮК П.И. Использование модификации r -алгоритма для нахождения глобального минимума полиномиальных функций // Кибернетика и систем. анализ. – 1997. – № 4. – С. 28–49.
32. ШОР Н.З., СТЕЦЮК П.И. Использование r -алгоритма в задачах полуопределенного программирования // Праці міжнародної конференції "Питання оптимізації обчислень" (6–8 жовтня 1997 р., м. Київ). – С. 330–335.
33. BARDADYM T.A., SHOR N.Z., КНОПОВ P.S., ЛІКHOVID A.P., СТЕЦЮК P.I. On the Use of Nonsmooth Optimization Methods in Statistics and Stochastic Programming // Thesis of 16-th International Symposium on Mathematical Programming, Lausanne, 1997.

1998 год

34. СТЕЦЮК П.И. Линейные операторы в квазиньютоновских методах // Теория и приложения методов оптимизации. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 1998. – С. 3–8.
35. СЕРГИЕНКО И.В., МИХАЛЕВИЧ М.В., СТЕЦЮК П.И., КОШЛАЙ Л.Б. Межотраслевая модель планирования структурно-технологических изменений // Кибернетика и систем. анализ. – 1998. – № 3. – С. 3–17.
36. BARDADYM T.A., SHOR N.Z., ZHURBENKO N.G., LIKHOVID A.P., STETSYUK P.I. Solver for Linear Two-Stage Fixed Recourse Problems // Thesis of OR98 International Conference on Operations Research, Zurich, 1998. – P. 65–66.

1999 год

37. СТЕЦЮК П.И. К методам эллипсоидов // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 1999. – С. 27–33.
38. 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.
39. ШОР Н.З., СТЕЦЮК П.И., КРЫЛОВ С.В. Программы нахождения оптимальных по объему эллипсоидов, описанных вокруг набора точек // Thesis of International Conference "Dynamical Systems Modelling and Stability Investigation" (May 25-29, 1999, Kyiv). – P. 62.

2000 год

40. СТЕЦЮК П.И. Об одном методе для нахождения допустимой точки выпуклого неравенства // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2000. – С. 3–10.
41. STETSYUK P. Method FEG2P1 to Find ε -Solutions of Convex Inequalities // Thesis of U.S.-Ukrainian Workshop "Recent Advances in Non-Differentiable Optimization" (May 15-18, 2000, Kyiv, Ukraine). – P. 36.

42. СТЕЦЮК П.И., КОЛЕСНИК Ю.Г. К вопросу выбора метода аппроксимации результатов измерений // Интеллектуальные информационно-аналитические системы и комплексы. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2000. – С. 62–67.
43. СТЕЦЮК П.И., КОЛЕСНИК Ю.Г., ПЕПЕЛЯЕВ В.А. Об оценке коэффициентов производственной функции Кобба-Дугласа при неокласических критериях // Теорія оптимальних рішень. Моделювання та керування в умовах невизначеності. – Київ: Ін-т кібернетики ім. В.М. Глушкова НАН України, 2000. – С. 14–23.
44. ШОР Н.З., СТЕЦЮК П.И., КРЫЛОВ С.В. Нахождение глобальных минимумов полиномиальных функций с использованием двойственных квадратичных оценок // Вестник Международного Соломонова университета. – 2000. – № 4. – С.217–233.
45. СЕРГИЕНКО И.В., ШИЛО В.П., СТЕЦЮК П.И. Приближенные алгоритмы для решения задачи нахождения максимально независимого множества // Компьютерная математика. – Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 2000. – С. 4–20.

2001 год

46. СТЕЦЮК П.И. К методам решения плохообусловленных систем линейных уравнений // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2001. – С. 9–15.
47. СТЕЦЮК П.И. Об ускорении сходимости методов эллипсоидов // Труды XII Байкальской международной конференции (Иркутск, Байкал, 24 июня – 1 июля 2001 г.). Том 1. Математическое программирование. – Иркутск, 2001. – С. 61–66.
48. PARDALOS P., BUTENKO S., SERGIENKO I.V., SHYLO V.P., STETSYUK P.I. Approximate and Exact Algorithms for Determination of a Maximum Independent Set // Thesis of WISTCIS Kick-Off Conference "Teleworking in Research, Medicine and Business"(April 19-21, 2001, Kiev, Ukraine). – P. 18.
49. СТЕЦЮК П.И. Линейная модель для нахождения пропускных способностей компонент надежной сети // Збірник наукових праць „Компьютерна математика. Оптимізація обчислень.“ – Київ: Ін-т кібернетики ім. В.М.Глушкова НАН України, 2001. – Т. 1. – С. 376–384.

50. 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“ (October 1-4, 2001, Kyiv, Ukraine). – P. 36.
51. SHYLO V.P. AND STETSYUK P.I. Using GES and r -algorithm for Solving Multidimensional Knapsak Problems // Abstracts of Second International Workshop „Recent Advances in Non-Differentiable Optimization“ (October 1-4, 2001, Kyiv, Ukraine). – P. 38.
52. BUTENKO S., PARDALOS P.M., SERGIENKO I.V., SHYLO V., STETSYUK P. Approximate and Exact Algorithms for the Maximum Independent Set // "MaxClique '01" Conference, University of Klagenfurt, Austria, May 31 – June 2, 2001.

2002 год

53. SHOR N.Z. AND STETSYUK P.I. Constructing Utility Functions by Methods of Nondifferentiable Optimization // in: A.S. Tangian, AND J. Gruber (Eds.) *Constructing and Applying Objective Functions*, Lecture Notes in Economics and Mathematical Systems. – Vol. 510. – Berlin.: Springer-Verlag. – 2002. – P. 215–232.
54. 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.
55. 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). – К.: НТТУ "КПІ". – 2002. – С. 399.
56. STETSYUK P.I. On one interesting ellipsoid // Международная конференция, посвященная 65-летию со дня рождения Б.Н.Пшеничного. Тезисы докладов. – Киев: Издательство НТТУ "КПИ", 2002. – С. 53-54.
57. СТЕЦЮК П.І. Про одну модифікацію методу еліпсоїдів // Міжнародна школа-семінар "Теорія прийняття рішень", Ужгород, 7–12 жовтня 2002. Праці школи семінару. – С. 64.

58. СТЕЦЮК П.И., БУХАНЦОВ Д.М. К ускорению метода эллипсоидов с помощью использования шарового слоя // Збірка тез міжнародної конференції "Обчислювальна та прикладна математика", присвяченої 80-річчю академіка І.І. Ляшка (9.09.–10.09.2002) – Київ, 2002 – С. 94.
59. СТЕЦЮК П.И., БУХАНЦОВ Д.М. К ускорению метода эллипсоидов с помощью использования шарового слоя // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2002. – С. 63–70.
60. BUTENKO S., PARDALOS P.M., SERGIENKO I.V., SHYLO V., STETSYUK P. Finding Maximum Independent Sets in Graphs Arising from Coding Theory // Proceedings of the 17th Symposium on Applied Computing. – ACM Press, 2002. – P. 542–546.
61. СТЕЦЮК П.И., КОЛЕСНИК Ю.С., ЛЕЙБОВИЧ М.М. О робастности метода наименьших модулей // Компьютерная математика. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2002. – Выпуск 2. – С. 114–123.

2003 год

62. СТЕЦЮК П.И. Модификация метода эллипсоидов // Информационный бюллетень Ассоциации математического программирования. №10. Научное издание. – Екатеринбург: УрО РАН, 2003. – С. 216–217.
63. СТЕЦЮК П.И. Приближенный метод эллипсоидов // Кибернетика и системный анализ. – 2003. – № 3. – С. 141–146.
64. СТЕЦЮК П.И. Об одном эллипсоиде для внешней аппроксимации n -мерного полушара // Компьютерная математика. – Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 2003. – Выпуск 2. – С. 144–151.
65. СТЕЦЮК П.И., БУХАНЦОВ Д.М. Нахождение квадратичных функций Ляпунова для наилучших оценок решения линейной системы // Міжнародна наукова конференція "ШОСТИ БОГОЛЮБОВСЬКІ ЧИТАННЯ", м. Чернівці, Україна, 26-30 серпня 2003 р.

66. СТЕЦЮК П.И., КОЛЕСНИК Ю.С., БЕРЕЗОВСКИЙ О.А. Об одном методе нахождения L_p -решения системы линейных уравнений // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 2003. – С. 83–90.
67. 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.
68. МІЦА О.В., СТЕЦЮК П.І. Задача знаходження оптимальних параметрів однорідного оптичного покриття // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 2003. – С. 127–134.

2004 год

69. БАБАК О.В., ГАСАНОВ А.С., ЛЕЙБОВИЧ М.М., СТЕЦЮК П.И. Об одном подходе к использованию метода наименьших модулей при построении линейных моделей // Кибернетика и системный анализ. – 2004. – №3. – С. 109–115.
70. GOLDENGORIN B., KUZ'MENKO V., STETSYUK P., TSO M. Pricing by an Allocation Model with Different Types of Discount // Book of Abstracts, (CO 2004, 28–31 March 2004). – Lancaster University, Lancaster, UK. – P. 31.
71. ШОР Н.З., СТЕЦЮК П.И., БЕРЕЗОВСКИЙ О.А. Двойственные оценки для специальной оптимизационной задачи квадратичного типа на многообразии Штиффеля // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2004. – С. 3–10.
72. ШОР Н.З., СТЕЦЮК П.И., БЕРЕЗОВСКИЙ О.А. Нахождение оценок глобального экстремума в задачах на многообразии Штиффеля // Автоматика-2004. Материалы 11-й международной конференции по автоматическому управлению, 27–30 сентября 2004, Национальный университет пищевых технологий. – Киев, 2004. – Т. 1. – С. 47.
73. КУЗЬМЕНКО В.Н., ГОЛЬДЕНГОРИН Б.И., ТСО М., СТЕЦЮК П.И. Сравнение двух субградиентных методов при нахождении оценок

для задач размещения // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 2004. – С. 111–119.

74. СТЕЦЮК П.І., МИЦА О.В. До використання методів першого порядку в задачах оптимізації багат шарових оптичних плівок // II-а Міжнародна школа-семинар "Теорія прийняття рішень", Ужгород, 27 вересня – 2 жовтня 2004 р. Праці школи-семинару. – С. 88.

2005 год

75. СТЕЦЮК П.И., МИЦА А.В. О задачах оптимизации параметров для многослойных оптических покрытий // Кибернетика и системный анализ. – 2005. – № 4. – С. 107–115.
76. СТЕЦЮК П.И., МИЦА А.В. О вычислении градиента в задаче синтеза оптических покрытий // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2005. – С. 127–133.
77. SHOR N. AND 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, February 2-4, 2005, University of Florida, Gainesville, FL.
78. СТЕЦЮК П.И. О функционально избыточных ограничениях для булевых оптимизационных задач квадратичного типа // Кибернетика и системный анализ. – 2005. – № 6. – С. 168–172.
79. Задачі оптимального проектування надійних мереж / Шор Н.З., Сергієнко І.В., Шило В.П., Стецюк П.І., Парасюк І.М., Лебедева Т.Т., Лаптінін Ю.П., Журбенко М.Г., Бардадим Т.О., Шаріфов Ф.А., Лиховид О.П., Березовський О.А., Мірошніченко В.М. – Київ, Наукова думка, 2005. – 230 с.

2006 год

80. СТЕЦЮК П.И. Новые модели квадратичного типа для задачи о максимальном взвешенном разрезе графа // Кибернетика и системный анализ. – 2006. – №1. – С. 63–75.

81. СТЕЦЮК П.И., БЕРЕЗОВСКИЙ О.А. Лагранжевая оценка для максимального разреза графа с заданными количествами вершин в обоих подмножествах разбиения // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2006. – С. 31-38.
82. СТЕЦЮК П.И., ПАРДАЛОС П.М. Об уточнении лагранжевых двойственных оценок в бинарных и булевых квадратичных задачах // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 2006. – С. 145–153.
83. STETSYUK P.I., BARDADYUM T.O. The Use of Odd Sets of Variables to Improve Bounds in Boolean Quadratic Problems // Abstracts of International Conference in Honor of the 70th Birthday of Academician I.V.Sergienko "Applied Optimization and Metaheuristic Innovations"(July 19-21, 2006, Yalta, Ukraine)
84. МІЦА О.В., СТЕЦЮК П.І. Дослідження оптимізації спектральних характеристик семишарового оптичного покриття з чергуючими шарами // III-а Міжнародна школа-семинар "Теорія прийняття рішень", Ужгород, 2 – 7 жовтня 2006 р. Праці школи-семинару. – С. 73–74.
85. П.И.СТЕЦЮК П.И., ПАРДАЛОС П.М., КРОШКО Д.Л. О новых лагранжевых двойственных оценках для числа устойчивости графа // Компьютерная математика. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2006. – Выпуск 3. – С. 149–158.

2007 год

86. СТЕЦЮК П.И. Об уточнении верхних оценок Н.З.Шора для числа независимости неориентированного графа // Информационный бюллетень Ассоциации математического программирования. № 11. Научное издание. – Екатеринбург: УрО РАН, 2007. – С. 80–81.
87. СТЕЦЮК П.И., ЧУМАКОВ Б.М. О свойствах одной верхней оценки Н.З.Шора для взвешенного числа устойчивости графа // Праці міжнародного симпозіума "Питання оптимізації обчислень (ПОО–XXXIII)". – К.: Ін-т кибернетики ім. В.М.Глушкова НАН України, 2007. – С. 271-272.

88. СТЕЦЮК П.И., БУТЕНКО С.И., БЕРЕЗОВСКИЙ О.А. Об одной верхней оценке для взвешенного числа устойчивости графа // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 2007. – № 6. – С. 80–89.

2008 год

89. БЕРЕЗОВСКИЙ О.А., СТЕЦЮК П.И. Об одном способе нахождения двойственных квадратичных оценок Шора // Кибернетика и системный анализ. – 2008. – № 2. – С. 89–99.
90. СТЕЦЮК П.И. Оценка Шора для взвешенного числа устойчивости W_p -совершенных графов // Материалы международной научной конференции "Математическое моделирование, оптимизация и информационные технологии", 19–21 марта 2008, Кишинэу, Академия транспорта, информатики и коммуникаций. – С. 112-121.
91. СТЕЦЮК П.И. О новых свойствах оценок Шора для взвешенного числа устойчивости графа // Праці міжнародної конференції "50 років Інституту кібернетики ім. В.М. Глушкова НАН України". – К.: Інститут кібернетики імені В.М. Глушкова НАН України. – 2008. – С. 164–173.
92. СТЕЦЮК П.И., БУТЕНКО С.И., ЛИХОВИД А.П. LP-ориентированная верхняя оценка для числа устойчивости графа на основе p -колес // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2008. – № 7. – С. 34–44.
93. STETSYUK P. On interesting properties of Shor's upper bounds for weighted stability number of a graph // Abstracts of International Conference "Discrete and Global Optimization"(July 31 – August 2, 2008, Yalta, Ukraine).
94. ЛУКНОВУД О., STETSYUK P. On numerical experiments with two LP-oriented upper bounds for the stability number of a graph // Abstracts of International Conference "Discrete and Global Optimization"(July 31 – August 2, 2008, Yalta, Ukraine).
95. СТЕЦЮК П.И., ПИЛИПОВСКИЙ А.В. Математическая модель оптимальной загрузки мощностей энергосистемы с учетом их маневренности // Праці IV міжнародної школи-семінару "Теорія прийняття рішень". – Ужгород: УжНУ, 2008. – С. 159.

2009 год

96. СТЕЦЮК П.И., ЛИХОВИД А.П. Об ЛП-ориентированных верхних оценках для взвешенного числа устойчивости графа // Кибернетика и системный анализ. – 2009. – № 1. – С. 157–170.
97. СЕРГИЕНКО И.В., МИХАЛЕВИЧ М.В., СТЕЦЮК П.И., КОШЛАЙ Л.Б. Модели и информационные технологии для поддержки принятия решений при проведении структурно-технологических преобразований // Кибернетика и систем. анализ. – 2009. – № 2. – С. 26–49.
98. СТЕЦЮК П.И., МИХАЛЕВИЧ М.В., ПИЛИПОВСКИЙ А.В. Применение методов негладкой оптимизации для планирования структурно-технологических изменений // Матеріали ХІ міжнародної науково-технічної конференції "Системний аналіз та інформаційні технології", 26-30 травня 2009 р. – НТТУ КПІ, Київ. – С. 137.
99. Методи негладкої оптимізації у спеціальних задачах класифікації / П.І. Стецюк, О.А. Березовський, М.Г. Журбенко, Д.О. Кропотов. – Київ, 2009. – 28 с. – (Препр./НАН України. Ін-т кібернетики ім.В.М.Глушкова; 2009–1).
100. ЖУРБЕНКО М.Г., СТЕЦЮК П.І. Субградієнтні методи змінної метрики для розв'язування яружних задач оптимізації – Київ, 2009. – 27 с. – (Препр. / НАН України. Ін-т кібернетики імені В.М.Глушкова; 2009-3).
101. СТЕЦЮК П.И., ЛИХОВИД А.П., ПИЛИПОВСКИЙ А.В. О решении одного класса оптимизационных задач нахождения нагрузок энергетических объектов с помощью NEOS-программ // Праці міжнародного симпозиуму "Питання оптимізації обчислень (ПОО-XXXV)". – Київ: Ін-т кібернетики ім. В.М. Глушкова НАН України, 2009. – Т.2. – С. 350–354.
102. СТЕЦЮК П.И. Субградиентные методы переменной метрики, использующие шаг Агмона-Моцкина и одноранговый эллипсоидальный оператор // Труды АТИК – 2007-2008. – Кишинэу: Эврика, 2009. – Т. I (XII). – С. 16–25.
103. СТЕЦЮК П.И., ЛИХОВИД А.П., ПИЛИПОВСКИЙ А.В. Задачи оптимизации для выбора электрических нагрузок в энергосистеме

// Теорія оптимальних рішень. – Київ: Ін-т кібернетики ім. В.М. Глушкова НАН України, 2009. – № 8. – С. 136–141.

104. BUTENKO S., PARDALOS P.M., SERGIENKO I.V., SHYLO V., STETSYUK P. Estimating the size of Correcting Codes using Extremal Graph Problems // In C. Pearce and E. Hunt, editors, Optimization: Structure and Applications. – Springer, 2009. – P. 227–243.

2010 год

105. СТЕЦЮК П.И., НУРМИНСКИЙ Е.А. Негладкий штраф и субградиентные алгоритмы для решения задачи проекции на политоп // Кибернетика и системный анализ. – 2010. – №1. – С. 59-63.
106. СТЕЦЮК П.И. Оценки Шора в квадратичных экстремальных задачах // Материалы международной научной конференции "Математическое моделирование, оптимизация и информационные технологии", 24–26 марта 2010, Кишинэу, Академия транспорта, информатики и коммуникаций. – С. 19–25.
107. STETSYUK P., NURMINSKI E. On a modification of ellipsoid method // Материалы международной научной конференции "Математическое моделирование, оптимизация и информационные технологии", 24–26 марта 2010, Кишинэу, Академия транспорта, информатики и коммуникаций. – С. 37-42.
108. МИЦА А.В., СТЕЦЮК П.И. Использование г-алгоритма для оптимизации трех оптических многослойных покрытий // Материалы международной научной конференции "Математическое моделирование, оптимизация и информационные технологии", 24–26 марта 2010, Кишинэу, Академия транспорта, информатики и коммуникаций. – С. 100–108.
109. СТЕЦЮК П.И., КОШЛАЙ Л.Б., ПИЛИПОВСКИЙ А.В. О задаче оптимального соотношения между спросом и добавленной стоимостью в моделях Леонтьева // Теорія оптимальних рішень. – 2010. – № 9. – С. 136–143.
110. СТЕЦЮК П.И., КОШЛАЙ Л.Б. Оптимальная нормированная структура спроса и добавленной стоимости в продуктивной модели Леонтьева // Кибернетика и систем. анализ. – 2010. – № 5. – С. 51–59.

111. СТЕЦЮК П.И., ЛЯШКО В.И., НУРМИНСКИЙ Е.А. Точная ЛП-оценка для взвешенного числа устойчивости t -совершенных графов // Журнал обчислювальної та прикладної математики. – 2009. – № 3(99). – С. 106–115.
112. СТЕЦЮК П.И., ЗОЛОТЫХ Н.Ю. Бинарный квадратичный многогранник и его аппроксимации // Журнал обчислювальної та прикладної математики. – 2010. – № 2(101). – С. 139–149.
113. СТЕЦЮК П.И. Оценки Н.З.Шора в квадратичных экстремальных задачах и их применение в комбинаторной оптимизации // Российская конференция "Дискретная оптимизация и исследование операций": Материалы конференции (Алтай, 27 июня – 3 июля 2010). – Новосибирск: Изд-во Ин-та математики, 2010. – С. 60–64.
114. БЕРЕЗОВСКИЙ О.А., СТЕЦЮК П.И. Задачи негладкой безусловной оптимизации для линейного и квадратичного классификаторов с максимальным зазором // V-а Міжнародна школа-семинар "Теорія прийняття рішень", Ужгород, 27 вересня – 1 жовтня 2010. Праці школи-семинару. – С. 15–16.
115. ЗОЛОТЫХ Н.Ю., СТЕЦЮК П.И. О внешней аппроксимации бинарного и булевого квадратичных многогранников // V-а Міжнародна школа-семинар "Теорія прийняття рішень", Ужгород, 27 вересня – 1 жовтня 2010. Праці школи-семинару. – С. 103–104.
116. KOSHLAI L., STETSYUK P. Application of Nonlinear Optimization to the Problems of Structural Changes during Transition // 8th EUROPT Workshop on Advances in Continuous Optimization, Aveiro, Portugal, July 9-10, 2010. – P. 28.
117. STETSYUK P., KOSHLAI L., PYLYPOVSKYI O. Decision Support System for Planning Structural and Technological Changes // 8th EUROPT Workshop on Advances in Continuous Optimization, Aveiro, Portugal, July 9-10, 2010. – P. 29.
118. П.И.СТЕЦЮК, Л.Б.КОШЛАЙ Об одной экстремальной задаче для связи прямой и двойственной моделей Леонтьева // Крымская осенняя математическая школа симпозиум (КРОМШ), 18-29 сентября 2010 года. – Севастополь: КНЦ НАНУ, 2010. – С. 56.
119. СТЕЦЮК П.І. Алгоритми недиференційовної оптимізації та їх застосування // В монографії „Стан та перспективи розвитку інформатики в Україні“ – Київ: Наукова думка, 2010. – С.183–208.

2011 год

120. СТЕЦЮК П.И. Метод amsq2p для овражных выпуклых функций // Информационный бюллетень Ассоциации математического программирования. № 12. Научное издание. – Екатеринбург: УрО РАН, 2011. – С. 57–58.
121. СТЕЦЮК П.И. Релаксационный субградиентный метод минимизации овражных выпуклых функций // Проблемы теоретической кибернетики. Материалы XVI Международной конференции (Нижний Новгород, 20–25 июня 2011 г.) / Под ред. Ю.И. Журавлева. – Нижний Новгород: Изд-во Нижегородского госуниверситета, 2011. – С. 449–453.
122. СТЕЦЮК П.И. Субградиентные методы с преобразованием пространства для минимизации овражных выпуклых функций // Международная конференция „Современные проблемы прикладной математики и механики: теория, эксперимент и практика“, посвященная 90-летию со дня рождения академика Н.Н. Яненко. Тезисы докладов. – Новосибирск, Академгородок, 2011. – С. 57.
123. СТЕЦЮК П.И. Субградиентные методы с преобразованием пространства для минимизации овражных выпуклых функций // Современные проблемы прикладной математики и механики: теория, эксперимент и практика [Электронный ресурс] / Международная конференция, посвященная 90-летию со дня рождения академика Н.Н. Яненко, Новосибирск, Россия, 30 мая – 4 июня 2011 г., Новосибирск, ИВТ СО РАН, 2011, № гос. регистрации – 0321101160, Режим доступа: <http://conf.nsc.ru/files/conferences/niknik-90/fulltext/37828/53979/Stetsyuk.pdf>, свободный. – Загл. с экрана (дата обращения: 07.11.2011).
124. Віділ А.Ю., Пилиповський О.В., Стецюк П.І. Паралельний алгоритм мултистарту на основі субградієнтного методу для розв'язування системи нелінійних рівнянь // Праці міжнародної молодіжної математичної школи "Питання оптимізації обчислень (ПОО-XXXVII)". – К.: Ін-т кібернетики ім. В.М.Глушкова НАН України, 2011. – С. 28–29.
125. СТЕЦЮК П.И., Бондаренко А.В. О спектральных свойствах модели Леонтьева // Теорія оптимальних рішень. – 2011. – № 10. – С. 84–90.

126. Бондаренко А.В., Кошлай Л.Б., Стецюк П.И. Об одной задаче нахождения оптимальной структуры спроса и добавленной стоимости // Математичні методи, моделі та інформаційні технології в економіці. – Матеріали II Міжнародної науково-методичної конференції. – Чернівці: ДрукАрт, 2011. – С. 26–27.
127. Стецюк П.И., Кошлай Л.Б. Об одной экстремальной задаче для связи прямой и двойственной моделей Леонтьева // Спектральные и эволюционные задачи. – 2011. – Т. 2, № 2. – С. 164–169.
128. П.И.Стецюк, Л.Б.Кошлай Об одной оптимизационной задаче для моделей Леонтьева // XXII ежегодная междунар. конферен. (KROMSH- 2011), (Крым, Севастополь, 18-29 сентября 2011), КНЦ НАНУ, 2011. – С. 51.
129. KOSHLAI L., STETSYUK P. Application of Nonlinear Optimization to the Problems of Economic Reforms during Transition // 82nd Annual Scientific Conference of the International Assosiation of Applied Mathematics and Mechanics, Graz University of Technology, Graz, Austria (April 18–21, 2011). – P. 350.
130. Стецюк П. И. О спектральных свойствах матриц Леонтьева // Статистика. Моделирование. Оптимизация: сборник трудов Всероссийской конференции (Челябинск, 28 ноября – 3 декабря 2011 г.). – Челябинск: Издательский центр ЮУрГУ, 2011. – С. 173–178.
131. NURMINSKI E.A., STETSYUK P.I. Nonsmooth optimization and projection problem // II International Conference "Optimization and Applications (OPTIMA-2011), Abstracts. – Petrovac, Montenegro, September 25 – October 2, 2011. – М.: Вычислительный центр им. А.А. Дородницына РАН, 2011. – С. 167–168.
132. Математичні методи оптимізації та інтелектуальні комп'ютерні технології моделювання складних процесів і систем з урахуванням просторових форм об'єктів/ Грицик В.В., Шевченко А.І., Кісельова О.М., Яковлев С.В., Відюк П.І., Гіль М.І., Крак Ю.В., Романова Т.Є., Куляс А.І., Стецюк П.І. – Донецьк: ІПП "Наука і освіта", 2011. – 480 с.
133. Горбачук В.М., Гаркуша Н.І., Стецюк П.І., Єгоров Ю. До розрахунку транзитних обсягів і ставок // PDMU-2011 (Алупка). – К.: КНУ ім. Т. Шевченка, 2011. – С. 61–62.

2012 год

134. СЕРГИЕНКО И.В., СТЕЦЮК П.И. О трех научных идеях Н.З.Шора // Кибернетика и системный анализ. – 2012. – № 1. – С. 4–22.
135. СТЕЦЮК П.И. О важных датах НДО и ЛП в 2012 году // Материалы международной конференции „Математика, экономика, менеджмент: 100 лет со дня рождения Л. В. Канторовича“. – СПб.: ООО „ТАИС“, 2012. – С. 67–69.
136. СЕРГИЕНКО И.В., СТЕЦЮК П.И. Три центральные идеи Н.З.Шора // Математическое моделирование, оптимизация и информационные технологии: материалы 3-й Междунар. науч. конф. (Кишинэу, 19-23 марта 2012 г.). – Кишинэу: Эврика, 2012. – С. 468–475.
137. СТЕЦЮК П.И. Ускоренные по Шору модификации субградиентного метода Поляка // Математическое моделирование, оптимизация и информационные технологии: материалы 3-й Междунар. науч. конф. (Кишинэу, 19-23 марта 2012 г.). – Кишинэу: Эврика, 2012. – С. 509–519.
138. СТЕЦЮК П.И., БАРДАДЫМ Т.А., БЕРЕЗОВСКИЙ О.А., ЖУРБЕНКО Н.Г., ЧУМАКОВ Б.М., ШОР Е.И. К 75-летию со дня рождения выдающегося ученого – академика НАН Украины Наума Зуселевича Шора // Кибертония. – 2012. – № 1. – С. 44–45.
139. СТЕЦЮК П.И. Ускорение субградиентного метода Поляка // Теорія оптимальних рішень. – 2012. – № 11. – С. 151–160.
140. СТЕЦЮК П.И. О решении системы линейных уравнений с двусторонними ограничениями на переменные // Алгебра и линейная оптимизация. Тезисы международной конференции, посвященной 100-летию С.Н.Черникова. Екатеринбург, 14–19 мая 2012 года. – Екатеринбург: Изд-во „УМЦ-УПИ“, 2012. – С. 155–157.
141. СТЕЦЮК П.І., ЖУРБЕНКО М.Г., ЛИХОВИД О.П. Математичні моделі та програмне забезпечення в задачах енергетики. – К.: ПП "Ательє "Поліграфічний комплекс", 2012. – 64 с.
142. СТЕЦЮК П.И. Shor's acceleration for Polyak's subgradient method (in Russian)² // Конструктивный негладкий анализ и смежные

²Ускоренный по Шору субградиентный метод Поляка

- вопросы. Тезисы докладов международной конференции. – СПб.: Изд-во Санкт-Петербургского университета, 2012. – P. 157–160.
143. STETSYUK P. Polyak's accelerated subgradient method // 10th EUROPT Workshop on Advances in Continuous Optimization, 5–7 July 2012, Siauliai, Lithuania. – P. 53.
 144. PYLYPOVSKYI O., KOSHLAI L., STETSYUK P. Model-driven DSS for economic reforms during transition // 10th EUROPT Workshop on Advances in Continuous Optimization, 5–7 July 2012, Siauliai, Lithuania. – P. 43.
 145. STETSYUK P. On one nonlinear optimization problem for Leontief models // 25th European Conference on Operational Research, 8–11 July 2012. Vilnius, Lithuania. – P. 141.
 146. PYLYPOVSKYI O., KOSHLAI L., STETSYUK P. DSS for technological planning // 25th European Conference on Operational Research, 8–11 July 2012. Vilnius, Lithuania. – P. 98.
 147. СТЕЦЮК П.И. Ускоренные модификации субградиентного метода Поляка для овражных выпуклых функций // В книге „Стохастическое программирование и его приложения“ / П.С. Кнопов, В.И. Зоркальцев, Я.М. Иванько и др. – Иркутск: Институт систем энергетики им. Л.А. Мелентьева СО РАН, 2012. – С. 160–184.
 148. СТЕЦЮК П.И. Оптимальные нормированные векторы конечного продукта и добавленной стоимости в продуктовой модели Леонтьева // В книге „Стохастическое программирование и его приложения“ / П.С. Кнопов, В.И. Зоркальцев, Я.М. Иванько и др. – Иркутск: Институт систем энергетики им. Л.А. Мелентьева СО РАН, 2012. – С. 483–492.
 149. СТЕЦЮК П.И., КОШЛАЙ Л.Б. Ускоренный субградиентный метод Поляка // Моделирование, управление и устойчивость (MCS-2012): межд. конф.; Севастополь, 10–14 сентября 2012 г. / отв. ред. О.В. Анашкин; Таврический нац. ун-т имени В.И. Вернадского. – Симферополь: ДИАЙПИ, 2012. – С. 186–187.
 150. СТЕЦЮК П.И., КОШЛАЙ Л.Б. Об одном субградиентном методе с преобразованием пространства // Крымская Осенняя Математическая Школа (КРОМШ–2012). XXIII ежегодная международная конференция. Тезисы докладов. – Симферополь: издательство КНЦ НАНУ, 2012. – С. 65–66.

151. Лиховид А.П., Ляшко В.И., Стецюк П.И. Задача линейного программирования с булевыми переменными для оптимальной суточной загрузки энергосистемы // VI-а Міжнародна школа-семинар "Теорія прийняття рішень", Ужгород, 1 – 6 жовтня 2012. Праці школи-семинару. – С. 129–130.
152. СТЕЦЮК П.И., ИВЛИЧЕВ А.В. Тестовые эксперименты с r -алгоритмом для мегабайтных систем линейных уравнений с двухсторонними границами на переменные // VI-а Міжнародна школа-семинар "Теорія прийняття рішень", Ужгород, 1 – 6 жовтня 2012. Праці школи-семинару. – С. 186–187.
153. ПАНКРАТОВ А.В., РОМАНОВА Т.Е., СТЕЦЮК П.И. Задача оптимальной кластеризации 2D-объектов в круговом контейнере // Матеріали Х міжнародної науково-практичної конференції "Математичне та програмне забезпечення інтелектуальних систем (MPZIS-2012)", Дніпропетровський національний університет, 21-23 листопада 2012р. – С. 225–226.

2013 год

154. STETSYUK P., ROMANOVA T., PANKRATOV A., KOVALENKO A. Optimal Packing Circular Cylinders into a Cylindrical Container Taking into Account Behavior Constraints // Proc. 10th ESICUP Meeting. – Lille (France). – 2013. – P. 21–22.
155. СТЕЦЮК П.И. Алгоритми недиференційовної оптимізації та лагранжеві двойсті оцінки в складних екстремальних задачах. – Автореф. дис. докт. фіз.-мат. наук., Ін-т кібернетики ім. В.М.Глушкова НАН України. – Київ, 2013. – 32 с.
156. СТЕЦЮК П.И. Алгоритмы недифференцируемой оптимизации и лагранжевые двойственные оценки в сложных экстремальных задачах // Рукопись диссертации на соискание ученой степени доктора физико-математических наук. – Ин-т кибернетики им. В.М.Глушкова НАН Украины. – Киев, 2013. – 305 с.
157. КОВАЛЕНКО А.А., ПАНКРАТОВ А.В., РОМАНОВА Т.Е., СТЕЦЮК П.И. Упаковка круговых цилиндров в цилиндрический контейнер с учетом специальных ограничений поведения системы // Журнал обчислювальної та прикладної математики. – 2013. – № 1 (111). – С. 126–134.

158. НЕНАХОВ Э.И., РОМАНОВА Т.Е., СТЕЦЮК П.И. Равновесная упаковка кругов в круг минимального радиуса // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2013. – С. 143–153.
159. СТОЯН Ю.Г., ПАНКРАТОВ А.В., РОМАНОВА Т.Е., СТЕЦЮК П.И. Методи розв'язання задач кластеризації 2D-об'єктів у крузі мінімального радіуса // Доповіді НАН України. – 2013. – № 8. – С. 51–57.
160. ДОЛГИХ В.Н., СТЕЦЮК П.И. О применении методов негладкой оптимизации для исследования эффективности сложных экономических систем // Фінансова система України: проблеми та перспективи розвитку в умовах трансформації соціально-економічних відносин: збірник тез доповідей Міжнародної науково-практичної конференції (16-18 травня 2013 р., м. Севастополь) / Севастопольський інститут банківської справи Української академії банківської справи Національного банку України ; Таврійський національний університет імені В.І. Вернадського. – Сімферополь : ТНУ ім. В.І. Вернадського, 2013. – С. 104–105.
161. СТЕЦЮК П.И., РОМАНОВА Т.Е. Равновесная упаковка шаров в шар минимального радиуса // Материалы международной конференции "Дискретная оптимизация и исследование операций" (Новосибирск, 24-28 июня 2013 г.). – Новосибирск: Издательство Института математики, 2013. – С. 142.
162. СТЕЦЮК П.И. Ускоренные по Шору модификации метода Поляка для овражных функций // Всероссийская молодежная школа-семинар "Дискретные модели и методы принятия решений": Материалы школы-семинара (г. Новосибирск, 21-23 июня 2013). – Новосибирск: Изд-во Ин-та математики, 2013. – С. 142–153.
163. СТЕЦЮК П.И., ПЕТРУХИН В.А., БУГРОВ Н.В., ХРИПКО К.Ю. Метод эллипсоидов и условно-оптимальный маршрут // Міжнародна наукова конференція "Сучасна інформатика: проблеми, досягнення та перспективи розвитку", присвяченої 90-річчю від дня народження академіка В.М. Глушкова (12-13 вересня 2013 року). – Київ: Інститут кибернетики імені В.М. Глушкова НАН України, 2013. – С. 111-113.
164. КОВАЛЕНКО А.А., ПАНКРАТОВ А.В., РОМАНОВА Т.Е., СТЕЦЮК П.И. Упаковка круговых цилиндров в цилиндрический контейнер

- с учетом специальных ограничений // Праці міжнародної наукової конференції "Питання оптимізації обчислень (ПОО-XL)", присвяченої 90-річчю від дня народження академіка В.М. Глушкова. – Київ: Інститут кібернетики імені В.М. Глушкова НАН України, 2013. – С. 123.
165. SERGIENKO I., SHYLO V., STETSUYUK P., STOYAN YU., ROMANOVA T. Computer technologies for placement and covering problems // Праці міжнародної наукової конференції "Питання оптимізації обчислень (ПОО-XL)", присвяченої 90-річчю від дня народження академіка В.М. Глушкова. – Київ: Інститут кібернетики імені В.М. Глушкова НАН України, 2013. – С. 240.
166. СТЕЦЮК П.И. Квадратичная задача для максимального сингулярного числа // Праці міжнародної наукової конференції "Питання оптимізації обчислень (ПОО-XL)", присвяченої 90-річчю від дня народження академіка В.М. Глушкова. – Київ: Інститут кібернетики імені В.М. Глушкова НАН України, 2013. – С. 255.
167. СТЕЦЮК П.И. γ -Алгоритмы: теория и практика // Матеріали XI Міжнародної науково-практичної конференції "Математичне та програмне забезпечення інтелектуальних систем (MPZIS-2013)" (20-22 листопада 2013 р., м. Дніпропетровськ). – С. 227–232.
168. СТЕЦЮК П.И., НУРМИНСКИЙ Е.А., СОЛОМОН Д.И. Транспортная задача и ортогональное проектирование на линейные многообразия // Материалы V-ой международной научной конференции "Транспортные системы и логистика", Кишинэу, 11-13 декабря 2013 года. – Кишинэу: Эврика, 2013. – С. 251–263.
169. СТЕЦЮК П.И., ЖИДКОВ В.А. О двух задачах оптимизации пропускных способностей дуг отказоустойчивой сети // Материалы V-ой международной научной конференции "Транспортные системы и логистика", Кишинэу, 11-13 декабря 2013 года. – Кишинэу: Эврика, 2013. – С. 300–309.

2014 год

170. СТЕЦЮК П.И. Методы эллипсоидов и γ -алгоритмы. – Кишинэу, Эврика, 2014. – 488 с.

171. СТЕЦЮК П.И., ЭММЕНЕГГЕР Ж.-Ф. Максимальное сингулярное число матрицы и его экономическая интерпретация // Кибернетика и системный анализ. – 2014. – № 3. – С. 51–57.
172. СТЕЦЮК П.И., РОМАНОВА Т.Є., СУБОТА І.О NLP-задача упаковки гомотетичних еліпсів в прямокутний контейнер // Теорія оптимальних рішень. – 2014. – С. 139–146.
173. СТЕЦЮК П.И., РОМАНОВА Т.Е., ШАЙТХАУЭР Г. О глобальном минимуме целевой функции в задаче равновесной упаковки кругов // Доповіди НАН України. – 2014. – № 6. – С. 53–57.
174. СТЕЦЮК П.И., ЛЯШКО В.И., ЯРЕМЧУК С.И. О решении минимаксных задач размещения источников физического поля // Журн. обчислювальної та прикладної математики. – 2014. – № 3(117). – С. 140–150.
175. БЕРЕЗОВСКИЙ О.А., ЖУРБЕНКО Н.Г., СТЕЦЮК П.И. Алгоритмы построения линейных бинарных классификаторов при неточных измерениях // Компьютерная математика. – Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 2014. – № 2. – С. 133–138.
176. КОВАЛЕНКО А.А., ПАНКРАТОВ А.В., РОМАНОВА Т.Е., СТЕЦЮК П.И. Балансная компоновка цилиндрических объектов: математические модели и методы решения // Радиоэлектроника и информатика. – 2014. – № 4. – С. 70–79.
177. ДМИТРЕНКО В.В., ЕРШОВ А.В., СТЕЦЮК П.И., ЛИХОВИД А.П., ЛАПТИН Ю.П., ШВАРЦ Д.Р., МЕКЛЕР А.А., КАВСАН В.М. Определение молекулярных подклассов глиобластом на основе анализа экспрессии генов // Цитология и генетика. – 2014. – Т.48. № 6. – С. 45–55.
178. STOYAN YU., STETSYUK P., ROMANOVA T. Optimal Balanced Packing Using Phi-Function Technique // In S. Butenko, E.L. Pasilio, and V. Shylo (Editors), Examining Robustness and Vulnerability of Networked Systems, pages 251–271. IOS Press, 2014.
179. СТЕЦЮК П.И., МІЦА О.В., ПЕЦКО В.І. Мережні інформаційні технології: методичні рекомендації до вивчення курсу. – Ужгород: Видавництво УжНУ <Говерла>, 2014. – 65 с.

180. СОЛОМОН Д.И., СТЕЦЮК П.И., ЧУМАКОВ Б.М. 4-ая международная научная конференция "Математическое моделирование, оптимизация и информационные технологии"(ММОТИ-2014) // Кибернетика и системный анализ. – 2014. – № 4. – С. 186–187.
181. СТЕЦЮК П.И., РОМАНОВА Т.Е., КОВАЛЕНКО А.А. Комп'ютерна програма Balance circular packing // Свідоцтво про реєстрацію авторського права на твір № 55689. Україна. Міністерство освіти і науки. Державний департамент інтелектуальної власності. – Дата реєстрації 21.07.2014.
182. СУББОТА И.А., РОМАНОВА Т.Е., СТЕЦЮК П.И. Комп'ютерна програма Optimal packing of homothetic oriented ellipses // Свідоцтво про реєстрацію авторського права на твір № 55690. Україна. Міністерство освіти і науки. Державний департамент інтелектуальної власності. – Дата реєстрації 21.07.2014.
183. СТЕЦЮК П.И., РОМАНОВА Т.Е., КОВАЛЕНКО Г.А. Комп'ютерна програма "Рівноважна упаковка кругів" // Свідоцтво про реєстрацію авторського права на твір № 56609. Україна. Міністерство освіти і науки. Державний департамент інтелектуальної власності. – Дата реєстрації 23.09.2014.
184. СЕРГИЕНКО И.В., СТЕЦЮК П.И. Две ЛП-задачи с булевыми переменными для отказоустойчивой сети // Информатика та системні науки (ІСН–2014): матеріали V Всеукр. наук.-практ. конф. (м. Полтава, 13–15 березня 2014 року). – Полтава: ПУЕТ, 2014. – С. 284–287.
185. STETSYUK P., ROMANOVA T., SCHEITHAUER G. On the global minimum in a balanced circular packing problem // Материалы 4-й Межд. науч. конф. "Математическое моделирование, оптимизация и информационные технологии": (Кишинэу, 25–28 марта 2014 г.). – Кишинэу: Эврика, 2014. – Т. I. – С. 206–214.
186. ЗОРКАЛЬЦЕВ В.И., ПЕРЖАВИНСКИЙ С.М., СТЕЦЮК П.И. Поиск нормальных решений СЛАУ методом внутренних точек // Математическое моделирование, оптимизация и информационные технологии: материалы 4-й Междунар. науч. конф. (Кишинэу, 25–28 марта 2014 г.). – Кишинэу: Эврика, 2014. – Т. II. – С. 202–212.
187. РОМАНОВА Т.Е., СУББОТА И.А., СТЕЦЮК П.И. Задача упаковки гомотетичних еліпсів в прямокульному контейнері мінімаль-

- ных размеров // Математическое моделирование, оптимизация и информационные технологии: материалы 4-й Междунар. науч. конф. (Кишинэу, 25–28 марта 2014 г.). – Кишинэу: Эврика, 2014. – Т. II. – С. 396–405.
188. СТЕЦЮК П.И. 2D-эллипсоид и r-алгоритмы // Математическое моделирование, оптимизация и информационные технологии: материалы 4-й Междунар. науч. конф. (Кишинэу, 25–28 марта 2014 г.). – Кишинэу: Эврика, 2014. – Т. II. – С. 466–471.
189. СТЕЦЮК П.И., ЖИДКОВ В.А., ФЕСЮК А.В. Octave-функция OrtPro: назначение и особенности использования // Системный анализ и информационные технологии: материалы 16-й Междунар. науч. конф. SAIT 2014 (Киев, 26–30 мая 2014 г.) / ННК "ИПСА" НТУУ "КПИ". – К.: ННК "ИПСА" НТУУ "КПИ", 2014. – С.156–157.
190. СТЕЦЮК П.И. Алгоритм описанных эллипсоидов для задачи о наилучшем линейном классификаторе // Тези VII Міжнародної науково-технічної конференції "Інформаційно-комп'ютерні технології 2014" (29–30 травня 2014 року). – Житомир: ЖДТУ, 2014. – С. 57–58.
191. СТЕЦЮК П.И. 1-d и 2-d эллипсоиды в выпуклом программировании // Тезисы докладов XVI Байкальской международной школы-семинара "Методы оптимизации и их приложения". – Иркутск, ИСЭМ СО РАН. – 2014. – С. 118.
192. Stetsyuk P. Circumscribed 2d-ellipsoid and Shor's r-algorithm // Proceedings of abstracts of reports presented at the V International Conference on Optimization Methods and Applications "Optimization and applications" (OPTIMA-2014) held in Petrovac, Montenegro, September 28 – October 4, 2014. – P. 183–184.
193. ROMANOVA T., STETSYUK P., KOVALENKO A. Optimal balanced circular packing problem // Proceedings of abstracts of reports presented at the V International Conference on Optimization Methods and Applications "Optimization and applications" (OPTIMA–2014) held in Petrovac, Montenegro, September 28 – October 4, 2014. – P. 158–159.
194. СТЕЦЮК П.И. Субградієнтний алгоритм з розтягом простору для задачі розділення двох полієдрів // Праці VII-ої міжнародної

- школи-семінару "Теорія прийняття рішень". – Ужгород, УжНУ, 2014. – С. 244–245.
195. КОВАЛЕНКО А.А., ПАНКРАТОВ А.В., РОМАНОВА Т.Е., СТЕЦЮК П.И. Модели и методы решения задач балансной компоновки цилиндров // Матеріали XII Міжнародної науково-практичної конференції "Математичне та програмне забезпечення інтелектуальних систем (MPZIS–2014)" (19–21 листопада 2014 р., м. Дніпропетровськ). – С. 122–123.
196. СТЕЦЮК П.И., ФЕСЮК А.В. Об одной задаче квадратичного программирования // Матеріали XII Міжнародної науково-практичної конференції "Математичне та програмне забезпечення інтелектуальних систем (MPZIS–2014)" (19–21 листопада 2014 р., м. Дніпропетровськ). – С. 220–221.
197. СТЕЦЮК П.И. Чего следует ожидать от γ -алгоритма Шора при решении задач энтропийно-линейного программирования большой размерности? // Труды 57-й научной конференции МФТИ "Актуальные проблемы фундаментальных и прикладных наук в современном информационном обществе". – М.: МФТИ, 2014.
198. СТЕЦЮК П.И. Максимальное сингулярное число в статической модели затраты-выпуск // Матеріали Всеукраїнської науково-практичної конференції "В.М. Глушков – піонер кібернетики" (2014 р., м. Київ). К.: Видавництво "Політехніка", 2014. – С. 70–72.

2015 год

199. Институциональные и технологические изменения в странах с рыночной и переходной экономикой / Стецюк П.И. Бортис Г., Эмменеггер Ж.-Ф. и др. – К.: Видавничий дім "Києво-Могилянська Академія", 2015. – 336 с.
200. СТЕЦЮК П.И., ИВЛИЧЕВ А.В., ИЩЕНКО А.А. О сходимости $\gamma\mu(\alpha)$ -алгоритма // Компьютерная математика. – Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 2015. – № 1. – С. 142–152.
201. СТЕЦЮК П.И., ГАСНИКОВ А.В. NLP-программы и γ -алгоритм в задаче энтропийно-линейного программирования // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 2015. – С. 73–78.

202. КОВАЛЕНКО А.А., РОМАНОВА Т.Е., СТЕЦЮК П.И. Задача балансной компоновки 3D-объектов: математическая модель и методы решения // Кибернетика и системный анализ. – 2015. – № 4. – С. 71–81.
203. СТЕЦЮК П.И. Булева задача для найкоротшого k-вершинного циклу у повному графі // Сучасні проблеми прикладної математики та інформатики. Збірник наукових праць. – Львів: Львівський національний університет імені Івана Франка. – 2015. – С. 297–299.
204. ЗОРКАЛЬЦЕВ В.И., ПЕРЖАБИНСКИЙ С.М., СТЕЦЮК П.И. Поиск нормальных решений СЛАУ при двусторонних ограничениях на переменные методом внутренних точек // Кибернетика и системный анализ. – 2015. – № 6. – С. 71–80.
205. СТЕЦЮК П.И., ФЕСЮК А.В. Двойственный алгоритм решения задачи сепарабельного квадратичного программирования с одним ограничением и границами на переменные // Питання прикладної математики і математичного моделювання. – Дніпропетровськ, 2015. С. 191–200.
206. СТЕЦЮК П.И., ЛЕФТЕРОВ А.В., ФЕДОСЕЕВ А.И. Кратчайший k-вершинный путь // Компьютерная математика. – Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 2015. – № 2. – С. 3–11.
207. ГАСНИКОВ А.В., ДВУРЕЧЕНСКИЙ П.Е., КАМЗОЛОВ Д.И., НЕСТЕРОВ Ю.Е., СПОКОЙНЫЙ В.Г., СТЕЦЮК П.И., СУВОРИКОВА А.Л., ЧЕРНОВ А.В. Поиск равновесий в многостадийных транспортных моделях // Труды МФТИ. – 2015. – Т. 7, № 4. – С. 143–155.
208. СТЕЦЮК П.И., ЛИХОВИД О.П. Комп'ютерна програма "A parallel algorithm for a balanced circular packing problem" // Свідоцтво про реєстрацію авторського права на твір № 62184. Україна. Міністерство освіти і науки. Державний департамент інтелектуальної власності. – Дата реєстрації 20.10.2015.
209. ГАСНИКОВ А.В., ДВУРЕЧЕНСКИЙ П.Е., НЕСТЕРОВ Ю.Е., СПОКОЙНЫЙ В.Г., СТЕЦЮК П.И., СУВОРИКОВА А.Л., ЧЕРНОВ А.В. Поиск равновесий в многостадийных транспортных моделях <http://arxiv.org/ftp/arxiv/papers/1506/1506.00292.pdf>
210. СТЕЦЮК П.И. Общая схема метода эллипсоидов // Информационный бюллетень Ассоциации математического программирования

- № 13. 15-я Всероссийская конференция "Математическое программирование и приложения"(тезисы докладов) Екатеринбург: ИММ УрО РАН, 2015. – С. 59–60.
211. СТЕЦЮК П.И. Об одном обобщении классического метода эллипсоидов // информатика та системні науки (ІСН–2015): матеріали VI Всеукр. наук.-практ. конф. за міжнародною участю (м. Полтава, 19–21 березня 2015 року). – Полтава: ПУЕТ, 2015. – С. 335–337.
212. FEDOSIIEV O., LEFTEROV O., LYKHOVYD O., STETSYUK P. Optimization tools for selecting wine routes // Theses of the conference CAIM–2015, Stefan cel Mare University of Suceava, Romania, September 17–20, 2015. – P. 37–38.
213. СТЕЦЮК П.И., ЗОРКАЛЬЦЕВ В.М., ЭММЕНЕГГЕР Ж.-Ф. Неразложимость неотрицательных матриц // Праці міжнародної наукової школи-семінару "Питання оптимізації обчислень (ПОО–XLII)", присвяченої 85-річчю від дня народження академіка В.С. Михалевича. – Київ: Інститут кібернетики імені В.М. Глушкова НАН України, 2015. – С. 122–123.
214. FEDOSIIEV O., LEFTEROV O., LYKHOVYD O., STETSYUK P. Optimization tools for touristic routes // Материалы VI-ой международной научной конференции "Транспортные системы и логистика", Кишинэу, Республика Молдова, 27–30 октября 2015 года. – Кишинэу: Эврика, 2015. – С. 341–346.
215. GASNIKOV A., DVURECHENSKY P., STETSYUK P., SUVORIKOVA A., CHERNOV A. Equilibriums in multistage transport problems // Proceedings of abstracts of reports presented at the V International Conference on Optimization Methods and Applications "Optimization and applications"(OPTIMA–2015) held in Petrovac, Montenegro, September 27 – October 3, 2015. – P. 73–74.
216. СТЕЦЮК П.И. Две задачи о кратчайшем k -вершинном цикле // Матеріали XIII Міжнародної науково-практичної конференції "Математичне та програмне забезпечення інтелектуальних систем (MPZIS–2015)"(18–20 листопада 2015 р., м. Дніпропетровськ). – С. 215–221.
217. СТЕЦЮК П.И. О задачах М.В.Михалевича для межотраслевых структурно-технологических изменений // Кіберсоціум – міф чи

реальність?: зб. матеріалів IV Всеукр. наук.-практ. конф. "Глушковські читання", Київ, 2 грудня 2015 року. – Київ: Вид-во "Політехніка", 2015. – С. 146–148.

2016 год

218. СТЕЦЮК П.И. Формулировки задач для кратчайшего k-вершинного пути и кратчайшего k-вершинного цикла в полном графе // Кибернетика и системный анализ. – 2016. – № 1. – С. 78–82.
219. STETSUYUK P., ROMANOVA T., SCHEITHAUER G. On the global minimum in a balanced circular packing problem // Optimization Letters. – 2016, № 10, P. 1347–1360.
220. СТЕЦЮК П.І., Андрияш М.М. Метод Сааті для аналізу показників податкової трансформації // Компьютерная математика. – Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 2016. – № 1. – С. 37–45.
221. СТЕЦЮК П.И., Долинский Э.С., Парасюк И.И. О кратчайшем k-вершинном пути в ориентированном графе // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 2016. – С. 95–102.
222. СТЕЦЮК П.І., Слабоспицька О.О., Ушакова О.О. Максимальні незалежні множини вершин графа та їх застосування в керуванні проектами // Питання прикладної математики і математичного моделювання. – Д.: РВВ ДНУ, 2016. – Вип. 16. – С. 151–162.
223. Гасников А.В., Двуреченский П.Е., Спокойный В.Г., Стецюк П.И., Суворикова А.Л. Суперпозиция метода балансировки и универсального градиентного метода для поиска энтропийно-регуляризованного барицентра Вассерштейна и равновесий в многостадийных моделях транспортных потоков // Труды МФТИ. – 2016. – Т. 8, № 3. – С. 5–24.
224. СТЕЦЮК П.І., Лиховид О.П., ФЕСЮК О.В. NLP-програми для ELD-задач завантаження енергосистеми // Компьютерная математика. – Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 2016. – № 2. – С. 142–150.

225. СТЕЦЮК П.І., ХІМІЧ О.М., СИДОРУК В.А. Реалізація г-алгоритму на графічних процесорах // Комп'ютерна математика. – Киев: Ін-т кібернетики ім. В.М.Глушкова НАН України, 2016. – № 2. – С. 100–109.
226. STOYAN YU., ROMANOVA T., PANKRATOV A., KOVALENKO A., STETSYUK P. Balance layout problems: Mathematical modeling and nonlinear optimization // In book "Space Engineering. Modeling and Optimization with Case Studies" / Springer Optimization and its Applications, G. Fasano and J. Pinter (Editors), Springer, New York, Vol. 114. – 2016. – P. 369–400.
227. СТЕЦЮК П.И., ДОЛИНСКИЙ Э.С. Кратчайший k-вершинный путь в ориентированном графе // Информатика та системні науки (ІСН–2016): матеріали VII Всеукр. наук.-практ. конф. за міжнародною участю(м. Полтава, 10–12 березня 2016 року) /за ред. Ємця О.О. – Полтава: ПУЕТ, 2016. – С. 293–299.
228. STOYAN YU.G., ROMANOVA T.E., PANKRATOV A.V., KOVALENKO A.A., STETSYUK P.I. Balance layout problems: mathematical modeling and nonlinear optimization // Матеріали 5-й Межд. науч. конф. "Математическое моделирование, оптимизация и информационные технологии": (Кишинэу, 22–25 марта 2016 г.). – Кишинэу: Эврика, 2016. – Т. I. – С. 329–331.
229. СТЕЦЮК П.И. Максимальное сингулярное число для матрицы полных затрат // Математическое моделирование, оптимизация и информационные технологии: материалы 5-й Междунар. науч. конф. (Кишинэу, 22–25 марта 2016 г.). – Кишинэу: Эврика, 2016. – Т. II. – С. 327–336.
230. СТЕЦЮК П.И., ЛЕФТЕРОВ А.В., ЛИХОВИД А.П., ФЕДОСЕЕВ А.И. Оптимизационный сервис для выбора винных маршрутов // Математическое моделирование, оптимизация и информационные технологии: материалы 5-й Междунар. науч. конф. (Кишинэу, 22–25 марта 2016 г.). – Кишинэу: Эврика, 2016. – Т. II. – С. 337–344.
231. ХЛУД О.М., РОМАНОВА Т.Е., СТЕЦЮК П.И. Задача упаковки гомотетичних еліпсоїдов // Математическое моделирование, оптимизация и информационные технологии: материалы 5-й Междунар. науч. конф. (Кишинэу, 22–25 марта 2016 г.). – Кишинэу: Эврика, 2016. – Т. II. – С. 373–377.

232. СТЕЦЮК П.И., СОЛОМОН Д.И. Вычислительные аспекты задачи коммивояжера // Тези доповідей VIII Міжнародної науково-технічної конференції "Інформаційно-комп'ютерні технології 2016"(22–23 квітня 2016 року). – Житомир: ЖДТУ, 2016. – С. 91–92.
233. СТЕЦЮК П.И. Оптимальный 2d-эллипсоид и растяжение пространства по разности нормированных субградиентов // Праці VIII-ої міжнародної школи-семінару "Теорія прийняття рішень". – Ужгород, УжНУ, 2016. – С. 249–250.
234. СТЕЦЮК П.И. Об r-алгоритмах Шора // VIII Московская международная конференция по Исследованию Операций (ORM2016), Москва, 17–22 октября 2016. Труды. – М.: Издательство ФИЦ ИУ РАН, 2016. – Т. 2. – С. 60–62.
235. СТЕЦЮК П.И. Программа galgb5 для минимизации овражных выпуклых функций // Тезисы докладов XIV Международной научно-практической конференции "Математическое и программное обеспечение интеллектуальных систем (MPZIS–2016)", Днепр, 16–18 ноября 2016 г. – Д.: ДНУ, 2016. – С. 185–197.
236. СТЕЦЮК П.І., ЛИХОВИД О.П., ФЕСЮК О.В. Система Маневр-New для розв'язання оптимізаційних задач завантаження енергоблоків ТЕС // Міждисциплінарні дослідження актуальних проблем застосування інформаційних технологій в сучасному світі: зб. матеріалів V Всеукр. наук.-практ. конф. "Глушковські читання", Київ, 24 листопада 2016 року. – К: ТОВ НВП "Інтерсервіс", 2016. – С. 174–176.
237. СТЕЦЮК П.І., СТОВБА В.О., МАРТИНЮК І.С. Octave-програма dist2r для розділення двох полієдрів // Теоретичні та прикладні аспекти побудови програмних систем: матеріали міжнародної наукової конференції, м. Київ, 5–9 грудня 2016 р. / редкол.: М.С.Нікітченко, Д.Б.Буй та ін. – Кіровоград: "Центр оперативної поліграфії "Авангард". – 2016. – С. 227–230.

2017 год

238. STETSUYUK P.I. Shor's r-Algorithms: Theory and Practice. In: Optimization Methods and Applications: In Honor of the 80th Birthday of Ivan V. Sergienko. Ed. by S. Butenko, P.M. Pardalos, V. Shylo. – Springer. 2017. – P. 495–520.

239. СТЕЦЮК П.И. Субградиентные методы ralgb5 и ralgb4 для минимизации овражных выпуклых функций // Вычислительные технологии. – 2017. – Т. 22. – № 2. – С. 127–149.
240. СТЕЦЮК П.И., В.А. СТОВБА В.А., МАРТЫНЮК И.С. Алгоритмы метода эллипсоидов для нахождения L_p -решения системы линейных уравнений // Теорія оптимальних рішень. – Київ: Інститут кібернетики ім. В.М.Глушкова НАН України. – 2017. – С. 139–146.
241. МІЦА О.В., ПЕЦКО В.І., СТЕЦЮК П.І. Двоетапний метод для синтезу багаточарових оптичних систем // Комп'ютерна математика. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2017. – № 1. – С. 126–133.
242. СТЕЦЮК П.И. Теория и программные реализации g -алгоритмов Шора // Кибернетика и системный анализ. – 2017. – № 5. – С. 43–57.
243. СТЕЦЮК П.І., ФЕСЮК О.В. Використання g -алгоритму для розв'язання квадратичної ELD-задачі // Математичне та комп'ютерне моделювання. Серія: Фізико-математичні науки: зб. наук. праць. – Кам'янець-Подільський: Кам'янець-Подільський національний університет імені Івана Огієнка, 2017. – Випуск 15. – С. 225–231.
244. СТЕЦЮК П.І., БАРДАДИМ Т.О., ЛЯШКО В.І. Квадратична задача для максимального k -плекса в неорієнтованому графі // Журн. обчислювальної та прикладної математики. – 2017. – № 1(124). – С. 71–78.
245. STETSYUK P.I. 2d-Ellipsoid of optimal volume and its applications // In: L.N.Polyakova (ed.) Constructive Nonsmooth Analysis and Related Topics (Dedicated to the Memory of V.F. Demyanov) (CNSA). – IEEE, 2017. – P. 303–306.
246. СТЕЦЮК П.І., МІЦА О.В., СТРЕЛЮК О.В., ФЕСЮК О.В. Транспортна задача з обмеженнями на пропускні спроможності проміжних пунктів // Питання прикладної математики і математичного моделювання. – Д.: ДНУ, 2017. – Вип. 17. – С. 207–219.
247. СТЕЦЮК П.І., ЛЯШКО В.І., БАРДАДИМ Т.О. Властивості квадратичної задачі для максимального k -плекса у неорієнтованому графі // Наукові записки НаУКМА. – 2017. Том 198. Комп'ютерні науки. – С. 8–13.

248. СТЕЦЮК П.І., ФЕСЮК О.В., СИДОРУК В.А. Алгоритми розв'язання задачі сепарабельного квадратичного програмування // Комп'ютерна математика. – Київ: Ін-т кібернетики ім. В.М. Глушкова НАН України, 2017. – № 2. – С. 137–146.
249. СТЕЦЮК П.І., БИЛА Г.Д., СТОВБА В.А. Метод еліпсоїдів для знаходження L_p -решення системи лінійних рівнянь // Інформатика та системні науки (ІСН–2017): матеріали VIII Всеукраїнської науково-практичної конференції за міжнародною участю, (м. Полтава, 16–18 березня 2017 р.) / за ред. Ємця О.О. – Полтава: ПУЕТ, 2017. – С. 258–264.
250. СТЕЦЮК П.І., САВИЦКИЙ В.В. О робастности метода наименьших модулей для поиска дефектов в регулярных 3D-структурах // Матеріали Дев'ятнадцятого Міжнародного науково-практичного семінару "Комбінаторні конфігурації та їх застосування" присвяченого пам'яті д.ф.-м.н., професора Петренюка Анатолія Яковича (Кропивницький, 7–8 квітня 2017 року) / За ред. Г.П. Донець – Кропивницький, 2017. – С. 127–132.
251. СТЕЦЮК П.І., СТОВБА В.О. Метод еліпсоїдів для лінійної регресії // Обчислювальний інтелект (результати, проблеми, перспективи): праці міжнар. наук.-практ. конф., 16–18 травня 2017 р., Київ–Черкаси; наук. ред. В.Є. Снитюк. – К. ВПЦ "Київський університет", 2017. – С. 314–315.
252. СТЕЦЮК П.І. г-Алгоритмы Шора для сепарабельной задачи квадратичного программирования // Тезиси докладов международной конференции "Конструктивный негладкий анализ и смежные вопросы", посвященной памяти профессора В.Ф. Демьянова. Часть II. – СПб.: Издательство ВВМ, 2017. – С. 95–99.
253. СТЕЦЮК П.І. Оптимальный по объему 2d-эллипсоид и его приложения // Тезиси докладов международной конференции "Конструктивный негладкий анализ и смежные вопросы", посвященной памяти профессора В.Ф. Демьянова. Часть II. – СПб.: Издательство ВВМ, 2017. – С. 173–176.
254. СТЕЦЮК П.І. Метод еліпсоїдів с берегов Дніпра // Математичне та програмне забезпечення інтелектуальних систем: Тези доповідей XV Міжнародної науково-практичної конференції MPZIS–2017, 22–24 листопада 2017 р. – Д.: ДНУ, 2017. – С. 185.

255. СТЕЦЮК П.І., МАЗЮТИНЕЦЬ Г.В., МІЛЕСОВСЬКИЙ Б.І. АМРЛ-реалізація двоетапної транспортної задачі // Математичне та програмне забезпечення інтелектуальних систем: Тези доповідей XV Міжнародної науково-практичної конференції MPZIS-2017, 22–24 листопада 2017 р. – Д.: ДНУ, 2017. – С. 186–191.
256. СТЕЦЮК П.І., ФИШЕР А. г-Алгоритмы Шора и Octave-функция `ralgb5a` // Тези доповідей Міжнародної наукової конференції "Сучасна інформатика: проблеми, досягнення та перспективи розвитку", присвяченої 60-річчю заснування Інституту кібернетики імені В.М.Глушкова НАН України. Україна, Київ, 13–15 грудня 2017 року. – К: Інститут кібернетики імені В.М.Глушкова НАН України, 2017. – С. 143–146.
257. СТЕЦЮК П.І. О мультипликаторах Михалевича // Цифрова революція в соціально-економічній сфері: історія і перспективи: матеріали VI Всеукр. наук.-практ. конф. «Глушковські читання», 13 грудня 2017 р., м. Київ / Уклад.: А.А.Мельниченко, В.Д.Піхорович, П.А.Богдан. – Київ: КПІ ім. Ігоря Сікорського, Вид-во "Політехніка", 2017. – С. 155–156.

2018 год

258. СТЕЦЮК П.І., САВИЦКИЙ В.В. О поиске дефектов в регулярных 3D-структурах // Проблемы управления и информатики. – 2018. – № 2. – С. 33–48.
259. СТЕЦЮК П.І., СТОВБА В.А., ЖМУД А.А. Метод еліпсоїдов для знаходження рішення переопределенной СЛАУ // Теорія оптимальних рішень. – Київ: Інститут кібернетики ім. В.М.Глушкова НАН України. – 2018. – С. 115–123.
260. СТЕЦЮК П.І., ФЕСЮК О.В., БУТКЕВИЧ О.Ф. Опуклі квадратичні ЕД-задачі: властивості та субградієнтні алгоритми розв'язання // Комп'ютерна математика. – Киев: Ин-т кібернетики ім. В.М. Глушкова НАН України, 2018. – № 1. – С. 133–141.
261. СТЕЦЮК П.І., ФЕСЮК А.В., ХОМЯК О.Н. Обобщенный метод эллипсоидов // Кибернетика и системный анализ. – 2018. – № 4. – С. 70–80.

262. ФЕСЮК А.В., СТЕЦЮК П.І., БУТКЕВИЧ О.Ф. Використання системи MANEUVER-NEW для розв'язання задач оптимального завантаження енергоблоків теплових електростанцій // Технічна електродинаміка. – 2018. – № 4. – С. 94–97.
263. СТЕЦЮК П.І., ЛЯШКО В.І., МАЗЮТИНЕЦЬ Г.В. Двоетапна транспортна задача та її АМРЛ-реалізація // Наукові записки НаУ-КМА. Комп'ютерні науки. – 2018. – Т. 1. – С. 14–20.
264. СТЕЦЮК П.І., ТРЕГУБЕНКО С.С. Про двоетапну транспортну задачу з заданою кількістю проміжних пунктів // Матеріали Двадцятого Міжнародного науково-практичного семінару „Комбінаторні конфігурації та їх застосування“ (Кропивницький, 13–14 квітня 2018 року) / За ред. Г.П. Донець – Кропивницький, 2018. – С. 128–132.
265. СТЕЦЮК П.І., ТРЕГУБЕНКО С.С. Двоетапна транспортна задача з заданою кількістю проміжних пунктів // Тези доповідей ІХ Міжнародної науково-технічної конференції „Інформаційно-комп'ютерні технології 2018“ (20–21 квітня 2018 р.). – Житомир: Вид. О.О. Євенок, 2018. – С. 99–100.
266. СТЕЦЮК П.И. Об ускорении субградиентного метода с шагом Поляка // Динамические системы: устойчивость, управление, оптимизация = Dynamical systems: stability, control, optimization: материалы Междунар. науч. конф., посвященной 100-летию со дня рождения академика Е.А. Барбашина, Минск, 24–29 сент. 2018 г. – Минск: БГУ, 2018. – С. 206–208.
267. PETRO STETSYUK, VIKTOR STOVBA, ZHANNA CHERNOUSOVA. Subgradient Method with Polyak's Step in Transformed Space // Book of abstracts of reports presented at the IX International Conference on Optimization Methods and Applications „Optimization and applications“ (OPTIMA-2018) held in Petrovac, Montenegro, October 1 – October 5, 2018. – Dorodnicyn Computing Centre of FRC Computer Science and Control of Russian Academy of Science, 2018. – P. 161.

СОДЕРЖАНИЕ

Предисловие	6
-------------------	---

Р А З Д Е Л I

Двойственные оценки в квадратичных задачах
(пионерские работы Шора)

О методе получения оценок в квадратичных экстремальных задачах с булевыми переменными (<i>Шор, Давыдов, 1985</i>)	18
Квадратичные оптимизационные задачи (<i>Шор, 1987</i>)	26
Об одном подходе к получению глобальных экстремумов в полиномиальных задачах математического программирования (<i>Шор, 1987</i>)	44
Об одном классе оценок глобального минимума полиномиальных функций (<i>Шор, 1987</i>)	55
Задачи минимизации матричных функций и недифференцируемая оптимизация (<i>Шор, 1995</i>)	62

Р А З Д Е Л II

Двойственные оценки Шора

Использование модификации г-алгоритма для нахождения глобального минимума полиномиальных функций (<i>Шор, Стецюк, 1997</i>)	94
Использование г-алгоритма в задачах полуопределенного программирования (<i>Шор, Стецюк, 1997</i>)	123
Нахождение глобальных минимумов полиномиальных функций с использованием двойственных квадратичных оценок (<i>Шор, Стецюк, Крылов, 1997</i>)	130
Dual Solution of Quadratic-Type Problems by g-algorithm (subroutine DSQTPPr) (<i>Shor, Stetsyuk, 2001</i>)	148
Lagrangian bounds in multiextremal polynomial and discrete optimization problems (<i>Shor, Stetsyuk, 2002</i>)	150

Р А З Д Е Л III

Новые семейства функционально избыточных ограничений в квадратичных булевых задачах

- О функционально избыточных ограничениях для булевых
 оптимизационных задач квадратичного типа (*Стецюк, 2005*).....198
- Новые модели квадратичного типа для задачи о максимальном
 взвешенном разрезе графа (*Стецюк, 2006*).....204
- Об уточнении лагранжевых двойственных оценок в бинарных
 и булевых квадратичных задачах (*Стецюк, Пардалос, 2006*).....221
- Об уточнении верхних оценок Н.З. Шора для числа
 независимости графа (*Стецюк, 2007*)230
- Бинарный квадратичный многогранник и его
 аппроксимации (*Стецюк, Золотых, 2010*).....232
- О внешней аппроксимации бинарного и булевого
 квадратичных многогранников (*Золотых, Стецюк, 2010*).....245

Р А З Д Е Л IV

Оценки для числа устойчивости графа

- О новых лагранжевых двойственных оценках для числа
 устойчивости графа (*Стецюк, Пардалос, Крошко, 2006*).....250
- Об одной верхней оценке для взвешенного числа
 устойчивости графа (*Стецюк, Бутенко, Березовский, 2007*).....261
- Точная ЛП-оценка для взвешенного числа устойчивости
 t -совершенных графов (*Стецюк, Ляшко, Нурминский, 2009*).....271
- О новых свойствах оценок Шора для взвешенного числа
 устойчивости графа (*Стецюк, 2008*).....284
- ЛП-ориентированная верхняя оценка для числа устойчивости
 графа на основе p -колес (*Стецюк, Бутенко, Лиховид, 2008*)298
- Об ЛП-ориентированных верхних оценках для взвешенного числа
 устойчивости графа (*Стецюк, Лиховид, 2009*).....310
- Оценки Н.З. Шора в квадратичных экстремальных задачах и их
 применение в комбинаторной оптимизации (*Стецюк, 2010*).....333

Р А З Д Е Л V

Специальные квадратичные задачи (часть I)

Двойственные оценки для специальной оптимизационной задачи квадратичного типа на многообразии Штифеля (<i>Шор, Стецюк, Березовский, 2004</i>)	342
Лагранжевая оценка для максимального разреза графа с заданными количествами вершин в обоих подмножествах разбиения (<i>Стецюк, Березовский, 2006</i>)	351
Квадратичная оптимизационная задача для максимального k -плекса в неориентированном графе (<i>Стецюк, 2006</i>)	360
Квадратичная задача для максимального k -плекса в неориентированном графе (<i>Стецюк, Бардадим, Ляшко, 2017</i>)	378
Властивості квадратичної задачі про максимальний k -плекс у неорієнтованому графі (<i>Стецюк, Ляшко, Бардадим, 2017</i>)	388

Р А З Д Е Л VI

Специальные квадратичные задачи (часть II)

Оптимальная нормированная структура спроса и добавленной стоимости в продуктивной модели Леонтьева (<i>Стецюк, Кошлай, 2010</i>)	400
Об одной экстремальной задаче для связи прямой и двойственной моделей Леонтьева (<i>Стецюк, Кошлай, 2011</i>)	414
О спектральных свойствах матриц Леонтьева (<i>Стецюк, 2011</i>)	423
Максимальное сингулярное число матрицы и его экономическая интерпретация (<i>Стецюк, Эмменеггер, 2014</i>)	428
О глобальном минимуме целевой функции в задаче равновесной упаковки кругов (<i>Стецюк, Романова, Шайтхауэр, 2014</i>)	438
Задача равновесной упаковки кругов (<i>Стецюк, Лиховид, 2015</i>)	444
Список научных трудов (1982–2018)	466
Содержание	501

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

МИНИСТЕРСТВО ПРОСВЕЩЕНИЯ, КУЛЬТУРЫ И НАУКИ
РЕСПУБЛИКИ МОЛДОВА
АКАДЕМИЯ ТРАНСПОРТА, ИНФОРМАТИКИ И КОММУНИКАЦИЙ

П. И. СТЕЦЮК

**ДВОЙСТВЕННЫЕ ОЦЕНКИ
В КВАДРАТИЧНЫХ
ЭКСТРЕМАЛЬНЫХ ЗАДАЧАХ**

Издательство «Эврика», тел. 022-38-45-30
мун. Кишинэу, Республика Молдова
Мунчештское шоссе, 121а

Подписано в печать 02.11.2018
Формат 60x90/16
Усл. печ. листов 31,5
Тираж 300 экз.
Заказ № 22

Типография Академии Наук Молдовы
г. Кишинев, ул. Петру Мовилэ, 8



СТЕЦЮК ПЕТР ИВАНОВИЧ

Заведующий отделом методов негладкой оптимизации Института кибернетики имени В.М. Глушкова Национальной академии наук Украины, доктор физико-математических наук.

Родился в 1958 году в селе Кунисовцы, Городенковского района, Ивано-Франковской области. Окончил факультет управления и прикладной математики (1982 г.) и аспирантуру (1992 г.) Московского физико-технического института. С 1982 г. работает в Институте кибернетики НАНУ, где защитил кандидатскую (1996 г.) и докторскую (2013 г.) диссертации.

Сфера научных интересов – линейное и нелинейное программирование; методы минимизации негладких функций, их применение в блочных, матричных, полиномиальных и булевых задачах оптимизации; алгоритмы нахождения и уточнения двойственных оценок целевой функции в квадратичных многоэкстремальных задачах; математическое и программное обеспечение прикладных задач оптимизации.

Автор 6 книг, 8 разделов в монографиях, более 200 научных работ, учебных курсов "Методы негладкой оптимизации и математическое программирование", "Методы выпуклой оптимизации", "Оптимизационные алгоритмы и сетевые технологии", "Методы эллипсоидов и r -алгоритмы".

Лауреат Государственной премии Украины в области науки и техники (2016), премии им. академика В.С. Михалевича НАН Украины (2017).