Національна академія наук України

Інститут кібернетики імені В.М. Глушкова

СТЕЦЮК Петро Іванович

УДК 519.8

АЛГОРИТМИ НЕДИФЕРЕНЦІЙОВНОЇ ОПТИМІЗАЦІЇ ТА ЛАГРАНЖЕВІ ДВОЇСТІ ОЦІНКИ В СКЛАДНИХ ЕКСТРЕМАЛЬНИХ ЗАДАЧАХ

01.05.01 – теоретичні основи інформатики та кібернетики

Автореферат
дисертації на здобуття наукового ступеня
доктора фізико-математичних наук

Київ – 2013

Дисертацією є рукопис.
Робота виконана в Інституті кібернетики імені В.М. Глушкова НАН України.

Науковий консультант: доктор фізико-математичних наук, професор,
академік НАН України
Сергієнко Іван Васильович,
Інститут кібернетики імені В.М. Глушкова
НАН України, директор.

Офіційні опоненти: доктор технічних наук, професор,
член-кореспондент НАН України
Стоян Юрій Григорович,
Інститут проблем машинобудування
імені А.М. Підгорного НАН України
завідувач відділу,

доктор фізико-математичних наук, професор
Моклячук Михайло Павлович,
Київський національний університет
імені Тараса Шевченка, професор,

доктор фізико-математичних наук, доцент
Касьянов Павло Олегович,
Навчально-науковий комплекс
«Інститут прикладного системного аналізу»
Національного технічного університету України
"КПІ" НАН України та МОНМС України,
завідувач науково-дослідного відділу.

Захист відбудеться "____"________ 2013 р. о(об) ____ годині на засіданні спеціалізованої вченої ради Д 26.194.02 при Інституті кібернетики імені В.М. Глушкова НАН України за адресою:
03680, МСП, Київ-187, проспект Академіка Глушкова, 40.

З дисертацією можна ознайомитися в науково-технічному архіві
Інституту кібернетики імені В.М. Глушкова НАН України
за адресою:
03680, МСП, Київ-187, проспект Академіка Глушкова, 40.

Автореферат розісланий "____"____________2013 р.

Учений секретар
спеціалізованої вченої ради Вагіс О.А.

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. До мінімізації опуклих функцій з розривним градієнтом зводиться велика кількість проблем, що виникають при розв'язуванні складних задач математичного програмування. Володіння методами недиференційовної (негладкої) оптимізації дає можливість гнучко використовувати схеми декомпозиції з урахуванням специфіки задач великої розмірності, дозволяє ефективно отримувати двоїсті оцінки для задач дискретного й неперервно-дискретного програмування та для деяких класів багатоекстремальних задач. З'являється можливість використовувати негладкі функції штрафу, які дозволяють при скінченних значеннях штрафних параметрів отримувати задачу безумовної мінімізації, еквівалентну початковій задачі опуклого програмування. Техніко-економічні характеристики об'єктів, що підлягають оптимізації, зазвичай добре апроксимуються кусочно-гладкими функціями, і це також приводить до постановок задач оптимізації з негладкими функціями. Однак відсутність ефективних методів негладкої оптимізації ускладнювала розв'язування зазначених класів задач та змушувала або змінювати формулювання задачі, що могло погіршити відповідність моделі реальності, або використовувати різні прийоми згладжування функцій. Останнє не завжди приводить до успіху, тому що застосування згладження погіршує обумовленість функції та знижує обчислювальну стійкість навіть таких ефективних методів гладкої мінімізації, як квазиньютонівські та методи спряжених градієнтів.

Отже, область застосувань методів негладкої оптимізації досить широка, а розробка обчислювальних методів негладкої оптимізації – актуальна проблема. Гладка функція з дуже швидко змінюваним градієнтом близька за своїми властивостями до негладкої функції. Тому обчислювальні методи, розроблені для розв’язання задач негладкої оптимізації, виявляються ефективними і для оптимізації яружних гладких функцій.

В Інституті кібернетики імені В.М. Глушкова НАН України під керівництвом академіка Н.З. Шора розроблені сімейства алгоритмів негладкої оптимізації: методи узагальненого градієнтного спуску (субградієнтні методи); субградієнтні методи з розтягом простору в напрямку субградієнта (світове визнання отримав їх окремий випадок – метод еліпсоїдів); субградієнтні методи з розтягом простору в напрямку різниці двох послідовних субградієнтів (вони отримали назву r-алгоритмів). Як за надійністю, так і за часом розрахунків і точністю результатів, r-алгоритми виявилися конкурентноздатними у порівнянні з найбільш ефективними методами розв'язання гладких, але погано обумовлених задач. r-алгоритми дали змогу розв'язувати задачі опуклого програмування великої розмірності (з використанням алгоритмів декомпозиції), мінімаксні та матричні задачі негладкої оптимізації та інші. Проте теоретичне обгрунтування їх збіжності є недостатнім. Ще в 1982 році в одній із робіт Н.З. Шор та В.І. Гершович написали: "Теория всего класса алгоритмов с растяжением пространства далека от совершенства. Нам кажется достаточно реалистичной целью построение такого алгоритма, который по своей практической эффективности не уступал бы r-алгоритму и был столь же хорошо обоснован как метод эллипсоидов".

Особливим джерелом задач недиференційовної оптимізації є використання лагранжевих двоїстих оцінок цільової функції у неопуклих квадратичних моделях. Такі моделі зустрічаються в багатьох практичних задачах, до них зводяться задачі оптимізації з поліноміальними функціями, важливі класи NP-складних комбінаторних задач та екстремальних задач із теорії графів. Серед останніх слід відмітити задачі знаходження максимальної зваженої незалежної множини вершин графу, максимального розрізу графу, оптимальної бісекції графу, мінімального розбиття графу на k частин з фіксованим числом вершин у кожній частині тощо.

Дослідження цього класу задач проводились під керівництвом Н.З. Шора, починаючи з 1985 року, пізніше вони отримали назву техніки лагранжевих двоїстих оцінок для багатоекстремальних квадратичних задач. Лагранжеві оцінки можна поліпшувати, додаючи в модель так звані функціонально надлишкові квадратичні обмеження. Алгоритми знаходження лагранжевих двоїстих оцінок є альтернативою використанню методів внутрішніх точок для розв'язання задач напіввизначеного програмування (semidefinite programming), які характеризуються умовою невід'ємної визначеності деякої симетричної $(n \times n)$-матриці. Багато задач напіввизначеного програмування доцільно розглядати, як окремий випадок задач недиференційовної оптимізації. Дійсно, якщо елементи $(n \times n)$-матриці $X(u)$ є лінійними функціями від вектора невідомих параметрів $u \in R^m$, то умова невід'ємної визначеності матриці $X(u)$ еквівалентна опуклому негладкому обмеженню $\varphi(u) = -\lambda_n(X(u)) \le 0$, де $\lambda_n(X(u))$ – мінімальне власне число матриці $X(u)$.

Лагранжеві оцінки відіграють важливу роль для NP-складних задач, вони дають можливість виділити такі їх підкласи, які можна розв'язати за поліноміальний час. Надзвичайно цікавими виявилися результати, отримані Н.З. Шором для NP-складної задачі знаходження зваженого числа стійкості (незалежності) графу. Ця задача є однією із центральних у теорії інформації та кодування, проектуванні різних пристроїв за певних умов несумісності обмежень. Вона тісно пов'язана з відомими задачами вибору, розбиття множин, розфарбування графів і іншими комбінаторними задачами, що мають багато застосувань. Двоїсті оцінки Шора для зазначеної задачі тісно пов'язані з відомими числами Ловаса $\vartheta(G,w)$ та $\vartheta'(G,w)$, які мають важливе значення при обгрунтуванні результатів про поліноміальну розв'язуваність ряду задач у теорії графів.

Таким чином, розвиток методів мінімізації негладких опуклих функцій та техніки лагранжевих двоїстих оцінок є актуальним. Він пов'язаний з численними застосуваннями обох підходів при розвязанні складних задач оптимізації. Ці підходи є взаємопов'язаними, тому що задачі знаходження лагранжевих оцінок, як правило, зводяться до мінімізації негладких опуклих функцій і можуть успішно розв'язуватися сучасними методами недиференційовної оптимізації.

Зв'язок роботи з науковими програмами, планами, темами. Робота виконувалась у відповідності з планами наукових досліджень відділу № 120 Інституту кібернетики імені В.М. Глушкова НАН України в рамках наукових тем, в яких автор був керівником чи відповідальним виконавцем: «Розробка нових методів розв'язання складних дискретних і багатоекстремальних задач оптимізації та їх застосування» (номер держреєстрації 0102U003213, 2002–2006), «Розробка та обґрунтування нових ефективних чисельних методів розв'язування складних задач оптимізації» (номер держреєстрації 0104U000276, 2004–2007), «Розробити математичні та програмні засоби розв'язування деяких класів структурованих задач оптимізації» (номер держреєстрації 0108U000280, 2008–2012).

Мета і завдання дослідження. Метою дисертаційної роботи є подальший розвиток теорії та методів негладкої оптимізації, який пов'язаний із вирішенням наступних завдань:

– розробити та теоретично обгрунтувати субградієнтні алгоритми мінімізації негладких опуклих функцій, що базуються на використанні лінійних неортогональних перетворень простору;

– розробити та теоретично дослідити різного роду функціонально надлишкові обмеження у квадратичних екстремальних задачах, вивчити їх вплив на уточнення двоїстих оцінок цільової функції в квадратичних булевих та бінарних задачах;

– програмно реалізувати, експериментально дослідити субградієнтні алгоритми для задач математичного програмування та алгоритми знаходження лагранжевих двоїстих оцінок для квадратичних екстремальних задач;

– провести апробацію розроблених методів для визначення оптимальних параметрів різного роду систем (економічних, транспортних, енергетичних та інших).

Об'єкт дослідження – задачі негладкої оптимізації, багатоекстремальні квадратичні задачі, екстремальні задачі на графах.

Предмет дослідження – субградієнтні методи пошуку мінімуму негладких яружних функцій та алгоритми уточнення двоїстих оцінок в булевих та бінарних квадратичних задачах.

Методи дослідження включають опуклий аналіз та теорію двоїстості, лінійні неортогональні перетворення простору, еліпсоїдальну та поліедральну апроксимації, обчислювальний експеримент.

Наукова новизна отриманих результатів.

1. Розроблено наближений та прискорений методи еліпсоїдів – нові модифікації методу еліпсоїдів. Обгрунтована їх збіжність зі швидкістю геометричної прогресії, в якій знаменник залежить тільки від вимірності простору. Досліджено зв'язок створених методів з методом еліпсоїдів Юдіна – Немировського – Шора.

2. Розроблено та обгрунтовано два нові субградієнтні методи з перетворенням простору для знаходження точки мінімуму опуклої функції при відомому мінімальному значенні функції. Вони використовують крок Поляка в перетвореному просторі змінних та характеризуються монотонним зменшенням відстані до точки мінімуму в послідовно перетворених просторах змінних.

3. Запропоновано три нові способи побудови функціонально надлишкових обмежень у квадратичних екстремальних задачах з булевими та бінарними змінними. Перші два способи генерують квадратичні рівності, які відповідають введенню нових змінних у формі добутку вже існуючих змінних. Третій спосіб для бінарних змінних генерує квадратичні нерівності, які є наслідком того що квадрат суми "непарної" кількості ± 1 не менше одиниці.

4. На основі перших двох способів побудови функціонально надлишкових обмежень отримано уточнені лагранжеві двоїсті оцінки для числа стійкості та зваженого числа стійкості графу, для зваженого максимального розрізу графу, для задачі максимізації квадратичної функції з бінарними або булевими змінними. На основі третього способу побудована поліедральна апроксимація зверху бінарного квадратичного многогранника, яка при двох, трьох та чотирьох змінних є точною.

5. Знайдено нові властивості найкращої із оцінок Шора для зваженого числа стійкості графу. Вони пов'язані з таким підграфом, як p-колесо, і на його основі побудовано нове сімейство графів, назване W_p-досконалими графами. Для них найкраща оцінка Шора є точною і зважене число стійкості може бути знайдено за поліноміальний час.

Практичне значення отриманих результатів. Розроблені методи та алгоритми можна застосовувати для розв’язання складних проблем оптимізації. У роботі на їх основі отримано результати.

1. Розроблено алгоритм для знаходження L_p-розв'язку перевизначеної системи лінійних рівнянь при двосторонніх обмеженнях на компоненти розв'язку, алгоритм пошуку допустимої точки опуклої нерівності та три ітераційні методи розв'язання системи лінійних алгебраїчних рівнянь.

2. Для зваженого числа стійкості графу розроблено та програмно реалізовано ітераційний алгоритм знаходження верхніх оцінок, в якому многогранник стійких множин апроксимується задачею лінійного програмування з поетапним доповненням її скінченним числом обмежень, що відповідають порушеним лінійним нерівностям для непарних циклів та p-коліс. Алгоритм перевірено для графів, які містять від кількох сотень до кількох тисяч вершин.

3. Розроблено математичні моделі, методи та програмне забезпечення для розв'язання таких задач: задачі про максимальний розріз зваженого графа з заданими кількостями вершин у підмножинах; задачі пошуку глобального мінімуму квадратичної функції на многовидах Штіфеля; задачі знаходження оптимальних нормованих структур кінцевого випуску та доданої вартості у продуктивній моделі Леонтьєва; задач знаходження пропускних здатностей дуг надійної орієнтованої мережі з передачею потоків довільними шляхами та з передачею потоків по заданій множині допустимих шляхів; задач визначення оптимальних параметрів плоских багатошарових оптичних покриттів; задачі знаходження електричних навантажень енергоблоків в енергосистемі з можливістю керування завантаженням (маневреністю) окремого сімейства енергоблоків.

Особистий внесок здобувача. Всі наукові результати дисертаційної роботи отримані особисто або за участі автора. Шістнадцять робіт опубліковано без співавторів. У колективних монографіях дисертанту належать: в [1] – розділи 1.4 та 4.1, в [2] – розділ 3, в [3] – розділи 1, 2 та 3. У статтях, що написані в співавторстві, дисертанту належить: [8] – розробка та обгрунтування алгоритму; [11] – розробка та обгрунтування стартових умов для методу; [12] – розробка квадратичних моделей; [13–14] – розробка математичних моделей та аналітичного способу обчислення градієнтів; [17–19, 25, 29] – формулювання квадратичних задач та їх релаксованих аналогів; [20–22] – формулювання задач та теоретичне обгрунтування алгоритмів; [24] – постановка математичних моделей; [26] – формулювання негладкої задачі та доведення теореми; [23, 27, 28, 30, 31] – формулювання квадратичної задачі та доведення теорем для продуктивної матриці Леонтьєва, аналіз результатів тестових експериментів; [32] – опис методів та алгоритмів для трьох ідей Шора; [34] – опис обчислювальних експериментів. По статтях [38–41, 43] зроблено пленарні доповіді на міжнародних конференціях.

Апробація результатів дисертації. Основні ідеї, принципи, положення і результати досліджень пройшли апробацію на профільних семінарах Інституту кібернетики імені В.М. Глушкова НАН України та доповідались на конференціях, серед яких: "Recent Advances in Non-Differentiable Optimization" (Київ, 2000, 2001); XII Байкальська міжнародна конференція (Іркутськ, Росія, 2001); "Питання оптимізації обчислень" (Кацивелі, 2001, 2007, 2009); Міжнародна конференція, присвячена 65-річчю з дня народження Б.М. Пшеничного (Київ, 2002); "Теорія прийняття рішень" (Ужгород, 2002, 2004, 2008); "Математическое программирование и приложения" (Єкатеринбург, Росія, 2003, 2011); 11-а Міжнародна конференція по автоматичному управлінню, (Київ, 2004), "Systems Analysis, Data Mining and Optimization in Biomedicine" (Gainesville, USA, 2005); International Conference in Honor of the 70th Birthday of Academician I. V. Sergienko "Applied Optimization and Metaheuristic Innovations" (Ялта, 2006); "Математическое моделирование, оптимизация и информационные технологии" (Кишинэу, Молдова, 2008, 2010, 2012); "Discrete and Global Optimization" (Ялта, 2008); Міжнародна конференція "50 років Інституту кібернетики ім. В.М. Глушкова НАН України" (Київ, 2008); "Дискретная оптимизация и исследование операций" (Новосибірськ, Росія, 2010); Кримська осіння математична школа (КРОМШ) (Севастополь, 2010, 2011, 2012); "Математичні методи, моделі та інформаційні технології в економіці" (Чернівці, 2011); Міжнародна конференція, присвячена 90-річчю від дня народження академіка М.М. Яненка "Современные проблемы прикладной математики и механики: теория, эксперимент и практика" (Новосибірськ, Росія, 2011); II International Conference "Optimization and Applications (OPTIMA-2011)" (Petrovac, Montenegro, 2011); "Статистика, моделирование, оптимизация" (Челябінськ, Росія, 2011); 10th EUROPT Workshop on Advances in Continuous Optimization (Siauliai, Lithuania, 2012); 25th European Confrerence on Operational Reserch (Vilnius, Lithuania, 2012).

Публікації. Основні результати дисертації опубліковані у трьох монографіях [1–3], 30 статтях у фахових виданнях згідно з Переліком фахових видань України [4–33], трьох статтях у інших виданнях [34–36] та семи статтях у матеріалах міжнародних конференцій [37–43].

Структура та обсяг роботи. Дисертаційна робота складається із вступу, семи розділів, висновків та списку використаних джерел (210 найменувань). Загальний обсяг роботи становить 305 сторінок, основний текст роботи викладено на 275 сторінках. Робота містить 19 рисунків та 16 таблиць (без додатка).

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** проаналізовано стан наукової проблеми, обґрунтовано актуальність вибраної тематики, сформульовано мету і задачі дослідження, наведено зміст отриманих результатів.

Перший розділ присвячено огляду трьох центральних ідей Н.З. Шора у негладкій оптимізації: узагальненому градієнтному спуску (1962 р.), використанню лінійних неортогональних перетворень простору для покращення обумовленості яружних функцій (1969 р.), двоїстому підходу до отримання та уточнення оцінок цільової функції у неопуклих квадратичних моделях (1985 р.). Наведено опис застосування цих ідей у методах і алгоритмах, розроблених в Інституті кібернетики імені В.М. Глушкова НАН України. Описано зв'язок дисертаційної роботи із розвитком двох ідей Н.З. Шора (1969 та 1985), який полягає у розробці з використанням цих ідей нових субградієнтних алгоритмів із перетворенням простору, нових схем функціонально надлишкових обмежень у квадратичних екстремальних задачах з булевими та бінарними змінними, дослідженні нових властивостей однієї із верхніх оцінок Шора для зваженого числа стійкості графу.

Другий розділ присвячений новим модифікаціям методу еліпсоїдів, які представлені як методи з розтягом простору. Ці модифікації актуальні тим, що збігаються зі швидкістю геометричної прогресії, в якій знаменник залежить тільки від вимірності простору, і тому дозволяють знаходити розв'язки задач із заданою точністю та можуть бути використані для встановлення границь складності алгоритмів розв'язання задач математичного програмування. У розділі відображено зв'язок розроблених модифікацій методу еліпсоїдів із відомим класичним методом еліпсоїдів Юдіна – Немировського – Шора.

У підрозділі 2.1 наведено наближений метод еліпсоїдів, який полягає у наступному. На k-й ітерації апроксимуючий еліпсоїд E_{k+1} об'єму $vol(E_{k+1})$ будується таким чином, що $vol(E_{k+1}) < vol(E_k)$ та коефіцієнт зменшення об'єму

$$Q_n = \frac{vol(E_{k+1})}{vol(E_k)} = \left(1+\frac{1}{n^2}\right)^{n/2}\left(\sqrt{1+\frac{1}{n^2}}-\frac{1}{n}\right) < 1,$$

де n – вимірність простору. Запропонований метод має такі властивості:

1) для великих n коефіцієнт зменшення об'єму Q_n апроксимується такою ж асимптотичною формулою $Q_n \approx 1 - \dfrac{1}{2n}$, як і для аналогічного коефіцієнта в методі еліпсоїдів Юдіна – Немировського – Шора;

2) якщо $n = 1$, то $Q_1 = 2 - \sqrt{2} \approx 0.5858$, і наближений метод еліпсоїдів близький до методу дихотомії, який є аналогом методу еліпсоїдів в одновимірному випадку.

Збіжність наближеного методу еліпсоїдів обгрунтована теоремами 2.1 та 2.2, а результати аналізу його ефективності у порівнянні з методом еліпсоїдів для задач невеликих розмірів $2 \le n \le 10$ представлені у таблиці 2.1. У ній наведено кількість ітерацій, потрібну кожному з двох методів, для відшукання мінімуму опуклої функції з відносною точністю, яка дорівнює 10^{-10}, а також відношення коефіцієнтів зменшення об'єму для цих методів. З'ясовано, що для розв'язання задач, розміри яких належать зазначеному діапазону, все одно, який з двох методів використовувати.

У підрозділі 2.2 описано застосування наближеного методу еліпсоїдів до задачі знаходження L_p-розв'язків перевизначеної системи m лінійних рівнянь при двосторонніх обмеженнях на n компонент невідомого вектора параметрів ($m > n$). На основі наближеного методу еліпсоїдів розроблено та обгрунтовано алгоритм знаходження L_p-розв'язків, причому "настройка" всіх параметрів алгоритму виконується автоматично. Алгоритм використовує розтяг простору в напрямку векторів спеціально побудованого градієнтного поля. Якщо точка належить допустимій області, то вибирається субградієнт цільової функції у цій точці, а якщо точка знаходиться поза межами допустимої області, то вибирається субградієнт до максимально порушеного двостороннього обмеження на компоненти невідомого вектора.

Показано, що запропонований алгоритм знаходить L_p-розв'язок із заданою точністю за скінченну кількість ітерацій, яка квадратично залежить від n (теорема 2.3). Алгоритм можна успішно застосовувати для оцінки параметрів за критерієм найменших квадратів, критерієм найменших модулів, або мінімаксним чебишевським критерієм, якщо $n \le 10$ та $m \sim 10000$. Подібні задачі часто зустрічаються в різних областях прикладної математики: при обробці результатів спостережень, побудові та аналізі різного роду моделей (фізичних, біологічних, економічних, соціальних та ін.), при пошуку компромісних рішень у моделях з суперечливими даними тощо.

У підрозділі 2.3 наведено метод еліпсоїдів, прискорений за рахунок використання еліпсоїда мінімального об'єму, описаного навколо кульового шару (тіла, отриманого в результаті перетину кулі та множини всіх точок між двома паралельними гіперплощинами, одна з яких проходить через центр кулі, а відстань між площинами менше радіуса кулі). Прискорений метод еліпсоїдів реалізовано наступним чином. Якщо на поточній ітерації можна побудувати кульовий шар (в напрямку обчисленої на цій ітерації відсікаючої гіперплощини), використовуючи інформацію про відсікаючу гіперплощину із попередньої ітерації, то переходимо в центр мінімального за об'ємом еліпсоїда, описаного навколо цього кульового шару. Якщо таку побудову здійснити не вдається, то переходимо в центр мінімального за об'ємом еліпсоїда, що містить півкулю, тобто поступаємо так, як у методі еліпсоїдів.

Для прискореного методу еліпсоїдів доведена теорема 2.4, яка гарантує швидкість його збіжності не гіршу, ніж швидкість збіжності методу еліпсоїдів Юдіна – Немировського – Шора. У випадку, коли на ітераціях часто отримуємо "хороші" кульові шари, то швидкість збіжності прискореного методу еліпсоїдів значно більша, ніж у методу еліпсоїдів. Це підтверджено результатами обчислювальних експериментів для двох тестових прикладів, пов'язаних із пошуком точок мінімуму опуклих квадратичної та кусочно-лінійної функцій. Для яружних функцій наближений метод еліпсоїдів витрачає значно меншу кількість ітерацій, ніж метод еліпсоїдів. Для функції від двадцяти змінних такий виграш оцінюється у шістьвісім разів.

У підрозділі 2.4 подана геометрична інтерпретація ітераційних процесів, які використовують описані еліпсоїди, що містять перетин кулі та гіперплощини. На основі оператора розтягу простору ці процеси дають змогу будувати монотонні за об'ємом методи для знаходження стаціонарних точок у спеціальних задачах. Їх "граничні" варіанти є проективними методами, яким відповідає використання нескінченного коефіцієнта розтягу; вони знаходять розв'язок спеціальних задач не більше, ніж за n ітерацій, де n – вимірність простору змінних.

Наведено три приклади задач, для яких можуть бути використані проективні методи еліпсоїдів. Перший приклад пов'язаний із пошуком розв'язку системи лінійних рівнянь. У цьому випадку проективний метод можна представити як метод із розтягом простору в напрямках нормалей до гіперплощин у перетвореному просторі змінних. Послідовність вибору гіперплощин конкретизує ітераційний процес. Другий приклад пов'язаний із знаходженням точки мінімуму x^* опуклої функції $f(x)$, для якої виконується рівність $m(f(x)-f^*)=(x-x^*,\partial f(x))$ при всіх $x\in E^n$ та заданому m. (Тут і далі E^n – евклідів простір вимірності n із скалярним добутком (x,y)). За умови апріорного знання $f^*=f(x^*)$ проективний метод еліпсоїдів конкретизується набором точок, в яких будуть обчислюватися субградієнти $\partial f(x)$. Розтяг простору буде проводитися в напрямку чергового субградієнта у перетвореному просторі змінних. Третій приклад пов'язаний із знаходженням стаціонарних точок квадратичної функції, для якої проективні методи не вимагають апріорного знання f^*. Структура цих методів буде визначатися набором точок, в яких буде обчислюватися градієнт, та набором довільних лінійно незалежних напрямків. Операція розтягу простору змінних реалізується в напрямку різниці двох послідовних градієнтів, другий з яких обчислюється в точці, де досягається екстремум функції за обраним напрямком руху з першої точки.

У **третьому розділі** розглянуті субградієнтні методи з перетворенням простору для мінімізації яружних функцій. У центрі уваги тут два субградієнтні методи з перетворенням простору для знаходження точки мінімуму опуклої функції при відомому мінімальному значенні функції. Вони використовують крок Поляка в перетвореному просторі змінних та гарантують монотонне зменшення відстані до точки мінімуму в послідовно перетворених просторах змінних. Перетворення простору реалізуються за допомогою однорангового еліпсоїдального оператора. Трудомісткість однієї ітерації методів така ж, як і в r-алгоритмів. Ці методи можна розглядати як удосконалення запропонованого Б.Т. Поляком в 1969 році субградієнтного методу. В них довжина кроку в напрямку антисубградієнта обчислюється (а не підбирається як в r-алгоритмах за допомогою розв'язання одновимірної задачі мінімізації функції за напрямком), причому для її обчислення потрібно лише знання мінімального значення функції.

У підрозділі 3.1 розглянуто субградієнтний метод Поляка та проведено аналіз його повільної збіжності для яружних функцій. Нехай $f(x)$ – опукла функція та її субградієнт $\partial f(x)$ задовільняє умові:

$$(x-x^*,\partial f(x)) \geq m(f(x)-f^*) \quad \forall x\in E^n, \quad m\geq 1. \tag{1}$$

Тут x^* – єдина точка мінімуму функції $f(x)$; f^* – мінімальне значення функції $f(x)$: $f^*=f(x^*)$. Параметр m вважається відомим, він введений для того, щоб метод можна було застосовувати для мінімізації спеціальних класів опуклих функцій, наприклад, для квадратичної функції при $m=2$. При значенні $m=1$ забезпечується виконання умови (1) для довільної опуклої функції.

Нехай відомо значення f^*. Метод Поляка є ітераційним процесом вигляду:

$$x_{k+1}=x_k-h_k\frac{\partial f(x_k)}{\|\partial f(x_k)\|}, \quad h_k=\frac{m\left(f(x_k)-f^*\right)}{\|\partial f(x_k)\|}, \quad k=0,1,2,\ldots, \tag{2}$$

призначеним для знаходження точки x^*. Величину h_k називають кроком Поляка або кроком Агмона – Моцкіна – Шенберга. Вперше такий крок був використаний у незалежних роботах Agmon (1964 р.) та Motzkin, Schoenberg (1964 р.) у релаксаційному (субградієнтному) методі пошуку хоча б одного розв'язку сумісної системи лінійних нерівностей. Величина h_k гарантує зменшення відстані до точки x^* на кожній ітерації методу.

При $m=1$ метод (2) має такий геометричний зміст. Функція $f(x)$ апроксимується лінійною функцією $\tilde f(x)=f(x_k)+(f'(x_k),x-x_k)$, і крок h_k вибирається таким, щоб ця апроксимуюча функція стала рівною f^*, тобто $\tilde f(x_{k+1})=f^*$. Для одновимірного випадку метод Поляка збігається із методом Ньютона для розв'язання рівняння $f(x)=f^*$.

Теорема 3.1. *Для всіх точок, що генеруються методом* (*2*), *справедливі нерівності*

$$\|x_{k+1}-x^*\|^2 \leq \|x_k-x^*\|^2-\left(\frac{m(f(x_k)-f^*)}{\|\partial f(x_k)\|}\right)^2, \quad k=0,1,\ldots,.$$

Теорема 3.1 відрізняється від теореми Поляка (1969) розширенням застосування методу Поляка для спеціальних класів функцій, у яких $m>1$.

Швидкість збіжності метода Поляка для яружних функцій пов'язана з величиною кута між двома послідовними субградієнтами: $\partial f(x_k)$ и $\partial f(x_{k+1})$. Чим ближче цей кут до 180°, тим більш повільною буде швидкість збіжності. Якщо на кожній ітерації позбавлятися від тупого кута між послідовними субградієнтами у перетвореному просторі змінних, то для яружних функцій слід очікувати підвищення швидкості збіжності методу.

У підрозділі 3.2 подано два таких методи, які реалізовані за наступним правилом. Якщо кут між двома векторами, що визначають гіперплощини локалізації точки x^*, є тупим, то простір перетворюється так, щоб цей кут став прямим у перетвореному просторі змінних. Це реалізується за допомогою однорангового еліпсоїдального оператора, який у матричному вигляді подамо формулою

$$T_1(\xi,\eta)=I-\frac{1}{1-(\xi,\eta)^2}\left(\left(1-\sqrt{1-(\xi,\eta)^2}\right)\eta-(\xi,\eta)\xi\right)\eta^T.$$

Тут $\xi,\eta \in E^n$ – вектори, такі що $\|\xi\|=1$, $\|\eta\|=1$ та їх скалярний добуток задовільняє умові $(\xi,\eta)^2 \neq 1$, I – одинична матриця розміру $n\times n$. Для оператора $T_1(\xi,\eta)$ існує обернений:

$$T_1^{-1}(\xi,\eta) = I + \frac{1}{\sqrt{1-(\xi,\eta)^2}}\left(\left(1-\sqrt{1-(\xi,\eta)^2}\right)\eta - (\xi,\eta)\xi\right)\eta^T .$$

Застосування оператора $T_1(\xi,\eta)$ забезпечує перетворення в кулю спеціального еліпсоїда, описаного навколо тіла W, яке отримано в результаті перетину кулі та двох напівпросторів, які проходять через центр кулі. У випадку тупого кута між нормалями напівпросторів цей еліпсоїд містить тіло W та є мінімальним за об'ємом у сімействі описаних еліпсоїдів, центр яких співпадає з центром кулі.

Якщо φ – кут між векторами ξ та η тупий, то спеціальний еліпсоїд $Ell(x_0,a,b,r)$ містить тіло W (рис. 1, a). Його об'єм менший, ніж об'єм кулі $S(x_0,r)$, і коефіцієнт зменшення об'єму дорівнює $\sqrt{1-(\xi,\eta)^2} = \sqrt{1-(\cos\varphi)^2}$. Перетворити еліпсоїд $Ell(x_0,a,b,r)$ в кулю можна за допомогою оператора $T_1(\xi,\eta)$. У перетвореному просторі еліпсоїд стане кулею $S(y_0,r)$, а образи векторів ξ та η стануть взаємно ортогональними (рис. 1, b).

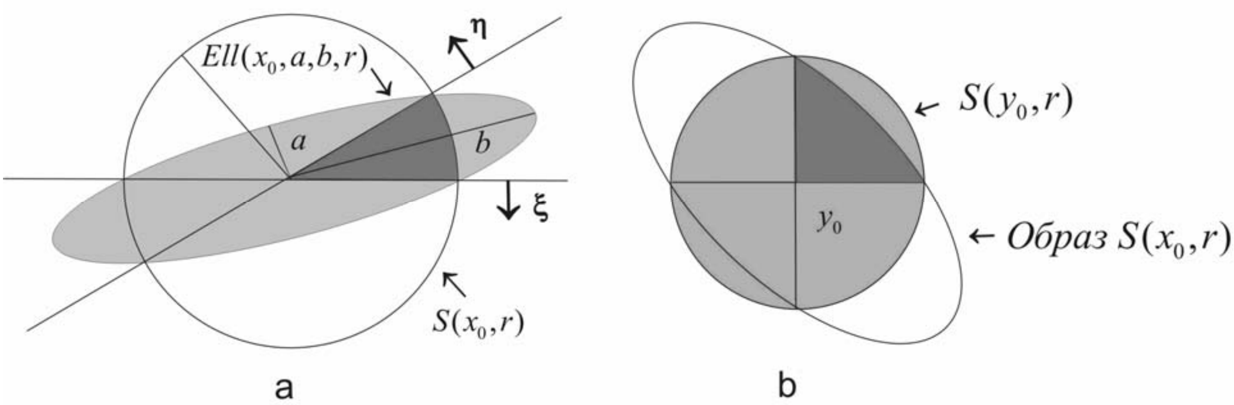

Рис. 1. Еліпсоїд $Ell(x_0,a,b,r)$ до і після перетворення простору

Використання оператора $T_1(\xi,\eta)$ у субградієнтному процесі дозволяє "розширити" конус напрямків зменшення функції у перетвореному просторі змінних, подібно тому як це відбувається в r-алгоритмах. Однак, при цьому розтяг простору виконується у напрямку різниці двох нормованих субградієнтів, і близьким до напрямку різниці двох субградієнтів він буде тільки тоді, коли норми субградієнтів близькі. Дійсно, для того щоб перетворити спеціальний еліпсоїд у кулю потрібно розтягнути простір у напрямку $\dfrac{\xi-\eta}{\|\xi-\eta\|}$ з коефіцієнтом $\alpha_1 = \dfrac{1}{\sqrt{1+(\xi,\eta)}} > 1$ та стиснути простір в ортогональному напрямку $\dfrac{\xi+\eta}{\|\xi+\eta\|}$ з коефіцієнтом $\alpha_2 = \dfrac{1}{\sqrt{1-(\xi,\eta)}} < 1$.

У першому методі перетворення простору спрямовано на те, щоб тупий кут між субградієнтами $\partial f(x_k)$ та $\partial f(x_{k+1})$ у перетвореному просторі змінних зробити прямим. Цей підхід реалізовано у методі amsg2, який подамо формулами:

$$x_{k+1} = x_k - h_k B_k \frac{B_k^T \partial f(x_k)}{\left\| B_k^T \partial f(x_k) \right\|}, \quad h_k = \frac{m(f(x_k) - f^*)}{\left\| B_k^T \partial f(x_k) \right\|}, \quad k = 0,1,2,\ldots, \tag{3}$$

де матриця $B_0 = I$, матриця B_{k+1} розмірністю $n \times n$ обчислюється за правилом:

$$B_{k+1} = \begin{cases} B_k, & \text{якщо } \mu_k \geq 0, \\ B_k + \left(B_k \eta\right) \xi_{k+1}^T, & \text{в іншому випадку,} \end{cases} \tag{4}$$

і

$$\mu_k = \xi_k \xi_{k+1}^T, \; \xi_k = \frac{B_k^T \partial f(x_k)}{\left\| B_k^T \partial f(x_k) \right\|}, \; \xi_{k+1} = \frac{B_k^T \partial f(x_{k+1})}{\left\| B_k^T \partial f(x_{k+1}) \right\|},$$

$$\eta = \left(\frac{1}{\sqrt{1-\mu_k^2}} - 1 \right) \xi_{k+1} - \frac{\mu_k}{\sqrt{1-\mu_k^2}} \xi_k.$$

У назві методу (3), (4) частина "ams" вказує на спосіб регулювання кроку в напрямі нормованого антисубградієнта, а "g2" вказує на те, що ams-крок (Agmon-Motzkin-Schoenberg step, крок Поляка) здійснюється у просторі змінних, перетвореному за допомогою двох послідовних субградієнтів (g2).

Теорема 3.2. *Нехай* $A_k = B_k^{-1}$, $A_{k+1} = B_{k+1}^{-1}$. *Для всіх точок, що генеруються методом* (*3*), (*4*), *справедливі нерівності*

$$\left\| A_{k+1}(x_{k+1} - x^*) \right\|^2 \leq \left\| A_k (x_k - x^*) \right\|^2 - \left(\frac{m(f(x_k) - f^*)}{\left\| B_k^T \partial f(x_k) \right\|} \right)^2, \; k = 0,1,\ldots,.$$

Теорема 3.2 гарантує монотонне зменшення відстані до точки мінімуму в послідовно перетворених просторах змінних.

Метод amsg2 фактично є методом Поляка, прискореним за рахунок антияружного прийому, подібного тому, який використано в r-алгоритмах Шора. Це надало методу ефективності при мінімізації яружних функцій. Так, наприклад, для кусочно-лінійної функції $f_1(x_1, x_2) = |x_1| + t|x_2|$ при будь-якому значенні параметра $t > 1$ та довільній стартовій точці x_0 метод amsg2 знаходить точку мінімуму $x^* = (0,0)$ не більше, ніж за три ітерації. Проте метод Поляка збігається до x^* зі швидкістю геометричної прогресії із знаменником $\sqrt{1 - 1/t^2}$ і тому вимагає великої кількості ітерацій навіть при порівняно невеликих значеннях t.

Прискореним за відношенням до попереднього є метод amsg2p, у якому ams-крок використовується у просторі змінних, перетвореному за допомогою двох послідовних субградієнтів (g2) та агрегатного вектора (p). Перетворення простору реалізується за допомогою однорангового еліпсоїдального оператора і тільки на тих ітераціях методу, коли тупим є хоча би один із двох кутів: кут між двома послідовними субградієнтами або кут між останнім субградієнтом та агрегатним вектором, який є опуклою комбінацією обчислених на попередніх ітераціях субградієнтів. Якщо поточний агрегатний вектор не дозволяє зменшити об'єм еліпсоїда локалізації x^*, то для наступної ітерації він автоматично оновлюється.

Метод amsg2p для довільного значення $f_{min} \ge f^*$ знаходить таку точку, де значення функції $f(x)$ менше або дорівнює $f_{min} + \varepsilon$, або гарантує, що точка, де значення $f(x)$ дорівнює f_{min}, не існує у кулі заданого радіуса. У першому випадку метод amsg2p знаходить точку $x_\varepsilon^* \in \{x : f(x) - f_{min} \le \varepsilon\}$ та відповідний їй номер ітерації k_ε^*, а у другому – зупиняється з повідомленням "точки не існує". Метод amsg2p полягає в наступному.

На ітерації $k = 0$ задамо: початкове наближення $x_0 \in E^n$, початковий радіус r_0 такий, що $\|x_0 - x^*\| \le r_0$, та досить мале $\varepsilon > 0$. Обчислимо $f(x_0)$ та $\partial f(x_0)$. Якщо $f(x_0) - f_{min} \le \varepsilon$, то $x_\varepsilon^* = x_0$, $k_\varepsilon^* = 0$ і роботу алгоритму закінчено. Інакше покладемо $h_0 = \dfrac{m(f(x_0) - f_{min})}{\|\partial f(x_0)\|}$, $\xi_0 = \dfrac{\partial f(x_0)}{\|\partial f(x_0)\|} \in E^n$, $p_0 = 0 \in E^n$, $B_0 = I$ – одинична $n \times n$– матриця. Перейдемо до наступної ітерації.

Нехай на k-й ітерації отримані $x_k \in E^n$, h_k, r_k, $\xi_k \in E^n$, $p_k \in E^n$, B_k – матриця $n \times n$. Для $(k+1)$-ї ітерації виконаємо пп. 1–5.

1. Обчислимо $t_k = h_k / r_k$. Якщо $t_k > 1$, то "точки не існує", і закінчення роботи алгоритму. Інакше покладемо $r_{k+1} = r_k \sqrt{1 - t_k^2}$ і обчислимо чергове наближення

$$x_{k+1} = x_k - h_k B_k \xi_k.$$

2. Обчислимо $f(x_{k+1})$ і $\partial f(x_{k+1})$. Якщо $f(x_{k+1}) - f_{min} \le \varepsilon$, то $x_\varepsilon^* = x_{k+1}$, $k_\varepsilon^* = k+1$ і закінчення роботи алгоритму. Інакше покладемо

$$\xi_{k+1} = \frac{B_k^T \partial f(x_{k+1})}{\|B_k^T \partial f(x_{k+1})\|}, \quad h_{k+1} = \frac{m(f(x_{k+1}) - f_{min})}{\|B_k^T \partial f(x_{k+1})\|}.$$

3. Обчислимо $\lambda_1 = -p_k^T \xi_{k+1}$ та $\lambda_2 = -\xi_k^T \xi_{k+1}$. Покладемо

$$p_{k+1} = \begin{cases} \dfrac{\lambda_1}{\sqrt{\lambda_1^2 + \lambda_2^2}} p_k + \dfrac{\lambda_2}{\sqrt{\lambda_1^2 + \lambda_2^2}} \xi_k, & \text{якщо } \lambda_1 > 0 \text{ и } \lambda_2 > 0, \\ p_k, & \text{якщо } \lambda_1 > 0 \text{ и } \lambda_2 \le 0, \\ \xi_k, & \text{якщо } \lambda_1 \le 0 \text{ и } \lambda_2 > 0, \\ 0, & \text{якщо } \lambda_1 \le 0 \text{ и } \lambda_2 \le 0. \end{cases}$$

4. Обчислимо $\mu_k = p_{k+1}^T \xi_{k+1}$. Якщо $-1 < \mu_k < 0$, то обчислимо

$$B_{k+1} = B_k + (B_k \eta) \xi_{k+1}^T, \text{ де } \eta = \left(\frac{1}{\sqrt{1 - \mu_k^2}} - 1 \right) \xi_{k+1} - \frac{\mu_k}{\sqrt{1 - \mu_k^2}} p_{k+1}$$

та перерахуємо

$$h_{k+1} = \frac{h_{k+1}}{\sqrt{1-\mu_k^2}}, \qquad p_{k+1} = \frac{1}{\sqrt{1-\mu_k^2}}(p_{k+1} - \mu_k \xi_{k+1}).$$

Інакше покладемо $B_{k+1} = B_k$ та $p_{k+1} = 0$.

5. Перейдемо до наступної ітерації з x_{k+1}, h_{k+1}, r_{k+1}, ξ_{k+1}, p_{k+1}, B_{k+1}.

Теорема 3.3. *Нехай* $A_k = B_k^{-1}$, $A_{k+1} = B_{k+1}^{-1}$. *Якщо* $f_{min} \ge f^*$, *то для всіх точок, які генеруються методом* amsg2p, *справедливі нерівності*

$$\left\| A_{k+1}(x_{k+1} - x^*) \right\|^2 \le \left\| A_k (x_k - x^*) \right\|^2 - \left(\frac{m(f(x_k) - f_{min})}{\left\| B_k^T \partial f(x_k) \right\|} \right)^2, \quad k = 0,1,2,...,.$$

Теорема 3.3 означає, що на кожній ітераціїї у перетвореному просторі змінних відстань до точки мінімуму зменшується. Завдяки цьому для кожної ітерації $k > 1$ має місце співвідношення

$$\left\| A_k (x_k - x^*) \right\|^2 \le \left\| x_0 - x^* \right\|^2 - \sum_{i=0}^{k-1} \left(\frac{m(f(x_i) - f_{min})}{\left\| B_i^T \partial f(x_i) \right\|} \right)^2 = r_0^2 - \sum_{i=0}^{k-1} h_i^2 = r_k^2,$$

за допомогою якого забезпечується достатня умова відсутності точки x_ε^* при $f_{min} < f^* - \varepsilon$. Це реалізовано в п. 1 методу amsg2p.

Метод amsg2p є конкурентоздатним за відношенням до r-алгоритму Шора – Журбенка, що підтверджено результатами комп'ютерних експериментів, проведених для стандартного набору тестових задач мінімізації яружних гладких та негладких опуклих функцій. Крім цього, метод amsg2p має перевагу над r-алгоритмом, тому що зберігає всі хороші властивості методу Поляка. Якщо поверхні рівня функції не є витягнутими, то метод amsg2p може збігатись всього за декілька ітерацій. Якщо при цьому не виконувалось перетворення простору, то метод amsg2p співпадає з методом Поляка. В роботі це підтверджено тестовими експериментами для квадратичних функцій різної степені яружності (табл. 3.3).

У підрозділі 3.3 наведено опис алгоритму feg2p1 для знаходження допустимої точки опуклої нерівності або перевірки існування такої точки. Алгоритм feg2p1 побудований на основі методу amsg2p, який доповнений критерієм зупинки, подібним до застосованого в п. 1 методу amsg2p. У підрозділі 3.4 алгоритм feg2p1 використано для побудови трьох ітераційних методів розв'язання системи лінійних алгебраїчних рівнянь. Ефективність розроблених методів підтверджена комп'ютерними експериментами з погано обумовленими системами лінійних рівнянь із матрицею Гільберта (табл. 3.4).

У підрозділі 3.5 запропоновано ДФП- та БФШ-оператори – лінійні оператори перетворення простору для двох квазіньютонівських методів: Давідона – Флетчера – Пауела та Бройдена – Флетчера – Шенно. Для цих методів дана геометрична інтерпретація градієнтного кроку в перетвореному просторі змінних. Для мінімізації опуклих функцій ДФП- та БФШ-оператори дають можливість будувати методи з перетворенням простору за типом r-алгоритмів із тим, щоб зменшити ступінь витягнутості поверхонь функції у перетвореному просторі змінних.

Четвертий розділ присвячено різним способам уточнення двоїстих оцінок у квадратичних екстремальних задачах з булевими $(0,1)$ та бінарними (± 1) змінними. Описано їх застосування до задач максимізації квадратичної функції з булевими та бінарними змінними, задачі про число стійкості (незалежності) неорієнтованого графу, задачі про максимальний розріз зваженого графу та ін. Перші два способи генерують усі можливі квадратичні рівності, які для бінарних або булевих квадратичних задач можна побудувати при введенні нових змінних у формі добутку вже існуючих змінних. Третій спосіб генерує всі можливі квадратичні нерівності, які випливають з того, що сума непарної кількості ± 1 не може дорівнювати нулю.

У підрозділі 4.1 розглянуто два способи уточнення двоїстих оцінок за допомогою сімейств функціонально надлишкових квадратичних обмежень-рівностей у бінарних та булевих квадратичних задачах. Їх суть полягає в наступному.

Нехай задано бінарні змінні $y_i^2 = 1, \quad i = 1, \ldots, n.$ Доповнимо їх новими бінарними змінними $y_{ij} = y_i y_j, \; 1 \le i < j \le n$, для кожної з яких справедлива рівність $y_{ij}^2 = 1$. Тоді для розширеної множини бінарних змінних маємо наступні сімейства функціонально надлишкових обмежень.

Для кожної двійки змінних y_i, y_j, де $1 \le i < j \le n$, маємо по два обмеження

$$y_i - y_{ij} y_j = 0, \quad y_j - y_{ij} y_i = 0, \tag{5}$$

які є наслідком рівностей $y_i = y_i y_j^2 = (y_i y_j) y_j = y_{ij} y_j$ та $y_j = y_j y_i^2 = (y_i y_j) y_i = y_{ij} y_i$.

Для кожної трійки змінних y_i, y_j та y_k, де $1 \le i < j < k \le n$, маємо по п'ять функціонально надлишкових квадратичних обмежень. Перші три з них

$$y_{ij} - y_{ik} y_{jk} = 0, \quad y_{ik} - y_{ij} y_{jk} = 0, \quad y_{jk} - y_{ij} y_{ik} = 0 \tag{6}$$

є наслідком очевидних рівностей $y_{ij} = y_i y_j = y_i y_j y_k^2 = (y_i y_k)(y_j y_k) = y_{ik} y_{jk}$, $y_{ik} = y_i y_k = y_i y_k y_j^2 = (y_i y_j)(y_j y_k) = y_{ij} y_{jk}$, $\quad y_{jk} = y_j y_k = y_j y_k y_i^2 = (y_i y_j)(y_i y_k) = y_{ij} y_{ik}$. Обмеження (6) доповнюються ще двома рівностями

$$y_{ij} y_k - y_{ik} y_j = 0, \quad y_{ij} y_k - y_{jk} y_i = 0, \tag{7}$$

які випливають із неоднозначності представлення добутку $y_i y_j y_k$:

$$y_i y_j y_k = (y_i y_j) y_k = y_{ij} y_k = (y_i y_k) y_j = y_{ik} y_j = (y_j y_k) y_i = y_{jk} y_i.$$

Зауважимо, що з неоднозначного представлення $y_i y_j y_k$ випливає три обмеження, проте одне з них є лінійно залежним від двох інших обмежень і не впливає на точність двоїстої оцінки.

Для кожної четвірки змінних y_i, y_j, y_k та y_l, де $1 \le i < j < k < l \le n$, маємо два функціонально надлишкові квадратичні обмеження

$$y_{ij} y_{kl} - y_{ik} y_{jl} = 0, \quad y_{ij} y_{kl} - y_{il} y_{jk} = 0, \tag{8}$$

які лінійно незалежні та випливають із неоднозначності представлення добутку всіх чотирьох змінних $y_i y_j y_k y_l = y_{ij} y_{kl} = y_{ik} y_{jl} = y_{il} y_{jk}$.

Аналогічно для розширеної множини булевих змінних $x_i^2 = x_i$, $i = 1,\ldots,n$, та $x_{ij} = x_i x_j$, $1 \le i < j \le n$, отримуємо наступні сімейства функціонально надлишкових обмежень. Для кожної двійки булевих змінних x_i, x_j маємо по два обмеження

$$x_{ij}x_i - x_{ij} = 0, \quad x_{ij}x_j - x_{ij} = 0, \tag{9}$$

які є наслідком очевидних рівностей

$$0 = x_i^2 - x_i = (x_i^2 - x_i)x_j = x_i(x_i x_j) - x_i x_j = x_i x_{ij} - x_{ij} = 0,$$
$$0 = x_j^2 - x_j = (x_j^2 - x_j)x_i = x_j(x_i x_j) - x_i x_j = x_j x_{ij} - x_{ij} = 0,$$

та справедливі при довільних i та j, таких, що $1 \le i < j \le n$.

Для кожної трійки булевих змінних x_i, x_j та x_k маємо по п'ять функціонально надлишкових квадратичних обмежень. Перші три обмеження

$$x_{ij}x_{ik} - x_i x_{jk} = 0, \quad x_{ij}x_{jk} - x_j x_{ik} = 0, \quad x_{ik}x_{jk} - x_k x_{ij} = 0 \tag{10}$$

є наслідком очевидних рівностей

$$0 = (x_i^2 - x_i)x_j x_k = (x_i x_j)(x_i x_k) - x_i(x_j x_k) = x_{ij}x_{ik} - x_i x_{jk} = 0,$$
$$0 = (x_j^2 - x_j)x_i x_k = (x_i x_j)(x_j x_k) - x_j(x_i x_k) = x_{ij}x_{jk} - x_j x_{ik} = 0,$$
$$0 = (x_k^2 - x_k)x_i x_j = (x_i x_k)(x_j x_k) - x_k(x_i x_j) = x_{ik}xjk - x_k x_{ij} = 0,$$

та справедливі при довільних i, j і k, таких що $1 \le i < j < k \le n$. Враховуючи неоднозначність представлення добутку $x_i x_j x_k$, введемо ще два обмеження, аналогічні обмеженням (7). Для четвірки змінних x_i, x_j, x_k і x_l маємо по два лінійно незалежні обмеження, аналогічні обмеженням (8).

Зазначені сімейства функціонально надлишкових обмежень використано для уточнення двоїстих оцінок у двох квадратичних екстремальних задачах: задачі максимізації квадратичної функції з бінарними змінними та задачі максимізації квадратичної функції з булевими змінними. Для обох квадратичних екстремальних задач побудовано їх аналоги з розширеними множинами змінних та повними сімействами функціонально надлишкових квадратичних обмежень. Показано, що розширеним квадратичним задачам відповідають більш точні верхні оцінки (теорема 4.2). Наведено тестові приклади, в яких нові оцінки є точними.

У підрозділі 4.2 для бінарних змінних описано третій спосіб побудови функціонально надлишкових обмежень, за допомогою якого генеруються квадратичні нерівності на основі властивості непарної кількості бінарних змінних ± 1. Введено бінарний квадратичний многогранник $BiQP_n$ (для опису множини допустимих розв'язків у задачі знаходження екстремуму квадратичної функції від n бінарних змінних) та побудовано його поліедральну апроксимацію зверху. Суть третього способу полягає в наступному.

Назвемо бінарною послідовність $l(k) = (l_1,\ldots,l_k)$, яка складається з $k \ge 2$ бінарних чисел. Будемо використовувати для неї ще і таке позначення: $l(k) \in \{-1,1\}^k$. Розглянемо окремо бінарну послідовність непарної довжини $l(2r+1) = (l_1,\ldots,l_{2r+1})$ та бінарну послідовність парної довжини $l(2r) = (l_1,\ldots,l_{2r})$, де $r = 1,2,\ldots,$.

Лема 4.1. *Для довільної бінарної послідовності непарної довжини* $l(2r+1)=(l_1,\ldots,l_{2r+1})$ *завжди справедлива квадратична нерівність*

$$S_{2r+1}^2=\left(\sum_{i=1}^{2r+1}l_i\right)^2=\left(l_1+l_2+\ldots+l_{2r+1}\right)^2\geq 1. \tag{11}$$

Для заданого r *кількість усіх можливих квадратичних нерівностей* (*11*) *дорівнює* 4^r.

Нехай вектор-рядок $y=(y_1,\ldots,y_k)$ складається з k бінарних змінних.

Лема 4.2. *Для вектора* $y\in\{-1,1\}^{2r+1}$, $r=1,2,\ldots$, *з непарною кількістю бінарних змінних та довільної бінарної послідовності непарної довжини* $l(2r+1)=(l_1,\ldots,l_{2r+1})$ *має місце квадратична нерівність*

$$\text{у першій формі } \left(\sum_{i=1}^{2r+1}l_iy_i\right)^2\geq 1 \text{ і другій формі } \sum_{i=1}^{2r}\sum_{j>i}^{2r+1}l_il_jy_iy_j\geq -r. \tag{12}$$

Для вектора $y\in\{-1,1\}^{2r}$, $r=1,2,\ldots$, *з парною кількістю бінарних змінних,* $l_0\in\{-1,1\}$, *та довільної бінарної послідовності парної довжини* $l(2r)=(l_1,\ldots,l_{2r})$ *справедлива квадратична нерівність*

$$\text{у першій формі } \left(l_0+\sum_{i=1}^{2r}l_iy_i\right)^2\geq 1 \text{ і другій формі } \sum_{i=1}^{2r}l_0l_iy_i+\sum_{i=1}^{2r-1}\sum_{j>i}^{2r}l_il_jy_iy_j\geq -r. \tag{13}$$

Для заданого r *кількість усіх можливих квадратичних нерівностей як вигляду* (*12*), *так і вигляду* (*13*) *дорівнює* 4^r.

Повні сімейства квадратичних нерівностей для двійок та трійок бінарних змінних включають по 4 нерівності. Для двійок бінарних змінних повне сімейство квадратичних нерівностей отримуємо з (13) при $r=1$. Воно містить такі нерівності:

$$\begin{cases}(+1+y_i+y_j)^2\geq 1,\\(-1+y_i+y_j)^2\geq 1,\\(+1-y_i+y_j)^2\geq 1,\\(+1+y_i-y_j)^2\geq 1,\end{cases}\quad\text{або їх аналоги}\quad\begin{cases}+y_i+y_j+y_iy_j\geq -1,\\-y_i-y_j+y_iy_j\geq -1,\\-y_i+y_j-y_iy_j\geq -1,\\+y_i-y_j-y_iy_j\geq -1.\end{cases}$$

Для трійок бінарних змінних повне сімейство квадратичних нерівностей, отримане з (12) при $r=1$, включає такі нерівності:

$$\begin{cases}(+y_i+y_j+y_k)^2\geq 1,\\(-y_i+y_j+y_k)^2\geq 1,\\(+y_i-y_j+y_k)^2\geq 1,\\(+y_i+y_j-y_k)^2\geq 1,\end{cases}\quad\text{або їх аналоги}\quad\begin{cases}+y_iy_j+y_iy_k+y_jy_k\geq -1,\\-y_iy_j-y_iy_k+y_jy_k\geq -1,\\-y_iy_j+y_iy_k-y_jy_k\geq -1,\\+y_iy_j-y_iy_k-y_jy_k\geq -1.\end{cases}$$

Повні сімейства квадратичних нерівностей для четвірок і п'ятірок бінарних змінних містять по $16=4^2$ різних нерівностей. Для четвірок вони випливають із (11) при $r=2$, а для п'ятірок – із (10) при $r=2$. За аналогічною схемою будуються квадратичні нерівності при $r=3$. Вони відповідають парним "шісткам" і непарним "сімкам" бінарних змінних та включають по $64=4^3$ нерівності.

Квадратичні нерівності у першій або в другій формах можна додавати до квадратичних екстремальних задач з бінарними змінними, тим самим покращуючи відповідні новим задачам лагранжеві двоїсті оцінки. Окрім цього, використання другої форми квадратичних нерівностей для парної та непарної кількості бінарних змінних спрощує перехід від квадратичних екстремальних задач з бінарними змінними до їх лінійних аналогів, отриманих релаксацією (ослабленням) бінарних змінних.

Дійсно, другі форми квадратичних нерівностей у (12) та (13) фактично підготовлені для використання змінних $y_{ij} = y_i y_j$, їх можна записати у вигляді лінійних нерівностей

$$\sum_{i=1}^{2r}\sum_{j>i}^{2r+1} l_i l_j y_{ij} \geq -r \quad \text{та} \quad \sum_{i=1}^{2r} l_0 l_i y_i + \sum_{i=1}^{2r-1}\sum_{j>i}^{2r} l_i l_j y_{ij} \geq -r.$$

При $r=1$, що відповідає трійкам та двійкам бінарних змінних, лінійні нерівності набувають наступного вигляду:

$$\begin{cases} +y_{ij} + y_{ik} + y_{jk} \geq -1, \\ -y_{ij} - y_{ik} + y_{jk} \geq -1, \\ -y_{ij} + y_{ik} - y_{jk} \geq -1, \\ +y_{ij} - y_{ik} - y_{jk} \geq -1, \end{cases} \quad \text{та} \quad \begin{cases} +y_i + y_j + y_{ij} \geq -1, \\ -y_i - y_j + y_{ij} \geq -1, \\ -y_i + y_j - y_{ij} \geq -1, \\ +y_i - y_j - y_{ij} \geq -1. \end{cases}$$

У теорії математичного програмування перша четвірка лінійних нерівностей відома як правило трикутника, а друга четвірка – як найпростіше правило для лінеаризації квадратичних булевих задач. Це правило випливає з релаксації бінарних змінних y_i і y_j:

$$y_i \in \{-1,1\} \quad \Rightarrow \quad -1 \leq y_i \leq 1 \quad \equiv \quad \begin{cases} 1 + y_i \geq 0, \\ 1 - y_i \geq 0, \end{cases}$$

$$y_j \in \{-1,1\} \quad \Rightarrow \quad -1 \leq y_j \leq 1 \quad \equiv \quad \begin{cases} 1 + y_j \geq 0, \\ 1 - y_j \geq 0, \end{cases}$$

та використання для релаксованих змінних очевидних нерівностей:

$$\begin{cases} (1+y_i)(1+y_j) \geq 0, \\ (1-y_i)(1-y_j) \geq 0, \\ (1-y_i)(1+y_j) \geq 0, \\ (1+y_i)(1-y_j) \geq 0, \end{cases} \quad \Rightarrow \quad \begin{cases} +y_i + y_j + y_i y_j \geq -1, \\ -y_i - y_j + y_i y_j \geq -1, \\ -y_i + y_j - y_i y_j \geq -1, \\ +y_i - y_j - y_i y_j \geq -1. \end{cases}$$

Із зростанням r, зокрема при $r=2$ (що відповідає п'ятіркам та четвіркам бінарних змінних), $r=3$ (що відповідає сімкам та шісткам бінарних змінних), лінійні нерівності мають аналогічний вигляд і їх можна вважати правилами п'ятикутника, семикутника і т. д.

Другі форми квадратичних нерівностей (12) та (13) використано при побудові поліедральної апроксимації зверху для бінарного квадратичного многогранника (binary quadratic polytope), який описує множину допустимих розв'язків задачі знаходження екстремуму квадратичної функції від n бінарних змінних. Бінарний квадратичний многогранник

$$BiQP_n = conv\{(y,z) \in \{-1,1\}^{n+\frac{n(n-1)}{2}} : z_{ij} = y_i y_j \ \forall (i,j), 1 \le i < j \le n\}$$

є опуклою оболонкою всіх можливих пар бінарних векторів $y \in \{-1,1\}^n$ та $z \in \{-1,1\}^{\frac{n(n-1)}{2}}$, де на компоненти вектора z накладено наступні обмеження: бінарна змінна z_{ij} дорівнює 1 тільки тоді, коли обидві бінарні змінні y_i та y_j приймають однакові значення, і дорівнює -1 тільки тоді, коли вони приймають різні значення. Многогранник $BiQP_n$ введено за аналогією з відомим булевим квадратичним многогранником (boolean quadratic polytope) М. Падберга

$$BoQP_n = conv\{(x,z) \in \{0,1\}^{n+\frac{n(n-1)}{2}} : z_{ij} = x_i x_j \forall (i,j) : 1 \le i < j \le n\},$$

який є опуклою оболонкою всіх можливих пар булевих векторів $x \in \{0,1\}^n$ та $z \in \{0,1\}^{\frac{n(n-1)}{2}}$, де значення компонент булевого вектора z визначаються іншою логічною функцією, ніж для многогранника $BiQP_n$.

Апроксимаційний многогранник $LBiQP_n$ включає, по-перше, всі можливі лінійні нерівності для послідовностей непарної довжини при всіх $r \in 1,\ldots,r_1$, де r_1 таке, що $2r_1 + 1 \le n$, і, по-друге, усі можливі лінійні нерівності для послідовностей парної довжини $r \in 1,\ldots,r_2$, де r_2 таке, що $2r_2 \le n$ (теорема 4.3). У теоремі вказано максимальну кількість лінійних нерівностей у многограннику $LBiQP_n$: для парних n ця кількість дорівнює $\frac{3(3^n-1)}{4} - n$, а для непарних n дорівнює $\frac{3^{n+1}-5}{4} - n$.

При $n = 2,3,4$ многогранники $LBiQP_n$ повністю збіглися із многраниками, які точно описують $BiQP_n$ за допомогою незвідних систем лінійних нерівностей. Останні були отримані М.Ю. Золотих за допомогою програми Skeleton (http://www.uic.nnov.ru/~zny/skeleton). Многогранник $BiQP_2$ описується 4 лінійними нерівностями для пари бінарних змінних. Система нерівностей для многогранника $BiQP_3$ включає 16 нерівностей: з них 4 для трійки та 3 по 4 для трьох можливих пар. У описі многогранника $BiQP_4$ використовується 56 нерівностей (із них 16 для четвірки, 4 по 4 для чотирьох трійок та 6 по 4 для шести пар змінних).

У підрозділі 4.3 запропоновано дві нові верхні лагранжеві двоїсті оцінки $\psi_b^*(G)$ та $\psi_c^*(G)$ для $\alpha(G)$ – числа стійкості графу G. Число стійкості неорієнтованого графу $G(V,E)$ дорівнює максимальній кількості вершин у незалежній (стійкій) множині його вершин, тобто у такій підмножині з множини вершин $V = \{1,\ldots,n\}$, де немає вершин, зв'язаних ребром із множини E. Нові оцінки пов'язані з квадратичною екстремальною задачею

$$\alpha(G)=\max\sum_{i=1}^{n}x_i \tag{14}$$

при обмеженнях

$$x_i x_j=0,\ \forall(i,j)\in E, \tag{15}$$

$$x_i^2-x_i=0,\ i=1,\ldots,n, \tag{16}$$

$$\begin{cases}2x_i x_j-x_i-x_j+x_{ij}=0,\\ x_{ij}^2-x_{ij}=0,\end{cases}\quad \forall i,j:1\le i<j\le n, \tag{17}$$

$$\begin{cases}2x_i x_{ij}-x_i+x_j-x_{ij}=0,\\ 2x_j x_{ij}+x_i-x_j-x_{ij}=0,\end{cases}\quad \forall i,j:1\le i<j\le n, \tag{18}$$

$$\begin{cases}2x_k x_{ij}-2x_j x_{ik}+x_j-x_k-x_{ij}+x_{ik}=0,\\ 2x_k x_{ij}-2x_i x_{jk}+x_i-x_k-x_{ij}+x_{jk}=0,\\ 2x_{ik}x_{jk}+x_{ij}-x_{ik}-x_{jk}=0,\\ 2x_{ij}x_{jk}-x_{ij}+x_{ik}-x_{jk}=0,\\ 2x_{ij}x_{ik}-x_{ij}-x_{ik}+x_{jk}=0,\end{cases}\quad \forall i,j,k:1\le i<j<k\le n, \tag{19}$$

$$\begin{cases}2x_{ij}x_{kl}-2x_{ik}x_{jl}-x_{ij}+x_{ik}+x_{jl}-x_{kl}=0,\\ 2x_{ij}x_{kl}-2x_{il}x_{jk}-x_{ij}+x_{il}+x_{jk}-x_{kl}=0,\end{cases}\quad \forall i,j,k,l:1\le i<j<k<l\le n. \tag{20}$$

Оцінка $\psi_b^*(G)$ відповідає задачі (14)–(19) з кількістю обмежень $O(n^3)$, а оцінка $\psi_c^*(G)$ – задачі (14)–(20) з кількістю обмежень $O(n^4)$. Остання задача побудована за допомогою функціонально надлишкових обмежень (5)–(8).

Теорема 4.4. *Для довільного графу* G *справедиві такі нерівності:*

$$\Psi_3^*(G)\ge\psi_b^*(G)\ge\psi_c^*(G)\ge\alpha(G).$$

Тут $\Psi_3^*(G)$ – найточніша із запропонованих Н.З. Шором верхніх оцінок, яка використовує $O(n^3)$ квадратичних обмежень та отримана доповненням задачі (14)–(16) функціонально надлишковими обмеженнями іншого вигляду.

Теорема 4.4 стверджує, що оцінка $\psi_b^*(G)$ (а тим більше оцінка $\psi_c^*(G)$*)* є не менш точною верхньою оцінкою для $\alpha(G)$, ніж оцінка Шора $\Psi_3^*(G)$. Це означає, що обидві оцінки є точними, якщо граф G є t-досконалим або h-досконалим, тобто $\psi_b^*(G)=\psi_c^*(G)=\alpha(G)$. Те що, для таких графів маємо $\Psi_3^*(G)=\alpha(G)$, доведено Н.З. Шором. Більш сильний результат про точність побудованих оцінок випливає з властивостей оцінки Шора, які розглянуто в підрозділі 5.1 дисертації.

Обчислювальні експерименти показали, що властивості нових оцінок сильніші, ніж сформульовані в теоремі 4.4. Так, наприклад, перша оцінка $\psi_b^*(G)$ уточнює число Ловаса $\theta'(G)$ як для графу в формі ікосаедра I_{12} (включає 12 вершин та 30 ребер), так і для доповнюючого до нього графу $\overline{I}_{12}$ (включає 12 вершин та 36 ребер). Для обох графів ця оцінка є однаковою $\psi_b^*(\overline{I}_{12})=\psi_b^*(I_{12})\approx 3.185116$.

Виявляється, що оцінка $\psi_c^*(G)$ є точною для обох графів, тобто $\psi_c^*(\overline{I}_{12}) = \alpha(\overline{I}_{12}) = 3$, $\psi_c^*(I_{12}) = \alpha(I_{12}) = 3$. Для порівняння зауважимо, що в роботі "*De Klerk E., Pasechnik D.V.* Approximation of the stability number of a graph via copositive programming // SIAM Journal of Optimization, 2002, Vol. 12, No 4, P. 875–892" для графу $\overline{I}_{12}$ отримано оцінку $(1+\sqrt{5})$, яка дорівнює числу Ловаса $\theta'(\overline{I}_{12})$. Вона знайдена за допомогою техніки "copositive programming", яка вважається однією з найкращих на сьогоднішній день для отримання подібних оцінок.

У підрозділі 4.4 подано три нові квадратичні екстремальні моделі для задачі про максимальний розріз зваженого графу (MaxCut-задача). Їх отримано з використанням функціонально надлишкових обмежень, якими доповнена найпростіша квадратична модель. Для неї верхню оцінку для максимального за вагою розрізу графу, який позначається $mc(G,w)$, побудовано Н.З. Шором та О.А. Березовським у 1995 році. Перша та друга запропоновані моделі використовують квадратичні рівності (6), а третя модель – квадратичні нерівності (10) при $r=1$. Для нових моделей відповідні їм верхні оцінки для $mc(G,w)$ є не гіршими, ніж найкраща з двох оцінок – Барахони – Маджуба (1986 р.) та Шора – Березовського (1995) .

Для перших двох моделей показано, що із них шляхом релаксації бінарних змінних можна отримати задачу лінійного програмування, яка відповідає верхній оцінці Барахони – Маджуба. Для третьої моделі встановлено, що така ж задача лінійного програмування отримується лінеаризацією квадратичних нерівностей (12). Якщо третю модель доповнити квадратичними обмеженнями (12) при $r=2,3,4,\ldots$, отримаємо більш точні верхні оцінки для $mc(G,w)$. Релаксуючи доповнені квадратичні моделі, отримуємо рекомендації для побудови "посилених" аналогів оцінки Барахони – Маджуба в MaxCut-задачах.

У **п'ятому розділі** розглядається ряд оцінок зверху для $\alpha(G,w)$ – зваженого числа стійкості графу. Центральним тут є підрозділ 5.1, у якому описано нові властивості однієї із оцінок, запропонованих Н.З. Шором. Це оцінка $\psi_1(G,w)$, яка відповідає квадратичній задачі:

$$\alpha(G,w) = \max \sum_{i\in V(G)} w_i x_i \tag{21}$$

при обмеженнях

$$x_i x_j = 0 \quad \forall (i,j) \in E(G), \tag{22}$$

$$x_i^2 - x_i = 0 \quad \forall i \in V(G), \tag{23}$$

$$x_i x_k + x_j x_k \le x_k \ \forall (i,j) \in E(G), k \ne i,j, \tag{24}$$

де булева змінна $x_i \in \{0,1\}$ дорівнює одиниці, якщо вершина i включається у стійку (незалежну) множину, та дорівнює нулю у протилежному випадку. Обмеження (22) означають, що дві вершини не можуть одночасно належати стійкій множині, якщо вони зв'язані ребром у графі G. Квадратичні обмеження (24) є функціонально надлишковими, і саме завдяки їм оцінка $\psi_1(G,w)$ співпадає з $\alpha(G,w)$ для W-досконалих графів. Більше того, вона є точною для більш складних графів, ніж W-досконалі. Пов'язано це з тим, що із обмежень (22)–(24) випливають лінійні нерівності (p-wheel constraints):

$$\sum_{i \in V(C_{2k+1})} x_i + k \sum_{j \in V(Q_p)} x_j \le k \quad \forall W_{2k+1+p} \in G, \tag{25}$$

які справедливі для многогранника стійких множин $STAB(G)$. Тут підграф W_{2k+1+p} є p-колесом (p-wheel). Його вершинами слугують вершини непарного циклу C_{2k+1} та кліки Q_p (повного підграфу, що містить p вершин), які не перетинаються. Множина ребер для W_{2k+1+p} включає всі ребра непарного циклу C_{2k+1}, всі ребра кліки Q_p, а також ребра, що зв'язують кожну вершину C_{2k+1} із усіма вершинами кліки Q_p. Приклад 2-колеса показано на рис. 2.

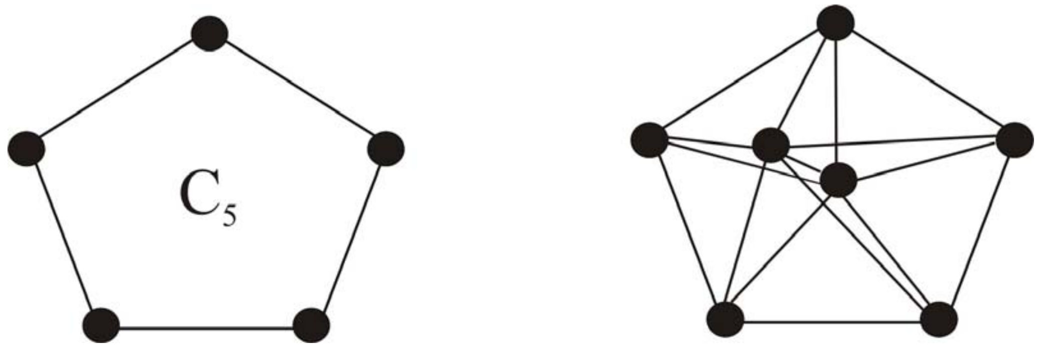

Рис. 2. Приклад 2-колеса на основі непарного циклу C_5

Нерівність (23) означає, що для кожного p-колеса із графу G в стійку множину може бути включена або одна із вершин кліки Q_p, або k вершин із непарного циклу C_{2k+1}. В частковому випадку, коли кліка Q_p містить всього одну вершину, з лінійних нерівностей для довільного p-колеса у формі (25) отримуються лінійні нерівності для звичайних коліс W_{2k+2}.

На основі p-колеса побудовано сімейство W_p-досконалих графів, для яких оцінка Шора є точною. Для них зважене число стійкості графу може бути знайдено за поліноміальний час. Окремими випадками сімейства W_p-досконалих графів є такі відомі сімейства, як досконалі, t-досконалі, h-досконалі та W-досконалі графи. Обговорюються також інші сімейства графів, для яких зважене число стійкості можна знайти за поліноміальний час.

У підрозділі 5.2 описано ітераційний алгоритм знаходження верхньої оцінки для зваженого числа стійкості, в якому многогранник стійких множин $STAB(G)$ апроксимується задачею лінійного програмування з поетапним доповненням її скінченним числом обмежень, що відповідають порушеним лінійним нерівностям для непарних циклів та p-коліс. Порушені непарні цикли отримано на основі алгоритму найкоротших шляхів у спеціальному графі. Алгоритм програмно реалізовано та експериментально досліджено для графів, що містять від кількох сотень до кількох тисяч вершин. Наведено результати тестових експериментів.

У підрозділах 5.3 та 5.4 для зваженого числа стійкості графу побудовано та обгрунтовано дві ЛП-оцінки $\alpha_{\Delta}^{*}(G,w)$ і $\alpha_{\nabla}^{*}(G,w)$, які дорівнюють оптимальному значенню цільової функції в задачах лінійного програмування з кількістю обмежень $O(n^3)$, де n – кількість вершин у графі. Оцінкам відповідають такі ЛП-задачі:

знайти

$$\alpha_{\Delta}^{*}(G,w) = \max \sum_{i\in V} w_i x_i \quad (26)$$

при обмеженнях

$$x_i + x_j - x_{ij} = 0 \ \ \forall (i,j)\in E, \quad (27)$$

$$\begin{cases} +x_{ij} + x_{ik} + x_{jk} \le 2, \\ -x_{ij} - x_{ik} + x_{jk} \le 0, \\ -x_{ij} + x_{ik} - x_{jk} \le 0, \\ +x_{ij} - x_{ik} - x_{jk} \le 0, \end{cases} \quad (28)$$

$$\forall i,j,k \in V : i<j<k,$$

$$0 \le x_i \le 1 \quad \forall i \in V. \quad (29)$$

знайти

$$\alpha_{\nabla}^{*}(G,w) = \max \sum_{i\in V} w_i x_i \quad (26')$$

при обмеженнях

$$x_{ij} = 0 \quad \forall (i,j)\in E, \quad (27')$$

$$\begin{cases} -x_{ij} - x_{ik} - x_{jk} + x_i + x_j + x_k \le 1, \\ +x_{ij} + x_{ik} - x_{jk} - x_i \le 0, \\ +x_{ij} - x_{ik} + x_{jk} - x_j \le 0, \\ -x_{ij} + x_{ik} + x_{jk} - x_k \le 0, \end{cases} \quad (28')$$

$$\forall i,j,k \in V : i<j<k,$$

$$0 \le x_i \le 1 \quad i \in V. \quad (29')$$

Доведено, що для зваженого числа стійкості довільного графу обидві ЛП-оцінки є не менш точними оцінками зверху, ніж відома верхня оцінка, пов'язана з многогранником непарних циклів (теореми 5.3 та 5.5). Наслідком цього факту є те, що оцінки $\alpha_{\Delta}^{*}(G,w)$ та $\alpha_{\nabla}^{*}(G,w)$ є точними для зваженого числа стійкості t-досконалих графів.

Шостий розділ містить опис трьох спеціальних квадратичних екстремальних задач та методів їх розв'язання за допомогою лагранжевих двоїстих оцінок. Особливістю цих задач є те, що цільова функція і обмеження є однорідними квадратичними функціями, тобто такими, які не включають лінійних членів. У цьому випадку алгоритми відшукання двоїстих оцінок зводяться до негладких задач безумовної оптимізації.

У підрозділі 6.1 розглянуто задачу знаходження максимального зваженого розрізу неорієнтованого графу з заданими кількостями вершин в обох підмножинах розбиття вершин графу. Для цієї задачі сформульовано відповідну квадратичну модель, розроблено та досліджено алгоритм отримання лагранжевої двоїстої оцінки, в зв'язку з чим доведено теорему 6.2, в якій за допомогою точних штрафних функцій показано, що знаходження оцінки зводиться до розв'язання безумовної задачі мінімізації опуклої функції. Описано програмну реалізацію розробленого алгоритму та наведено чисельні результати розрахунків верхніх та нижніх оцінок для зваженого розрізу графу Петерсена (10 вершин і 15 ребер) при трьох різних способах розбиття множини його вершин.

Підрозділ 6.2 присвячено задачі глобальної оптимізації, пов'язаній з мінімізацією спеціального вигляду квадратичної однорідної функції на многовиді Штіфеля (наборі k ортонормованих векторів із E^n, де $k \leq n$). Тут описано квадратичні оптимізаційні моделі для трьох варіантів цієї задачі, які відрізняються обмеженнями на компоненти ортонормованих векторів: компоненти можуть належати множинам дійсних, цілих або натуральних чисел. Для кожної задачі побудовано лагранжеві двоїсті оцінки. Для перших двох задач розроблено алгоритми їх знаходження за допомогою зведення до безумовних задач максимізації увігнутих негладких функцій. Для третьої задачі з булевими компонентами векторів доведено, що відповідна їй оцінка є точною (теорема 6.5). На тестових прикладах показано, що розроблені алгоритми можна ефективно використовувати при знаходженні глобального мінімуму на многовиді Штіфеля.

У підрозділі 6.3 розглянуто задачу нелінійного програмування для вивчення зв'язку між прямою та двоїстою моделями Леонтьєва. У задачі максимізується білінійна цільова функція при лінійних обмеженнях та двох квадратичних рівностях, за допомогою яких нормуються вектори кінцевого випуску та доданої вартості. Якщо матриця Леонтьєва (матриця коефіцієнтів прямих витрат) продуктивна, то оптимальний розв'язок задачі виражається через власні вектори, що відповідають максимальним власним числам деяких невід'ємних симетричних матриць. Якщо матриця Леонтьєва продуктивна та нерозкладна (кожна галузь, хоча б неявно, використовує продукцію всіх інших галузей), то задача має єдиний розв'язок, який інтерпретується як оптимальні нормовані структури кінцевого випуску та доданої вартості в продуктивній моделі Леонтьєва. Наведено розрахунки для 15-галузевих матриць Леонтьєва, які побудовано на основі таблиць 38-галузевого балансу України за 2003–2009 роки.

Наведені у підрозділі результати доповнюють арсенал засобів для аналізу якісних властивостей моделей Леонтьєва, який можна здійснити за допомогою чисел і векторів Фробеніуса. Вони дають можливість використовувати для аналізу економічної системи сингулярні числа та власні вектори симетричних матриць. За їх допомогою можна досліджувати зв'язок між витратами на виробництво продукції та цінами при її розподілі в економічній системі. Це можна зробити як для системи в цілому, так і для її окремих галузей, що дає можливість виявити такі коефіцієнти в матриці Леонтьєва, зміни яких необхідно відслідковувати в першу чергу.

Сьомий розділ містить математичні моделі та методи розв'язання ряду прикладних екстремальних задач, де істотну роль відіграють саме субградієнтні методи мінімізації негладких функцій.

У підрозділі 7.1 розглянуто математичні моделі, методи та програмне забезпечення для розв'язання двох задач пошуку мінімальних за сумарною вартістю пропускних здатностей дуг надійної (відказостійкої) орієнтованої мережі. У першій задачі вважається, що обсяги потоків передаються по довільних шляхах, які є допустимими у мережі; у другій задачі обсяги потоків передаються по заданій множині допустимих шляхів. В обох випадках необхідно забезпечити повноцінне функціонування мережі по передачі заданих обсягів потоків при одиничних, але довільних (із заданого списку) пошкодженнях мережі. Одиничним пошкодженням мережі називається такий її стан, при якому зменшується у задане число раз пропускна здатність однієї або декількох дуг мережі.

Математичні моделі представлено задачами лінійного програмування великої розмірності з блочною структурою матриці обмежень. Алгоритми розв’язання обох задач базуються на подвійному використанні двох схем декомпозиції (одна в одній). Зовнішньою є схема декомпозиції за змінними, якими є невідомі значення пропускних здатностей дуг орієнтованої мережі. Схема декомпозиції за обмеженнями використовується для розв’язання підзадач, які виникають при кожному пошкодженні за умови фіксованих значень пропускних здатностей дуг.

Програмне забезпечення реалізовано програмою ModelA для першої оптимізаційної задачі та програмою ModelC для другої задачі (мова програмування ФОРТРАН). У результаті роботи програм ModelA та ModelC для орієнтованої мережі отримуємо мінімальні за сумарною вартістю пропускні здатності дуг, що доповнюють вже існуючі, або достатню умову неможливості виконання всіх вимог по передачі обсягів потоків, щоб орієнтована мережа була надійною. Друге може бути за двох причин: а) структура мережі така, що задовольнити вимоги до потоків неможливо; б) структура пошкоджень у мережі така, що задовольнити вимоги до потоків неможливо.

У підрозділі 7.2 розглянуто дві задачі знаходження оптимальних параметрів плоских багатошарових оптичних покриттів, які сформульовано як багатоекстремальні задачі нелінійного програмування з цільовими функціями складного вигляду. Параметрами зазначених покриттів є коефіцієнти заломлення та оптичні товщини шарів у багатошаровій структурі. Характеристична матриця цієї структури розраховується за допомогою методу Абелле з використанням характеристичних матриць для всіх однорідних шарів, з яких складається багатошарова структура. Обговорено питання знаходження локальних екстремумів цих задач за допомогою субградієнтних методів та проведено аналіз способів обчислення субградієнта цільової функції у залежності від кількості шарів в оптичному покритті.

У підрозділі 7.3 розглянуто сімейство математичних моделей для визначення електричних навантажень паралельно працюючих енергоблоків у енергосистемі з можливістю керування завантаженням (маневреністю) окремого сімейства енергоблоків. Математичні моделі представлено спеціальними задачами нелінійного програмування з неперервними змінними, в яких потрібно мінімізувати нелінійну сепарабельну функцію при лінійних та опуклих обмеженнях на змінні.

У підрозділі 7.4 розглянуто проблему пошуку вектора мінімальної довжини в опуклій оболонці векторів скінченного набору точок у скінченновимірному евклідовому просторі. Вона зводиться до еквівалентної негладкої екстремальної задачі, для якої встановлюється величина штрафного параметра. Для розв'язання еквівалентної задачі пропонується використовувати алгоритми субградієнтного спуску з перетворенням простору.

ВИСНОВКИ

Дисертаційна робота є розвитком методів негладкої оптимізації на основі двох центральних ідей Н.З. Шора: використання лінійних неортогональних перетворень простору для покращення обумовленості яружних функцій (1969 р.) та застосування двоїстого підходу для отримання та уточнення оцінок цільової функції у неопуклих квадратичних моделях (1985 р.).

У дисертації запропоновано та досліджено нові методи мінімізації негладких опуклих функцій та нові способи удосконалення лагранжевих двоїстих оцінок в квадратичних екстремальних задачах з булевими та бінарними змінними. Це дозволило розробити нові ефективні алгоритми для ряду задач оптимізації та побудувати уточнені оцінки для NP-складних екстремальних задач на графах.

У рамках проведених досліджень отримані такі основні результати.

1. Розроблено наближений та прискорений методи еліпсоїдів – нові модифікації методу еліпсоїдів. Обгрунтована їх збіжність зі швидкістю геометричної прогресії, в якій знаменник залежить тільки від вимірності простору. Досліджено їх зв'язок з методом еліпсоїдів Юдіна – Немировського – Шора.

2. Розроблено та обгрунтовано два нові субградієнтні методи з перетворенням простору для знаходження точки мінімуму опуклої функції при відомому мінімальному значенні функції. Ефективність методів для яружних функцій підтверджена результатами обчислювальних експериментів.

3. На основі розроблених методів побудовано: алгоритм знаходження L_p-розв'язку перевизначеної системи лінійних рівнянь при двосторонніх обмеженнях на компоненти розв'язку; алгоритм знаходження допустимої точки опуклої нерівності або перевірки існування такої точки; три ітераційні методи знаходження розв'язку системи лінійних алгебраїчних рівнянь.

4. Запропоновано три способи побудови функціонально надлишкових обмежень в квадратичних екстремальних задачах з булевими та бінарними змінними. Перші два способи генерують квадратичні рівності, які відповідають введенню нових змінних у формі добутку вже існуючих змінних. Третій спосіб для бінарних змінних генерує квадратичні нерівності, які є наслідком того, що квадрат суми "непарної" кількості ± 1 не менше одиниці. На їх основі уточнено лагранжеві двоїсті оцінки для ряду квадратичних екстремальних задач (див. пункти 5, 6 та 7).

5. Для задачі максимізації квадратичної функції з бінарними або булевими змінними побудовано квадратичні екстремальні задачі, яким відповідають уточнені верхні оцінки. Побудована поліедральна апроксимація зверху для бінарного квадратичного многогранника, яка для двох, трьох та чотирьох змінних є точною.

6. Отримано нові верхні оцінки для числа стійкості (незалежності) графу, які уточнюють найточнішу із запропонованих Н.З. Шором верхніх оцінок. Побудовані нові квадратичні моделі для зваженого максимального розрізу графу, для яких відповідні верхні оцінки є не гіршими, ніж найкраща із двох оцінок – оцінка Барахони – Маджуба (1986 р.) та оцінка Шора – Березовського (1995 р.).

7. Побудовано дві нові верхні оцінки для зваженого числа стійкості графу, які є оптимальними значеннями цільової функції у задачах лінійного програмування із кількістю обмежень, що кубічно залежить від числа вершин у графі. Показано, що обидві оцінки є точними для t-досконалого графу.

8. Знайдено нові властивості найкращої із оцінок Шора для зваженого числа стійкості графу. Вони пов'язані з підграфом, відомим як p-колесо, та дозволили побудувати нове сімейство графів, яке названо W_p-досконалими. Для них найкраща оцінка Шора є точною та зважене число стійкості графа може бути знайдено за поліноміальний час.

9. Розроблено та програмно реалізовано ітераційний алгоритм знаходження верхніх оцінок для зваженого числа стійкості графу, в якому многогранник стійких множин апроксимується задачею лінійного програмування з поетапним доповненням її скінченним числом обмежень, відповідним порушеним лінійним нерівностям для непарних циклів та p-коліс. Алгоритм експериментально досліджено для графів, що містять декілька тисяч вершин.

10. Розроблено квадратичні екстремальні моделі та нові методи отримання оцінок для задачі про максимальний розріз зваженого графу з заданими кількостями вершин у підмножинах та для задачі знаходження глобального мінімуму квадратичної функції на многовиді Штіфеля.

11. Розроблено математичні моделі, методи та програмне забезпечення для таких задач: знаходження оптимальних нормованих структур кінцевого випуску та доданої вартості у продуктивній моделі Леонтьєва, знаходження пропускних здатностей дуг надійної (відказостійкої) орієнтованої мережі з передачею потоків довільними шляхами та з передачею потоків по заданій множині допустимих шляхів, знаходження оптимальних параметрів плоских багатошарових оптичних покриттів та знаходження навантажень енергосистеми з урахуванням маневреності окремого сімейства енергоблоків.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. Задачі оптимального проектування надійних мереж / Н.З. Шор, І.В. Сергієнко, П.І. Стецюк та ін. – К.: Наукова думка, 2005. – 230 с.
2. Математичні методи оптимізації та інтелектуальні комп'ютерні технології моделювання складних процесів і систем з урахуванням просторових форм об'єктів / В.В. Грицик, А.І. Шевченко, П.І. Стецюк та ін. – Донецьк: ІПІ "Наука і освіта", 2011. – 650 с.
3. Стецюк П.І. Математичні моделі та програмне забезпечення в задачах енергетики / П.І. Стецюк, М.Г. Журбенко, О.П. Лиховид. – К.: ПП «Ательє «Поліграфічний комплекс», 2012. – 64 с.
4. Стецюк П.И. Линейные операторы в квазиньютоновских методах / П.И. Стецюк // Теория и приложения методов оптимизации. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 1998. – С. 3–8.
5. Стецюк П.И. К методам эллипсоидов / П.И. Стецюк // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 1999. – С. 27–33.
6. Стецюк П.И. Об одном методе для нахождения допустимой точки выпуклого неравенства / П.И. Стецюк // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2000. – С. 3–10.

7. Стецюк П.И. К методам решения плохообусловленных систем линейных уравнений / П.И. Стецюк // Теория оптимальных решений. – Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 2001. – С. 9–15.
8.Стецюк П.И. К ускорению метода эллипсоидов с помощью использования шарового слоя / П.И. Стецюк, Д.М. Буханцов // Теория оптимальных решений. – 2002. – № 1. – С. 63–70.
9. Стецюк П.И. Приближенный метод эллипсоидов / П.И. Стецюк // Кибернетика и системный анализ. – 2003. – № 3. – С. 141–146.
10. Стецюк П.И. Об одном эллипсоиде для внешней аппроксимации n-мерного полушара / П.И. Стецюк // Компьютерная математика. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2003. – Вып. 2. – С. 144–151.
11. Стецюк П.И. Об одном методе нахождения L_p-решения системы линейных уравнений / П.И. Стецюк, Ю.С. Колесник, О.А. Березовский // Теория оптимальных решений. – 2003. – № 2. – С. 83–90.
12. Шор Н.З. Двойственные оценки для специальной оптимизационной задачи квадратичного типа на многообразии Штиффеля / Н.З. Шор, П.И. Стецюк, О.А. Березовский // Теория оптимальных решений. – 2004. – № 3. – С. 3–10.
13. Стецюк П.И. О вычислении градиента в задаче синтеза оптических покрытий / П.И. Стецюк, А.В. Мица // Теория оптимальных решений. – 2004. – № 3. – С. 127–133.
14. Стецюк П.И. О задачах оптимизации параметров для многослойных оптических покрытий / П.И. Стецюк, А.В. Мица // Кибернетика и системный анализ. – 2005. – № 4. – С. 107–115.
15. Стецюк П.И. О функционально избыточных ограничениях для булевых оптимизационных задач квадратичного типа / П.И. Стецюк // Кибернетика и системный анализ. – 2005. – № 6. – С.168–172.
16. Стецюк П.И. Новые модели квадратичного типа для задачи о максимальном взвешенном разрезе графа / П.И. Стецюк // Кибернетика и системный анализ. – 2006. – № 1. – С. 63–75.
17. Стецюк П.И. Лагранжевы оценки для максимального разреза графа с заданными количествами вершин в обоих подмножествах разбиения / П.И. Стецюк, О.А. Березовский // Теория оптимальных решений. – 2006. – № 5. – С. 31–38.
18. Стецюк П.И. Об уточнении лагранжевых двойственных оценок в бинарных и булевых квадратичных задачах / П.И. Стецюк, П.М. Пардалос // Теория оптимальных решений. – 2006. – № 5. – С. 145–153.
19. Стецюк П.И. О новых лагранжевых двойственных оценках для числа устойчивости графа / П.И. Стецюк, П.М. Пардалос, Д.Л. Крошко // Компьютерная математика. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2006. – Вып. 3. – С. 149–158.
20. Стецюк П.И. Об одной верхней оценке для взвешенного числа устойчивости графа / П.И. Стецюк, С.И. Бутенко, О.А. Березовский // Теория оптимальных решений. – 2007. – № 6. – С. 80–89.
21. Стецюк П.И. ЛП-ориентированная верхняя оценка для числа устойчивости графа на основе p-колес / П.И. Стецюк, С.И. Бутенко, А.П. Лиховид // Теория оптимальных решений. – 2008. – № 7. – С. 34–44.

22. Стецюк П.И. Об ЛП-ориентированных верхних оценках для взвешенного числа устойчивости графа / П.И. Стецюк, А.П. Лиховид // Кибернетика и системный анализ. – 2009. – № 1. – С.157–170.
23. Сергиенко И.В. Модели и информационные технологии для поддержки принятия решений при проведении структурно-технологических преобразований / И.В. Сергиенко, М.В. Михалевич, П.И. Стецюк, Л.Б. Кошлай // Кибернетика и системный анализ. – 2009. – № 2. – С. 26–49.
24. Стецюк П.И. Задачи оптимизации для выбора электрических нагрузок в энергосистеме / П.И. Стецюк, А.П. Лиховид, А.В. Пилиповский // Теория оптимальных решений. – 2009. – № 8. – С.136–141.
25. Стецюк П.И. Точная ЛП-оценка для взвешенного числа устойчивости t-совершенных графов / П.И. Стецюк, В.И. Ляшко, Е.А. Нурминский // Журнал обчислювальної та прикладної математики, – 2009. – № 3(99). – С.106–115.
26. Стецюк П.И. Негладкий штраф и субградиентные алгоритмы для решения задачи проекции на политоп / П.И. Стецюк, Е.А. Нурминский // Кибернетика и системный анализ. – 2010. – № 1. – С. 59–63.
27. Стецюк П.И. Оптимальная нормированная структура спроса и добавленной стоимости в продуктивной модели Леонтьева / П.И. Стецюк, Л.Б. Кошлай // Кибернетика и системный анализ. – 2010. – № 5. – С. 51–59.
28. Стецюк П.И. О задаче оптимального соотношения между спросом и добавленной стоимостью в моделях Леонтьева / П.И. Стецюк, Л.Б. Кошлай, А.В. Пилиповский // Теория оптимальных решений. – 2010. – № 9. – С. 136–143.
29. Стецюк П.И. Бинарный квадратичный многогранник и его аппроксимации / П.И. Стецюк, Н.Ю. Золотых // Журнал обчислювальної та прикладної математики. – 2010. – № 2(101). – С. 76–86.
30. Стецюк П.И. О спектральных свойствах модели Леонтьева / П.И. Стецюк, А.В. Бондаренко // Теория оптимальных решений. – 2011. – № 10. – С. 84–90.
31. Стецюк П.И. Об одной экстремальной задаче для связи прямой и двойственной моделей Леонтьева / П.И. Стецюк, Л.Б. Кошлай // Спектральные и эволюционные задачи. – 2011. – Т. 2. – № 2. – С. 164–169.
32. Сергиенко И.В. О трех научных идеях Н.З. Шора / И.В. Сергиенко, П.И. Стецюк // Кибернетика и системный анализ. – 2012. – № 1. – С. 4–22.
33. Стецюк П.И. Ускорение субградиентного метода Поляка / П.И. Стецюк // Теория оптимальных решений. – 2012. – № 11. – С. 151–160.
34. Shor N.Z. Lagrangian bounds in multiextremal polynomial and discrete optimization problems / N.Z. Shor, P.I. Stetsyuk // J. of Global Optimization. – 2002 – **23**. – P. 1–41.
35. Стецюк П.И. Субградиентные методы переменной метрики, использующие шаг Агмона – Моцкина и одноранговый эллипсоидальный оператор / П.И. Стецюк // Труды АТИК, 2007–2008. – Кишинэу: Эврика, 2009. – Т. I (XII). – С. 16–25.
36. 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. Иваньо и др. – Иркутск: Институт систем энергетики им. Л.А. Мелентьева СО РАН, 2012. – С. 160–184.

37. Стецюк П.И. Об ускорении сходимости методов эллипсоидов / П.И. Стецюк // Труды XII Байкальской международной конференции (Иркутск, Байкал, 24 июня – 1 июля 2001г.). Том 1. Математическое программирование. – Иркутск, 2001. – С. 61–66.
38. Стецюк П.И. Оценка Шора для взвешенного числа устойчивости W_p-совершенных графов / П.И. Стецюк // Материалы международной научной конференции "Математическое моделирование, оптимизация и информационные технологии", 19–21 марта 2008, Кишинэу, Академия транспорта, информатики и коммуникаций. – С. 112–121.
39. Стецюк П.И. О новых свойствах оценок Шора для взвешенного числа устойчивости графа / П.И. Стецюк // 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.
40. 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.
41. Стецюк П.И. 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.
42. Стецюк П.И. 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.
43. Сергиенко И.В., Стецюк П.И. Три центральные идеи Н.З. Шора / И.В. Сергиенко, П.И. Стецюк // Математическое моделирование, оптимизация и информационные технологии: материалы 3-й Междунар. науч. конф. (Кишинэу, 19–23 марта 2012 г.): Кишинэу: Эврика, 2012. – С. 468–475.

АНОТАЦІЯ

Стецюк П.І. Алгоритми недиференційовної оптимізації та лагранжеві двоїсті оцінки в складних екстремальних задачах. – Рукопис.

Дисертація на здобуття наукового ступеня доктора фізико-математичних наук за спеціальністю 01.05.01 – теоретичні основи інформатики та кібернетики. – Інститут кібернетики імені В.М. Глушкова НАН України, Київ, 2013.

Дисертація присвячена розробці та дослідженню субградієнтних методів з перетворенням простору та способів уточнення двоїстих оцінок в квадратичних екстремальних задачах.

У роботі розроблено нові модифікації методу еліпсоїдів та субградієнтні методи з перетворенням простору для знаходження точки мінімуму опуклої функції при відомому її мінімальному значенні. З'ясовано ефективність розроблених

субградієнтних методів у застосуванні до яружних функцій. Запропоновано три нові способи побудови функціонально надлишкових обмежень у квадратичних задачах з булевими та бінарними змінними. Перші два способи генерують квадратичні рівності, які відповідають введенню нових змінних у формі добутку вже існуючих змінних. Третій спосіб для бінарних змінних генерує квадратичні нерівності, які є наслідком того що квадрат суми непарної кількості ± 1 не менше одиниці. На основі цих способів отримано уточнені двоїсті оцінки зваженого числа стійкості графу, зваженого максимального розрізу графу, для задачі максимізації квадратичної функції з бінарними або булевими змінними. Побудована верхня поліедральна апроксимація бінарного квадратичного многогранника, яка при двох, трьох та чотирьох змінних є точною. Побудовано сімейство W_p-досконалих графів, для якого найкраща з оцінок Н.З. Шора є точною для зваженого числа стійкості графа. Створені методи застосовано для знаходження оптимальних параметрів економічних, транспортних та енергетичних систем.

Ключові слова: недиференційовна оптимізація, субградієнтний метод, метод еліпсоїдів, квадратична екстремальна задача, двоїста оцінка, функціонально надлишкові обмеження, число стійкості графа, максимальний розріз графа.

АННОТАЦИЯ

Стецюк П.И. Алгоритмы недифференцированной оптимизации и лагранжевые двойственные оценки в сложных экстремальных задачах. – Рукопись.

Диссертация на соискание ученой степени доктора физико-математических наук по специальности 01.05.01 – теоретические основы информатики и кибернетики. – Институт кибернетики имени В.М. Глушкова НАН Украины, Киев, 2013.

Диссертация посвящена разработке и исследованию новых субградиентных методов с преобразованием пространства и новых способов уточнения лагранжевых двойственных оценок в квадратичных экстремальных задачах с булевыми и бинарными переменными.

В работе предложены приближенный и ускоренный методы эллипсоидов – новые модификации метода эллипсоидов. Обоснована их сходимость со скоростью геометрической прогрессии, в которой знаменатель зависит только от размерности пространства. Исследована их связь с методом эллипсоидов Юдина – Немировского – Шора. На основе приближенного метода эллипсоидов построен алгоритм нахождения L_p-решения переопределенной системы линейных уравнений при двусторонних ограничениях на компоненты решения.

Разработаны и обоснованы два новые субградиентные методы с преобразованием пространства для нахождения точки минимума выпуклой функции при известном минимальном значении. Они используют шаг Поляка (шаг Агмона – Моцкина – Шенберга) в преобразованном пространстве и монотонно уменьшают расстояние до точки минимума в последовательно преобразованных пространствах переменных. Эффективность методов для овражных функций подтверждена результатами вычислительных экспериментов. На основе второго метода

построены алгоритм нахождения допустимой точки выпуклого неравенства или проверки существования допустимой точки в шаре начальной локализации и три итерационных метода нахождения решения системы линейных алгебраических уравнений, эффективность которых подтверждена численными расчетами для плохообусловленных задач.

Предложены три способа построения функционально избыточных ограничений в квадратичных экстремальных задачах с булевыми и бинарными переменными. Первые два способа генерируют квадратичные равенства, которые соответствуют введению новых переменных в форме произведения уже существующих переменных. Третий способ для бинарных переменных генерирует квадратичные неравенства, которые являются следствием того, что квадрат суммы нечетного количества ± 1 не меньше единицы. На их основе уточнены лагранжевые двойственные оценки для ряда задач.

Для задачи максимизации квадратичной функции с бинарными или булевыми переменными построены квадратичные экстремальные задачи, которые индуцируют более точные верхние оценки. Построена полиэдральная аппроксимация сверху для бинарного квадратичного многогранника, которая для двух, трех и четырех переменных является точной. Построены новые верхние оценки для числа устойчивости (независимости) графа, которые уточняют самую точную из предложенных Н.З. Шором верхних оценок. Построены новые квадратичные модели для взвешенного максимального разреза графа, для которых соответствующие верхние оценки являются не хуже, чем лучшая из двух оценок – Барахоны – Маджуба (1986г.) и Шора – Березовского (1995г.). Построены две новые верхние оценки для взвешенного числа устойчивости графа, они являются оптимальными значениями целевой функции в задачах линейного программирования с количеством ограничений, которое кубически зависит от числа вершин в графе. Показано, что обе оценки являются точными для t-совершенного графа.

Найдены новые свойства самой точной из оценок Шора для взвешенного числа устойчивости графа. Они связаны с подграфом, известным как p-колесо, и позволили построить семейство W_p-совершенных графов, для которых оценка Шора является точной. Для этого семейства взвешенное число устойчивости графа может быть найдено за полиномиальное время. Полученные результаты позволили построить итерационный алгоритм, в котором многогранник устойчивых множеств аппроксимируется задачей линейного программирования с поэтапным дополнением ее конечным количеством ограничений, соответствующих нарушенным линейным неравенствам для нечетных циклов и p-колес. Алгоритм программно реализован и экспериментально исследован для графов, содержащих несколько тысяч вершин.

Разработаны квадратичные экстремальные модели и методы получения оценок для задачи о максимальном разрезе взвешенного графа с заданными количествами вершин в подмножествах разбиения и для задачи нахождения глобального минимума квадратичной однородной функции на многообразии Штифеля (наборе k ортонормированных векторов из евклидового пространства E^n, $k \le n$). Разработаны математические модели, методы и программное обеспечение для решения таких задач: нахождения оптимальных нормированных структур конечного выпуска и добавленной стоимости в продуктивной модели Леонтьева; нахождения пропускных способностей дуг надежной (отказоустойчивой) ориентированной сети с передачей потоков произвольными путями и с передачей потоков по заданному множеству допустимых путей; нахождения оптимальных параметров плоских многослойных оптических покрытий; нахождения нагрузок энергосистемы с учетом маневренности отдельного семейства энергоблоков.

Ключевые слова: недифференцируемая оптимизация, субградиентный метод, метод эллипсоидов, квадратичная экстремальная задача, двойственная оценка, функционально избыточные ограничения, число устойчивости графа, максимальный разрез графа.

ABSTRACT

Stetsyuk P.I. Nondifferentiable optimization algorithms and Lagrangian dual bounds in complex extremal problems. – Manuscript.

Thesis submitted for a Doctor’s degree in physical and mathematical sciences by speciality 01.05.01 – theoretical bases of informatics and cybernetics. – V.M. Glushkov Institute of Cybernetics of National Academy of Science of Ukraine, Kyiv, 2013.

The thesis is devoted to development and investigation of new subgradient methods with space transformation and new techniques to improve Lagrangian dual bounds in extremal quadratic problems.

New modifications of ellipsoid method are developed. Their convergence at a geometric rate, depending only on space dimension, was proved. New subgradient methods with space transformation to find minimum point of convex functions with the known value were developed. The efficiency of the methods for ravine functions is given. Three new methods for constructing superfluous constraints in quadratic extremal problems with boolean and binary variables were proposed. The first two methods generate quadratic equalities, which correspond to introduction of new variables in the form of the product of the existing variables. The third method for binary variables generates quadratic inequalities, which are a consequence of the fact that a square of the sum of odd number of ± 1 is not less than one. Using the developed methods the improved Lagrangian dual bounds for the stability number and weighted stability number of the graph, for the maximum weighted cut of the graph, for the problem of maximizing a quadratic function of binary or boolean variables were received. The upper polyhedral approximation of binary quadratic polytope, which for two, three and four variables is exact, was constructed. The developed methods and algorithms were applied to find the optimal parameters for economic, transport, energy systems.

Key words: nondifferentiable optimization, subgradient method, ellipsoid method, quadratic optimization problem, dual bound, superfluous constraints, stability number of graph, maximum cut problem.

Підп. до друку 12.02.2013. Формат 60х84/16. Папір офс.
Цифровий друк. Ум. друк. арк. 1,86. Ум. фарбо-відб. 2,09.
Обл.-вид. арк. 2,0. Зам. 14. Тираж 100 прим.

Редакційно-видавничий відділ з поліграфічною дільницею
Інституту кібернетики імені В.М. Глушкова НАН України
03680, МСП, Київ-187, проспект Академіка Глушкова, 40