3.5. Алгоритми недиференційовної оптимізації та їх застосування

Підрозділ містить огляд основних результатів школи недиференційовної (негладкої) оптимізації Інституту кібернетики ім. В.М.Глушкова НАН України. Засновником цього наукового напрямку та його лідером заслужено вважається академік Наум Зуселевич Шор (1937–2006).

До мінімізації опуклих функцій з розривним градієнтом зводиться велике число проблем, що виникають при розв'язуванні складних задач математичного програмування. Володіння методами недиференційовної оптимізації дає можливість гнучко використовувати схеми декомпозиції (за змінними, обмеженнями, ресурсами тощо), які враховують специфіку задач великої розмірності, дозволяє ефективно одержувати двоїсті оцінки в задачах дискретного й неперервно-дискретного програмування та для деяких класів багатоекстремальних задач. З'являється можливість використовувати негладкі функції штрафу, що дозволяють при скінченних значеннях штрафних параметрів одержувати задачу безумовної мінімізації, еквівалентну початковій задачі опуклого програмування. Техніко-економічні характеристики об'єктів, що підлягають оптимізації, зазвичай добре апроксимуються кусочно-гладкими функціями від невідомих параметрів, і це також породжує задачі оптимізації з негладкими функціями. Відсутність ефективних методів негладкої оптимізації ускладнювало розв'язування вказаних класів задач та змушувало або змінювати формулювання задачі, що погіршувало відповідність моделі реальності, або використовувати різні прийоми згладжування. Останнє не завжди приводить до успіху, тому що застосування згладження погіршує обумовленість функції, яку мінімізують, та знижує обчислювальну стійкість навіть таких ефективних методів гладкої мінімізації, як квазиньютонівські та методи спряжених градієнтів.

Таким чином, область застосувань методів негладкої оптимізації є досить широкою, й розробці обчислювальних методів негладкої оптимізації слід надавати велику увагу. Як за надійністю, так і за часом розрахунків й точності результатів, методи негладкої оптимізації виявилися конкурентноздатними з найбільш ефективними методами рішення гладких, але погано обумовлених задач. Нижче буде наведено короткий огляд розроблених в Інституті кібернетики під керівництвом академіка Н.З. Шора сімейств алгоритмів недиференційовної оптимізації та їх численних застосувань. Дамо коротку характеристику описаних методів недиференційовної оптимізації.

1. Методи узагальненого градієнтного спуску (УГС) (розділ 3.5.1). Методи УГС послужили основою нового напрямку математичного програмування – чисельним методам негладкої оптимізації, історія розробки яких описана в численних наукових статтях та монографіях. Найбільш інтенсивно дослідження з методів УГС велися в 1962–1971 р.

2. Субградієнтні методи з розтягом простору в напрямку субградієнта (розділ 3.5.2), що у порівнянні з методами УГС мають прискорену збіжність. Ці методи дали теорії оптимізації унікальний алгоритм – метод еліпсоїдів, швидкість збіжності якого залежить лише від розмірності простору. Використання методу еліпсоїдів дозволило вирішити ряд важливих питань у теорії складності задач математичного програмування. Найбільш результативними ці дослідження виявилися в період з 1968 по 1986 р.

3. Субградієнтні методи з розтягом простору в напрямку різниці двох послідовних субградієнтів – r-алгоритми (розділ 3.5.3). На поточний момент r-алгоритми є могутнім практичним засобом розв'язування задач недиференційовної оптимізації. При мінімізації гладких функцій вони виявилися конкурентноздатними з найбільш вдалими реалізаціями методів спряжених напрямків та методів квазиньютоновського типу. Початок досліджень з r-алгоритмів припав на 1971–1974 рр. Потім вони активно продовжувалися аж донині.

4. Субградієнтні методи використовувалися в задачах оптимізації великої розмірності та в блочних задачах із різними схемами декомпозиції, для розв'язування мінімаксних та матричних задач оптимізації, для обчислення лагранжевих (двоїстих) оцінок у багатоекстремальних задачах оптимізації. Вони застосовувалися для розв'язування задач оптимального планування й проектування, синтезу мереж, відновлення зображень, еліпсоїдальній апроксимації і локалізації й ін. Основні джерела, що породжують задачі негладкої оптимізації, та питання застосування до них методів недиференційовної оптимізації описані в розділі 3.5.4.

5. Особливим джерелом задач недиференційовної оптимізації є підхід, пов'язаний із знаходженням лагранжевих (двоїстих) оцінок у неопуклих квадратичних моделях. Такі моделі зустрічаються в багатьох практичних задачах, і лагранжеві оцінки тут відіграють важливу роль, тому що дають можливість виділити такі підкласи з NP-складних задач, які можна розв'язати за поліноміальний час. Лагранжеві оцінки можна поліпшувати, додаючи в модель так звані функціонально надлишкові квадратичні обмеження. У розділі 3.5.5. описано техніку лагранжевих оцінок для багатоекстремальних квадратичних задач та подано приклади її використання для задач поліноміального типу й екстремальних задач на графах.

3.5.1 Узагальнений градієнтний спуск

Методи узагальненого градієнтного спуску (УГС) поклали початок новому напрямкові математичного програмування – чисельним методам негладкої оптимізації. Значну частину творчої спадщини Н.З. Шора займає теоретичне дослідження цих методів. Про суть проблем, пов'язаних з недиференційовністю та про внесок Н.З. Шора в їхнє рішення чудово написав Б.Т. Поляк [1]: "Основные алгоритмы минимизации гладких функций – градиентный и Ньютона – были построены на использовании линейной и квадратичной аппроксимации функции, задаваемой первыми членами ряда Тейлора. Однако, для недиференцируемой функции эта идея неприменима – такая функция не может быть хорошо аппроксимирована ни линейной, ни квадратичной функциями. ... Поэтому разработка методов минимизации негладких функций требует привлечения новых идей. Одна из них, принадлежащая Н.З. Шору, выглядит несколько неожиданно. Пишется прямой аналог градиентного метода с заменой градиента на произвольный субградиент $g_f(x)$ функции $f(x)$:

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

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

Наведемо ряд результатів по методу узагальненого градієнтного спуска з книги Н.З. Шора [2]. Ця книга практично відразу ж була переведена на англійську мову видавництвом "Шпрінгер" (1985).

Нехай $f(x)$ – опукла функція, визначена на евклідовому просторі E^n, X^* – множина мінімумів (вона може бути і пустою), $x^* \in X^*$ – точка мінімуму; $\inf f(x) = f^*$; $g_f(x)$ – субградієнт (довільний) функції $f(x)$ в точці x. Субградієнт $g_f(\bar{x})$ функції f в точці $\bar{x}$ є вектор $g_f(\bar{x})$ такий, що

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

З визначення субградієнта випливає, що при $f(x) < f(\bar{x})$, виконується

$$(-g_f(\bar{x}), x - \bar{x}) > 0. \tag{2}$$

Геометрично формула (2) означає, що антисубградієнт у точці $\bar{x}$ утворює гострий кут з будь-якою прямою, проведеною з точки $\bar{x}$ в напрямку точки x з меншим значенням $f(x)$. Звідси, якщо множина X^* є непустою та $\bar{x} \notin X^*$, то при зміщенні з точки $\bar{x}$ у напрямку $-g_f(\bar{x})$ з досить малим кроком відстань до X^* буде зменшуватись. Цей простий факт лежить в основі субградієнтного методу або методу узагальненого градієнтного спуска (УГС), уперше запропонованого в [3] для розв’язання транспортної задачі на мережі. Саме на цьому простому факті робиться наголос у наведеній вище цитаті з книги Б.Т. Поляка.

Методом узагальненого градієнтного спуску (УГС) називається процедура побудови послідовності $\{x_k\}_{k=0}^{\infty}$, де x_0 –- початкове наближення, а x_k обчислюються за такою рекурентною формулою: :

$$x_{k+1} = x_k - h_{k+1} \frac{g_f(x_k)}{\| g_f(x_k) \|}, \quad k = 0,1,2,\ldots; \tag{3}$$

тут $g_f(x_k)$ – довільний субградієнт функції $f(x)$ в точці x_k, h_{k+1} – кроковий множник. Якщо $g_f(x_k) = 0$, то x_k є точкою мінімуму функції $f(x)$, і процес зупиняється.

Починаючи з 1962 року Н.З. Шор розробив кілька варіантів УГС, в яких використовуються співвідношення (3). Частина з них послужила основою його кандидатської дисертації “Про структуру алгоритмів чисельного рішення задач оптимального планування й проектування” (1964р.). Найбільш повно результати з цієї тематики, отримані в період з 1962 по 1971 рік, відображені в монографії [2]. До неї, зокрема, було включено також ряд матеріалів докторської дисертації Н.З. Шора “Методы минимизации недиференцируемых функций и их приложения” (1970 р.). Найбільш загальний результат про збіжність УГС міститься в наступній теоремі [2].

Теорема 1 *Нехай $f(x)$ – опукла функція, визначена на E^n з обмеженою областю мінімумів X^*, $\{h_k\}$ $(k = 1,2,\ldots)$ – послідовність чисел, що має такі властивості:*

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

Тоді для послідовності $\{x_k\}$ $(k=1,2,\ldots)$, *побудованої згідно з формулою* (3), *для довільного* $x_0 \in E^n$ *існує дві можливості: або знайдеться таке* $k=\bar{k}$, *що* $x_{\bar{k}} \in X^*$, *або* $\lim\limits_{k\to\infty} \min\limits_{y\in X^*} \| x_k - y \| = 0$, $\lim\limits_{k\to\infty} f(x_k) = \min\limits_{y\in E^n} f(x) = f^*$.

При певних додаткових припущеннях Н.З. Шором і П.Р. Гамбурдом [4] були отримані варіанти УГС, що сходяться зі швидкістю геометричної прогресії.

Теорема 2 *Нехай* $f(x)$ *– опукла функція, визначена на* E^n, *і для всіх* $x \in E^n$ *при якомусь* $\varphi\,(0 \le \varphi < \pi/2)$ *виконується нерівність*

$$(g_f(x), x - x^*(x)) \ge \cos\varphi \| g_f(x) \| \cdot \| x - x^*(x) \|, \tag{4}$$

де $x^*(x)$ *– точка, що належить множині мінімумів функції* $f(x)$ *і лежить на найменшій відстані від* x. *Тоді, якщо при заданому* x_0 *вибрати величину* h_1, *що задовольняє нерівності*

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

визначити $\{h_k\}_{k=1}^{\infty}$ *згідно з рекурентною формулою*

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

де

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

та обчислити $\{x_k\}_{k=1}^{\infty}$ *по формулі* (3), *то або при деякому* k^* $g_f(x_{k^*})$ *і* x_{k^*} *належать області мінімумів, або при усіх* $k=1,2,\ldots$ *виконується нерівність*

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

Таким чином, якщо кут φ заздалегідь відомий, то, регулюючи крок згідно з формулами теореми 2, можна одержати збіжність до мінімуму зі швидкістю геометричної прогресії зі знаменником $q = r(\varphi)$.

У формулі (4) величина $\cos\varphi$ характеризує ступінь витягнутості поверхонь рівня функції $f(x)$. Якщо в деякому околі мінімуму функції $f(x)$ не існує такого кута $\varphi < \pi/2$, що для кожного x із цього околу виконується (4), то таку функцію називають *суттєво яружною*. При мінімізації суттєво яружних функцій наведений у теоремі 2 спосіб регулювання крокових множників застосувати неможливо. У цьому випадку варто використовувати універсальний спосіб вибору крокових множників, вказаний у теоремі 1.

Наступна теорема, аналогічна теоремі 2, сформульована безпосередньо в термінах, що характеризують ступінь "витягнутості" поверхонь рівня.

Теорема 3 *Нехай опукла функція* $f(x)$ *визначена на* E^n, x^* *–- єдина точка мінімуму* $f(x)$, *задані початкове наближення* x_0 *та числа* σ *і* h_1, *причому* $\sigma \ge \sqrt{2}$, $h_1 \ge \| x_0 - x^* \| \;/\sigma$. *Розглянемо множину*

$$Y = \left\{ y : \| y - x^* \| \le \sigma h_1 \right\}$$

Якщо для будь-якої пари точок x, $z \in Y$ *таких, що* $f(x) = f(z) \ne f(x^*)$, *виконується умова*

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

то послідовність $\{x_k\}_{k=0}^{\infty}$, *утворена за допомогою рекурентних формул* (3), *де* $h_{k+1} = h_k \sqrt{\sigma^2 - 1}/\sigma$, *збігається до* x^* *зі швидкістю геометричної прогресії:*

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

за винятком випадку, коли для деякого $k = \bar{k}$ $g_f(x_{\bar{k}}) = 0$, *тобто* $x_{\bar{k}} = x^*$.

У наступному варіанті методу узагальненого градієнтного спуска кроковий множник залишається протягом визначеного числа кроків постійним, а потім зменшується в два рази [2].

Теорема 4 *Нехай для опуклої функції* $f(x)$ *виконуються умови теореми* 1, $\sigma \ge 2$. *Розглянемо при заданому х ітеративний процес* (3), *де* $h_{k+1} = h_0 \cdot 2^{-\lfloor (k+1)/N \rfloor}$. *Тут* $\lfloor a \rfloor$ *– ціла частина числа* a. *При достатньо великому* h_0 *та* $N \ge 3\sigma^2 + 1$ *виконується нерівність*

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

Регулювання кроку відповідно з теоремою 4 було обґрунтовано в роботі Н.З. Шора і М.Б. Щепакіна (1967), де був розглянутий алгоритм розв’язання двохетапної задачі стохастичного програмування, оснований на застосуванні методу узагальненого градієнтного спуска. Однак, уперше це регулювання кроку було використано в методі УГС, що був запропонований у 1962 році для розв’язання транспортної задачі в мережевій формі [3]. Робота [3] була першим прикладом використання субградієнтного процесу для мінімізації опуклих недиференційовних функцій.

Методи УГС дали можливість розв’язати велику кількість задач виробничо-транспортного планування із застосуванням схем декомпозиції для задач великої розмірності. Докладну інформацію про ці задачі можна знайти в [2, 5]. Метод УГС також послужив основою для створення стохастичного аналога узагальненого градієнтного спуска [6], що має велике практичне значення, зокрема, при розв’язанні багатоетапних задач стохастичного програмування. В [5] описане застосування методу узагальненого стохастичного градієнта для розв’язання двохетапної стохастичної транспортної задачі, пов'язаної з визначенням обсягів складів однорідної продукції при випадковому попиті.

Методи УГС, отримані в Інституті кібернетики, одержали розвиток у роботах І.І. Єрьоміна [7] і Б.Т. Поляка [8] для розв’язання задач опуклого програмування з обмеженнями. До 1974 року роботи з методів УГС були мало відомі за кордоном, тому що були опубліковані російською мовою в малодоступних виданнях. Вони одержали популярність після публікації в [9] докладного огляду результатів і бібліографії робіт з недиференційовної оптимізації, виконаних у СРСР.

3.5.2 Субградієнтні методи з розтягом простору в напрямку субградієнта

При аналізі алгоритмів УГС, що сходяться зі швидкістю геометричної прогресії, суттєву роль грали верхні границі кутів між напрямком антиградієнта в даній точці й напрямком прямої, проведеної з цієї точки в точку мінімуму. Якщо верхня границя зазначених кутів дорівнює $\pi/2$, повільну збіжність УГС в рамках цього методу в яружних задачах покращити неможливо. Ситуацію можна змінити, використовуючи лінійні неортогональні перетворення простору змінних для покращення обумовленості задачі. У випадку коли антиградієнти утворять з напрямком на точку мінімуму кут, близький до $\pi/2$, доцільно застосувати операцію розтягу простору в напрямку градієнта для зменшення його “поперечної” складової. Ці евристичні міркування послужили основою створення сімейства методів субградієнтного типу з розтягом простору.

Операція розтягу простору в напрямку градієнта вперше була введена у роботі Н.З. Шора й В.І. Білецького [10] як евристична процедура для покращення обумовленості задачі. Потім у роботі [11] Н.З.Шор обґрунтував цей підхід та одержав ряд теоретичних результатів для субградієнтного процесу з розтягом простору в напрямку субградієнта (див. теорему 5). Операція розтягу простору змінних реалізується за допомогою оператора розтягу простору, що у векторній формі можна записати так:

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

де $(\cdot)^T$ означає транспонування, I_n – одинична матриця розміру $n \times n$, α – коефіцієнт розтягу простору, ξ – напрямок розтягу. Докладно властивості оператора $R_\alpha(\xi)$ викладені в монографії [2]. При описі алгоритмів використовується оператор $R_\beta(\xi)$, обернений до оператора розтягу простору $R_\alpha(\xi)$. Він має такий вигляд:

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

і в методах з розтягом простору змінних використовується для "стискування" простору субградієнтів.

Опишемо загальну схему алгоритмів субградієнтного типу з розтягом простору в напрямку субградієнта для мінімізації функції $f(x)$.

Задано $x_0 \in E^n$, $B_0 = A_0^{-1} = I_n$ (одиничну матрицю розміру $n \times n$). Після k кроків маємо $x_k \in E^n$, $B_k = A_k^{-1}$, A_k – матриця розміру $n \times n$ перетворення простору після k кроків. На $k+1$-м кроці виконуються такі операції.

1. Обчислюємо $g_f(x_k)$ (якщо $g_f(x_k) = 0$, процес зупиняється).

2. Визначаємо $\tilde{g}_k = g_{\varphi_k}(y_k) = B_k^* g_f(x_k)$, де $\varphi_k(y) = f(B_k y)$; $y_k = A_k x_k$; $\tilde{g}_k$ – узагальнений градієнт для функції $\varphi_k(y)$, визначеної в "розтягнутому" просторі.

3. Знаходимо

$$\xi_k = \tilde{g}_k / \|\tilde{g}_k\|; \quad x_{k+1} = x_k - h_{k+1} B_k \xi_k. \tag{5}$$

Формулі (5) відповідає зміщення по антисубградієнту в "розтягнутому" просторі: $A_k x_{k+1} = y_k - h_{k+1}\xi_k$.

4. Обчислюємо

$$B_{k+1} = A_{k+1}^{-1} = B_k R_{\beta_{k+1}}(\xi_k), \quad \beta_{k+1} = 1/\alpha_{k+1}. \tag{6}$$

Формулі (6) відповідає розтяг перетвореного простору в напрямку ξ_k: $A_{k+1} = R_{\alpha_{k+1}}(\xi_k)A_k$, $\alpha_{k+1} > 1$.

5. Переходимо до наступного кроку: $k+1 \to k+2$.

Основна складність при конструюванні працездатного алгоритму полягала у виборі коефіцієнтів розтягу простору α_k та стратегії зміни крокових множників h_k. Експерименти показали, що, вибираючи $\alpha = 2$ і $h_k = const$, для багатьох прикладів опуклих яружних функцій можна одержати чудові результати [10]. Авторами програм були Н.З. Шор і В.І. Білецький. На жаль, такий простий спосіб не завжди приводить до мети. При побудові інших варіантів алгоритмів, що удалося теоретично обґрунтувати, кроковий множник і коефіцієнти розтягу простору вибиралися таким чином, щоб послідовність відстаней до точки мінімуму у відповідних перетворених просторах не зростала. Цей принцип гарантує збіжність зі швидкістю геометричної прогресії за значенням функції. Для реалізації цього принципу необхідна додаткова інформація про функцію $f(x)$ – значення функції в точці мінімуму f^* і так звані "константи росту" M та N.

Теорема 5 *Нехай $f(x)$ – опукла функція, визначена на E^n і в деякому сферичному околі S_d точки мінімуму x^*: $S_d = \{x : \| x - x^* \| \le d\}$ субградієнт задовольняє двосторонню нерівність*

$$N(f(x) - f(x^*)) \le (g_f(x), x - x^*) \le M(f(x) - f(x^*)), \tag{7}$$

де $M \ge N$ – невід'ємні константи. Тоді якщо в алгоритмі прийняти:

$$x_0 \in S_d,$$

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

$$1 < \alpha_{k+1} \le \frac{(M+N)}{(M-N)}, k = 0,1,2,\dots,$$

то для всіх $k = 0,1,\dots$ виконується нерівність

$$\| A_k(x_k - x^*) \| \le d. \tag{8}$$

З нерівності (8) випливає локалізація x^* в еліпсоїді Φ_k з центром у точці x_k. Відношення об'ємів еліпсоїдів Φ_{k+1} та Φ_k задається таким рівнянням

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

У випадку квадратичної додатньо визначеної функції в нерівності (7) можна вибирати $M = N = 2$. Для кусочно-лінійної функції, надграфік якої являє собою конус з вершиною в точці (x^*, f^*), можна вибирати $M = N = 1$. Для цих випадків $\beta_{k+1} = \beta = 0$, і алгоритм збігається за число кроків, що не перевищує n. Розв'язок невиродженої системи n лінійних рівнянь з n невідомими $(a_i, x) + b_i = 0, i = 1,...n,$ можна замінити на знаходження мінімуму $f(x) = \max_{1 \le i \le n} |(a_i, x) + b_i|$. Якщо взяти $f^* = 0$, $\beta_k = 0$ і застосувати метод (5)–(6), одержимо алгоритм, який відповідає відомій скінченній процедурі розв'язання лінійних алгебраїчних систем – методові ортогоналізації градієнтів.

Узагальнення теореми 5 були отримані Н.З. Шором і для деяких класів неопуклих функцій, що виникають при розв'язанні систем нелінійних рівнянь $f_i(x) = 0$, $i = 1,\ldots,n$. Для $f(x) = \max|f_i(x)|$ можна показати, що для регулярної точки x^* (розв'язку системи) (тобто точки, де функції $f_i(x)$ неперервно диференційовні в цій точці, а якобіан системи $I(x^*)$ є відмінним від нуля), для будь-якого $\delta > 0$ знайдеться достатньо малий окіл $S_d(x^*)$, такий, що константи M та N у (7) можна взяти рівними відповідно $M = 1+\delta; N = 1-\delta; \beta = \dfrac{M-N}{M+N} = \delta$. Як показано в роботі [2], при застосуванні граничного варіанту алгоритму з $\beta = 0$ і відновленням після кожних n ітерацій (великий цикл) при звичайних припущеннях щодо диференційовності й регулярності для розв'язання систем нелінійних рівнянь можна одержати квадратичну швидкість збіжності (для великих циклів).

Сімейство алгоритмів з розтягом простору в напрямку субградієнта містить як окремий випадок так званий метод еліпсоїдів. Метод еліпсоїдів був запропонований Д.Б. Юдіним та А.С. Немировським [12] на основі методів послідовних відсікань, а також, незалежно, Н.З. Шором [13], як окремий випадок алгоритму з розтягом простору в напрямку субградієнта. В алгоритмі використовуються такі параметри: коефіцієнт розтягу простору вибирається постійним і рівним

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

а регулювання кроку здійснюється за правилом:

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

де n – розмірність простору, r – радіус кулі з центром у точці x_0, що містить точку x^*. Метод еліпсоїдів збігається зі швидкістю геометричної прогресії по відхиленню найкращого досягнутого значення $f(x)$ від оптимального, при цьому знаменник геометричної прогресії асимптотично залежить тільки від розмірності простору n

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

Роботи з методів еліпсоїдів були продовжені й виконувалися Н.З. Шором та В.І. Гершовичем. Сімейство цих алгоритмів одержало назву "методи еліпсоїдів з глибокими відсіканнями". Досить повний огляд результатів про цей напрямок міститься в монографії [5].

Метод еліпсоїдів використаний Л.Г. Хачіяном [14] для побудови й обґрунтування першого поліноміального алгоритму розв'язання задачі лінійного програмування з раціональними коефіцієнтами. Крім того, метод еліпсоїдів набув важливого застосування в теорії складності алгоритмів дискретної оптимізації [15]. Конгрес з математичного програмування, що проводився в Бонні в 1982 році за участю Н.З. Шора, був зокрема присвячений методові еліпсоїдів та його застосувань. У вибраних працях конгресу [16] була опублікована 28-ми сторінкова доповідь Н.З. Шора "Generalized gradient methods of nondifferentiable optimization employing space dilatation operations” з оглядом по методів негладкої оптимізації, розроблених в Інституті кібернетики.

Досвід застосування алгоритмів з розтягом простору в напрямку градієнта показав суттєве прискорення збіжності субградієнтних процесів при використанні операторів, що змінюють метрику простору. У той же час труднощі вибору крокових множників стимулювали пошук нових методів негладкої оптимізації зі змінною метрикою, у яких вибір крокового множника пов'язаний із пошуком мінімуму уздовж напрямку. Цей клас алгоритмів описаний у наступному розділі.

3.5.3 Субградієнтні методи з розтягом простору в напрямку різниці двох послідовних субградієнтів

При розв’язанні складних задач недиференційовної оптимізації середнього розміру (до кількох сотень змінних) особливо ефективними виявилися алгоритми субградієнтного типу з розтягом простору в напрямку різниці двох послідовних субградієнтов (r-алгоритми). Вони були запропоновані в 1971 році Н.З. Шором і М.Г. Журбенком в роботі [17]. По своїй структурі та трудомісткості ітерації r-алгоритми близькі до методів з розтягом простору в напрямку субградієнта. Але між ними існує важлива різниця: методи УГС із розтягом простору в напрямку субградієнта в принципі не можуть бути монотонними, а r-алгоритми при певному регулюванні крокових множників та коефіцієнтів розтягу простору можуть стати такими. Це пов'язано з таким геометричним фактом: якщо ми знаходимося на границі двох частин кусочно-гладкої поверхні рівня, а градієнти до цих гладких частин, обчислені в даній точці, утворюють тупий кут, то ніякий розтяг простору в напрямку градієнтів не може перетворити цей кут у гострий, – він може лише наближатися до $\pi/2$, залишаючись тупим. Застосовуючи метод із розтягом простору в напрямку субградієнта, неможливо одержати напрямок спадання функції у вигляді антиградієнта до одної з частин у розтягнутому просторі. У той же час розтяг простору в напрямку різниці двох зазначених градієнтів з достатнім коефіцієнтом розтягу перетворює тупий кут між градієнтами на гострий, тобто відповідні образи цих антиградієнтів у розтягнутому просторі стають напрямками спадання функції.

Приведемо загальну схему r-алгоритмів для мінімізації опуклої функції $f(x)$, визначеної на E^n. Будемо припускати, що $f(x)$ має обмежену область мінімумів X^*, так що виконується $\lim\limits_{\|x\|\to\infty} f(x) = +\infty$.

Вибираємо початкове наближення $x_0 \in E^n$ і неособливу матрицю B_0 (найчастіше в якості B_0 вибирають одиничну матрицю I_n або діагональну матрицю D_n із невід’ємними елементами на діагоналі, за допомогою якої здійснюється масштабування змінних).

Перший крок алгоритму робимо по формулі $x_1 = x_0 - h_0\eta_0$, де $\eta_0 = B_0 B_0^T g_f(x_0)$, h_0 – деякий кроковий множник, обраний з умови існування в точці x_1 субградієнта $g_f(x_1)$, такого, що $(g_f(x_1), \eta_0) \le 0$. При $B_0 = I_n$ маємо $\eta_0 = g_f(x_0)$, і перший крок збігається з ітерацією субградієнтного процесу.

Нехай у результаті обчислень після k ($k = 1,2,...$) кроків процесу отримані обчислені значення $x_k \in E^n$ й матриці B_k розміру $n \times n$. Опишемо $(k+1)$-й крок процесу.

1. Обчислюємо наступні величини: $g_f(x_k)$ – субградієнт функції $f(x)$ в точці x_k; $r_k = B_k^T(g_f(x_k) - g_f(x_{k-1}))$ – вектор різниці двох послідовних субградієнтів у перетвореному просторі.

Перехід від початкового простору до перетвореного задається формулою $y = A_k x$, де $A_k = B_k^{-1}$. Визначимо функцію $\varphi_k(y) = f(B_k y)$, тоді $g_{\varphi_k}(y) = B_k^T g_f(x)$. Таким чином, r_k є різниця двох субградієнтів функції $\varphi_k(y)$, обчислених у точках $y_k = A_k x_k$ і $\tilde{y}_k = A_k x_{k-1}$.

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

3. Задаємо величину β_k, обернену коефіцієнтові розтягу простору α_k перед $(k+1)$-м кроком.

4. Обчислюємо $B_{k+1} = B_k R_{\beta_k}(\xi_k)$, де $R_{\beta_k}(\xi_k)$ – оператор розтягу простору на $(k+1)$-му кроці. Відмітимо, що $B_{k+1} = A_{k+1}^{-1}$.

5. Знаходимо $\tilde{g}_k = B_{k+1}^T g_f(x_k)$ – субградієнт функції $\varphi_{k+1} = f(B_{k+1} y)$ в точці $y_{k+1} = A_{k+1} x_k$.

6. Визначаємо

$$x_{k+1} = x_k - h_k B_{k+1} \tilde{g}_k / \| \tilde{g}_k \|. \tag{9}$$

Крок алгоритму (9) відповідає кроку узагальненого градієнтного спуску в перетвореному під дією оператора A_{k+1} просторі. Дійсно, застосувавши до обох частин формули (9) оператор A_{k+1}, одержимо

$$y_{k+1} = A_{k+1} x_{k+1} = \bar{y}_k - h_k \tilde{g}_k / \| \tilde{g}_k \|, \tag{10}$$

де $\bar{y}_k = A_{k+1} x_k$.

7. Переходимо до наступного кроку або закінчуємо роботу алгоритму при виконанні деяких умов завершення.

Практична ефективність алгоритму багато в чому залежить від вибору крокового множника h_k. В r-алгоритмі h_k вибирається з умови наближеного пошуку мінімуму $f(x)$ уздовж напрямку, при цьому під час мінімізації опуклих функцій слід дотримуватися умови $h_k \ge h_k^*$, де h_k^* – значення крокового множника, що відповідає мінімуму уздовж напрямку. Тобто необхідно, щоб напрямок субградієнта в точці x_{k+1} утворював нетупий кут із напрямком спуску з точки x_k.

При мінімізації негладких опуклих функцій найбільш вдалими виявилися такі варіанти алгоритму. Коефіцієнти розтягу простори α_k вибираються в межах 2–3, для крокового множника h_k застосовується адаптивний спосіб регулювання. Задається деяке натуральне число m, постійні $q > 1$ і $t_0^0 > 0$ (після k кроків ця величина буде позначатися відповідно t_k^0). Рухаємося з точки x_k в напрямку спуску з кроком t_k^0 доти, поки не буде виконана умова завершення спуску уздовж напрямку або число кроків не стане рівним m. Умова завершення спуску може полягати в тому, що значення функції в черговій точці не менше, ніж значення функції в попередній точці; інший варіант такої умови – похідна за напрямком спуску в даній точці є від’ємною. Якщо пройшло m кроків, а умову завершення спуску не виконано, то замість t_k^0 запам'ятовуємо $t_k^1 = qt_k^0$, де $q > 1$, і продовжуємо спуск у тому ж напрямку з великим кроком. Якщо після чергових m кроків умову завершення спуску не виконано, то замість t_k^1 беремо $t_k^2 = qt_k^1$ і т.д. В силу припущення $\lim\limits_{\|x\|\to\infty} f(x) = +\infty$ після скінченної кількості кроків у визначеному напрямку обов'язково виконається умова завершення спуску. Постійна кроку $t_k^{p_k} = q^{p_k} t_k^0$ $(p \in \{0,1,2,\ldots\})$, що використовувалася на останньому кроці, вважається початковою при спуску в новому напрямку з точки x_{k+1}, тобто $t_{k+1}^0 = t_k^{p_k}$.

Як показали численні обчислювальні експерименти й практичні розрахунки, у більшості випадків при $\alpha \in [2,3]$, $m = 3$ та зазначеному вище способі регулювання h число кроків уздовж напрямку в середньому рідко перевершує 2, при цьому за n кроків r-алгоритму точність по функції, як правило, поліпшується в 3–5 разів.

У випадку мінімізації гладкої функції для прискорення збіжності можна застосовувати більш тонкі способи пошуку мінімуму уздовж напрямку руху, наприклад, квадратичну апроксимацію по трьох точках, процес "золотого сечения" тощо. У гладкому випадку добре зарекомендував себе адаптивний спосіб регулювання кроку уздовж напрямку, подібний приведеному вище, з невеликою модифікацією: якщо на даній ітерації функція вже після першого кроку прийняла більше значення, то кроковий множник домножується на задане число, що менше одиниці (наприклад, в межах 0,8 – 0,95). Це пов'язано з тим, що в гладкому випадку швидкість збіжності може виявитися більшою при точнішому знаходженні мінімуму уздовж напрямку, а додаткове подрібнення кроку сприяє збільшенню точності пошуку мінімуму уздовж напрямку.

Для задач гладкої оптимізації r-алгоритм по своїй формальній структурі близький до алгоритмів квазиньютонівського типу зі змінною метрикою. Так граничний варіант r-алгоритму з нескінченним коефіцієнтом розтягу (тобто $\beta = 0$) є проективним варіантом методу спряжених градієнтів. У роботі [17] показано, що граничний варіант r-алгоритму має квадратичну швидкість збіжності при звичайних умовах гладкості й регулярності. Відмітимо, що на основі використання операторів розтягу простору можна побудувати сімейство квазиньютонівських методів і при кінцевих значеннях коефіцієнтів розтягу [18, 19]. Ці алгоритми характеризуються чисельною стійкістю по відношенню до точності пошуку мінімуму уздовж напрямку.

У 2002 р. Н.З. Шором була розроблена монотонна модифікація r-алгоритму [20]. Від r-алгоритму з адаптивним регулюванням кроку вона відрізняється тим, що:

- крок у напрямку спуску визначається з умови

$$h_k^{\min} = \arg\min_{h_k} f(x_k - h_k B_k B_k^T g(x_k));$$

- якщо виявилося, що $h_k^{\min} = 0$, то $x_{k+1} = x_k$. У цьому випадку змінюється лише обернена матриці B_{k+1}. Крім того, якщо субградієнт g_{k+1} можна вибрати неоднозначно, то вибирається той, проекція якого на напрямок спуску мінімальна. Це дозволяє уникнути "пасток" у точках негладкості, що зв'язані з границями частин гладких функцій;
- використовується допоміжна спеціальна квадратична задача для одержання f_* – оцінки знизу для функції, що мінімізується.

У роботі [20] показано, що монотонна модифікація r-алгоритму дозволяє з високою точністю вирішувати широкий клас оптимізаційних задач, зокрема мінімаксні задачі, спеціальні квадратичні задачі, задачі максимального розрізу графа, задачі оптимального керування з дискретним часом, поліноміальні багатоекстремальні задачі.

За останні 30 років у ІК ім. В.М. Глушкова НАН України накопичено значний досвід розв'язання оптимізаційних задач за допомогою r-алгоритму. Зусиллями Н.З. Шора та його учнів М.Г. Журбенка, Л.П. Шабашової, В.І. Гершовича, П.І. Стецюка, О.П. Лиховида, О.В. Кунцевича та інших розроблено кілька модифікацій r-алгоритму для до розв'язання різноманітних задач оптимізації. r-алгоритм використовувався в задачах оптимізації великої розмірності, у блокових задачах з різними схемами декомпозиції, при розв'язанні мінімаксних та матричних задач оптимізації, для обчислення двоїстих лагранжевих оцінок у багатоекстремальних і комбінаторних задачах оптимізації. На практиці він застосовувався для розв'язання задач оптимального планування, оптимального проектування, синтезу й аналізу мереж, відновлення зображень, еліпсоїдальній апроксимації та локалізації тощо [2, 21, 22].

3.5.4 Застосування алгоритмів недиференційовної оптимізації

Приведемо головні джерела задач негладкої оптимізації та дамо коротку характеристику застосувань субградієнтних методів оптимізації до розв'язання цих класів задач.

По-перше, це задачі математичного програмування великої розмірності з блоковою структурою та порівняно невеликим числом зв'язків між блоками. Використання схем декомпозиції для розв'язання таких задач приводить до задач мінімізації, як правило, негладких функцій від з'єднуючих змінних або від множників Лагранжа. Блокові задачі розглянуті в монографіях [2, 5, 23, 24]. Так у монографіях [2, 5] розглянуто різноманітні задачі оптимального планування і проектування, задачі довгострокового та поточного виробничо-транспортного планування, багатоіндексні розподільчі (рос.: распределительные) задачі в лінійній, нелінійній і динамічній постановках тощо. Активну участь у розробці алгоритмів розв'язання цих задач взяли співробітники Інституту кібернетики М.Г. Журбенко, В.І. Гершович, Т.В. Бєлих, Л.В. Бєляєва, Г.І. Горбач та ін. Монографія Н.З. Шора та Д.І. Соломона [23] присвячена методам рішення блокових задач дробно-лінійного програмування великої розмірності. Тут розглянуто такі практичні задачі оптимізації: визначення клієнтурного плану та структури рухомого складу автотранспортних підприємств, маршрутизація автомобільних вантажних перевезень. Монографія [24] містить застосування схем декомпозиції при розв'язанні ряду задач оптимального проектування й маршрутизації в мережах з урахуванням відмов окремих компонентів мережі та зміни вимог до потоків. Це був останній великий проект[1] під керівництвом Н.З. Шора. Участь у цих роботах приймали співробітники його відділу П.І. Стецюк, Ю.П. Лаптін, М.Г. Журбенко, Т.О. Бардадим, О.П. Лиховид, О.А. Березовський.

По-друге, це задачі мінімізації функції максимуму. Нехай задане параметричне сімейство опуклих функцій, визначених на E^n, $\{f_\alpha(x)\}_{\alpha\in A}$. Функція максимуму $F(x)$ є результатом операції максимуму по параметру α:

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

Область визначення $dom F$ функції $F(x)$ збігається з такими значеннями $x \in E^n$, при яких $\{f_\alpha(x)\}$ обмежена зверху по α. Для кожного $\bar{x} \in dom F$ визначимо підмножину індексів

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

Субградієнтна множина $G_F(\bar{x})$ функції F в точці $\bar{x}$ визначається за формулою

$$G_F(\bar{x}) = \overline{conv}\{\cup_{\alpha\in I(\bar{x})} G_{f_\alpha}(\bar{x})\}, \tag{11}$$

де $\overline{conv}\{M\}$ позначає операцію знаходження мінімальної опуклої замкнутої множини, що містить M, $G_{f_\alpha}(\bar{x})$ – субградієнтні підмножини функцій f_α у точці $\bar{x}$, $\alpha \in I(\bar{x})$. Якщо всі функції $f_\alpha(\alpha \in I(\bar{x}))$ є диференційовними в точці $\bar{x}$, то множина $G_{f_\alpha}(\bar{x})$ складається з єдиної точки, що збігається з градієнтом $g_{f_\alpha}(\bar{x})$, і формула (11) приймає такий вигляд:

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

У випадку, коли $I(\bar{x})$ – скінченна множина, усі крайні точки множини $G_F(\bar{x})$ є градієнтами деяких функцій f_α, $\alpha \in I(\bar{x})$, у точці $\bar{x}$, і $G_f(\bar{x})$ являє собою опуклий багатогранник.

До цього класу задач відносяться задачі мінімізації функції максимуму, характерні для моделей ігрового характеру та багатокритеріальних моделей оптимального планування й дослідження операцій. До такого роду задач зводяться задачі розв’язання систем рівнянь і нерівностей, визначення коефіцієнтів нелінійної регресії з чебишевським критерієм мінімізації максимуму нев'язки. Інформацію про такі задачі можна знайти в монографії [2], де, наприклад, описані дві модифікації методу узагальненого градієнтного спуску з розтягом простору, розроблені Н.З. Шором і Л.П. Шабашовою для розв’язання мінімаксних задач. Ці модифікації застосовувалися до розв’язання практичних задач, пов'язаних із проблемою побудови оптимальних еталонів. Наведено результати для складної максимінної задачі, де число змінних дорівнювало 1200, а функцій, що використовуються в операції взяття мінімуму, було 1536. Другий приклад пов'язаний із розв’язанням задач інтерпретації гравіметричних спостережень (Н.З. Шор, І.Г. Овруцький).

Третє джерело негладких задач – це лагранжеві оцінки в задачах математичного програмування. Розглянемо досить загальну задачу математичного програмування, обмеження якої складаються з двох частин: одна має загальний вигляд умови приналежності $x \in X \subseteq E^n$, а друга визначається системою

---

[1] проект Науково-технологічного центру в Україні "Математичні та програмні засоби оптимального проектування структур надійних мереж" (№ 1625)

рівностей: знайти

$$f_0^* = \inf_{x \in X} f_0(x),\ X \subseteq E^n, \tag{12}$$

при обмеженнях

$$f_i(x) = 0,\ i = 1,\ldots,m. \tag{13}$$

Припустимо, що X – замкнута підмножина n-мірного евклідового простору, $f_i(x)$ – неперервні функції, визначені на X. Для оцінки f_0^* уведемо функцію Лагранжа

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

де $u = \{u_1,\ldots,u_m\}$ – вектор множників Лагранжа. Розглянемо оцінку

$$\psi(u) = \inf_{x \in X} L(x,u).$$

При будь-якому допустимому x та довільному u маємо $\psi(u) \le f_0(x)$, звідки випливає, що $\psi(u) \le f_0^*$.

Задача знаходження найкращої оцінки для оптимального значення задачі (12)–(13) у даному класі лагранжевих оцінок зводиться до розв'язання такої задачі оптимізації:

$$\psi^* = \sup_{u \in E^m} \psi(u).$$

Функція $\psi(u)$ є увігнутою як результат операції мінімізації по $x \in X$ параметричного сімейства лінійних по u функцій. Припустимо, що $\psi(u)$ – власна увігнута функція з непорожньою областю $dom\psi$, що має внутрішні точки. Нехай $\bar{x}$ – деяка внутрішня точка $dom\psi$, тобто $\bar{x} \in int\,dom\psi$. Тоді за правилами обчислення субградієнта від функції максимуму $\varphi(u) = -\psi(u)$ субградієнтна множина $G_\varphi(\bar{x})$ визначається в таким чином:

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

де $X(u)$ – множина всіх розв'язків локальної задачі $\inf_{x \in X} L(x,\bar{u})$; $F(x(\bar{u}))$ – вектор "нев'язок", що відповідає розв'язку $x(\bar{u})$,

$$F(x(\bar{u})) = \{f_1(x(\bar{u})),\ldots, f_m(x(\bar{u}))\}.$$

Якщо в точці $\bar{x}$ локальна задача має один розв'язок, то функція $\varphi(u)$ у відповідній точці є диференційовною, а її градієнт збігається з вектором нев'язок $\{-f_i(x(\bar{u}))\}_{i=1}^{m}$. Інакше градієнт функції $\varphi(u)$ у відповідній точці терпить розрив.

Обчислення оцінок виду ψ^*, коли функції $f_\nu(x), \nu = 0,1,\ldots,m$ в задачі (12)–(13) є квадратичними (можливо неопуклими), а $X \equiv E^n$, буде докладно розглянуто в розділі 3.5.5. Неопуклі квадратичні моделі зустрічаються в багатьох випадках, наприклад у поліноміальних задачах оптимізації та NP-складних екстремальних задачах на графах. Оцінка виду ψ^* для цих задач відіграє важливу роль, тому що дає можливість у рамках NP-складних неопуклих квадратичних задач, виділити такі підкласи, для яких задачу знаходження значення глобального мінімуму цільової функції можна розв'язати за поліноміальний час.

Четверте джерело – це задачі матричної оптимізації. Вони тісно пов'язані зі знаходженням оцінок виду ψ^* для неопуклих квадратичних задач. Найбільш типовими прикладами негладких функцій є максимальне (мінімальне) власне число симетричної матриці, а також сума (зважена сума) k найбільших власних чисел симетричної матриці.

Нехай S_n – клас дійсних симетричних $n \times n$-матриць. Нехай $X = \{x_{ij}\}_{i,j=1}^n$ – довільна матриця з S_n і

$$\lambda_1(X) \geq \lambda_2(X) \geq \ldots \geq \lambda_n(X),$$

власні числа матриці X, впорядковані за незростанням.

Для знаходження максимального власного числа $\lambda_1(X)$ відома формула Релєя-Рітца

$$\lambda_1(X) = \max_{\|y\|=1}(Xy, y) = \max_{\|y\|=1}\sum_{i,j=1}^n x_{ij} y_i y_j, \tag{14}$$

де $y = (y_1, \ldots, y_n) \in E^n$. $\lambda_1(X)$ – опукла функція, визначена на S_n, тому що формула (14) дає представлення цієї функції як максимума сімейства лінійних функцій при $\{x_{ij}\}_{i,j=1}^n$. Позначимо $Y^*(X)$ набір нормованих векторів y, при яких досягається максимум у (14), тобто $\lambda_1(X) = (Xy^*, y^*)$ для усіх $y^* \in Y^*(X)$. З (14) можна одержати множину субградієнтів $G_{\lambda_1}(\overline{X})$ функції $\lambda_1(\cdot)$ в точці $\overline{X}$:

$$G_{\lambda_1}(\overline{X}) = \overline{conv}\left\{\cup_{y \in Y^*(\overline{X})} yy^T\right\}.$$

Обчислення субградієнта $g_{\lambda_1}(\overline{X}) \in G_{\lambda_1}(\overline{X})$ може бути зведене до знаходження довільного $y^*(\overline{X}) \in Y^*(\overline{X})$ і визначено по формулі:

$$G_{\lambda_1}(\overline{X}) = \left\{y^*(\overline{X})[y^*(\overline{X})]^T\right\}$$

Якщо $\lambda_1(\overline{X})$ має коефіцієнт кратності, рівний одиниці, то $g_{\lambda_1}(\overline{X})$ єдиний, і функція $\lambda_1(X)$ – диференційовна в точці $\overline{X}$. Якщо ж коефіцієнт кратності $\lambda_1(\overline{X})$ більше одиниці, то функція $\lambda_1(X)$ – є недиференційовною в $\overline{X}$.

Аналогічні властивості мають і цікаві класи опуклих матричних функцій, визначених на симетричних матрицях $X \in S_n$, наприклад, сума k найбільших власних чисел:

$$S_{n,k}(X) = \sum_{r=1}^k \lambda_r(X), 1 \leq k \leq n,$$

та зважена сума k найбільших власних чисел:

$$S_{n,k}^w(X) = \sum_{i=1}^k w_i \lambda_i(X), \quad \text{де } w = (w_1, \ldots, w_k) \geq 0.$$

Якщо $w_1 \geq w_2 \geq \ldots \geq w_k$, то $S_{n,k}^w(X)$ є опуклою функцією, визначеною на S_n. Для опуклих функцій $S_{n,k}(X)$ і $S_{n,k}^{w}(X)$ множину субградієнтів можна представити таким способом:

$$G_{S_{n,k}}(X)=\overline{conv}\{\sum_{i=1}^{k} y_i y_i^T\} \text{ та } G_{S_{n,k}^{w}}(X)=\overline{conv}\{\sum_{i=1}^{k} w_i y_i y_i^T\},$$

де $\{y_i\}_{i=1}^{k}$ – довільна ортогональна система власних векторів матриці X. Кожному вектору y_i відповідає $\lambda_i(X)$. У загальному випадку, якщо для симетричної матриці X виконується $\lambda_k(X)>\lambda_{k+1}(X)$, то $S_{n,k}$ диференційовна в точці X, у противному випадку функція $S_{n,k}$ недиференційовна в точці X. Якщо $\lambda_k(X)=\lambda_{k+1}(X)$, то і функція $S_{n,k}^{w}$ є недиференційовною у точці X. Крім того, вона буде недиференційовною у точці X і тоді, коли існує такий індекс $i=1,\ldots,k-1$, для якого $w_i > w_{i+1}$ і $\lambda_i(X)=\lambda_{i+1}(X)$.

Матричні задачі мають практичне застосування в теорії стійкості й оптимізації динамічних систем. Наприклад, до них відносяться деякі задачі теорії стійкості динамічних систем, що описуються системами диференціальних рівнянь. Показниками функціонування таких систем є різні характеристики: час перехідного процесу, розбіжність траєкторій через збурювання, область стійкості тощо. Прямий метод Ляпунова теоретично дає можливість одержувати досить точні оцінки таких характеристик. Однак не існує загального конструктивного підходу до побудови функцій Ляпунова. Тому оптимізація за визначеними критеріями функцій Ляпунова становить значний інтерес. У монографіях [21, 22] дані приклади матричних задач, пов'язаних з одержанням оптимальної функції Ляпунова і методи їхнього розв'язання за допомогою алгоритмів недиференційовної оптимізації. До матричних задач відносяться й задачі побудови оптимальних за об'ємом вписаних у багатогранник та описаних навколо багатогранника еліпсоїдів. Оригінальні алгоритми для розв'язання цих задач розроблені Н.З. Шором разом з учнями С.І. Стеценком та О.А. Березовським. Ці результати ввійшли в монографію [22].

П'яте джерело складають задачі нелінійного програмування, для розв'язання яких використовується метод негладких штрафних функцій. Негладкі штрафні функції мають безсумнівну перевагу в порівнянні зі звичайними гладкими функціями штрафу: при використанні негладких штрафних функцій немає необхідності спрямовувати штрафні коефіцієнти до $+\infty$.

Шосте джерело – це задачі оптимального керування з неперервним та дискретним часом. Використання принципу максимуму або дискретного принципу максимуму в багатьох випадках приводить до задач мінімізації функцій з розривним градієнтом. Ці задачі можна розглядати як спеціальні задачі нелінійного програмування, для розв'язання яких можна застосувати схеми декомпозиції та метод негладких штрафних функцій.

Сьоме джерело задач – це задачі дискретного програмування або задачі змішаного дискретно-неперервного типу. Багато задач такого роду досить успішно можна розв'язувати методом гілок та границь з одержанням оцінок шляхом розв'язання двоїстої задачі. Двоїста задача звичайно виявляється задачею мінімізації опуклої кусочно-лінійної функції з величезним числом "кусків" при простих обмеженнях, тобто задачею негладкої оптимізації.

І нарешті, функції з розривним градієнтом можуть безпосередньо входити в модель задачі оптимального планування, проектування або дослідження операцій як результат кусочно-гладкої апроксимації техніко-економічних характеристик реальних об'єктів.

Слід зазначити, що з прикладної точки зору різкої границі між негладкими й гладкими функціями немає. З позицій прикладної математики та обчислювальної практики 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.

Численні застосування субградієнтних алгоритмів негладкої оптимізації для зазначених класів задач можна знайти в монографіях [2, 5, 21, 22]. На їхній основі розроблене програмне забезпечення для важливих класів задач оптимального планування, проектування і керування. Так, наприклад, задачі траспортного типу великої розмірності розв'язувалися для розподілу судів по лініях річкових перевезень, задачі розподільного типу особливо великої розмірності розв'язувалися при плануванні завантаження прокатних станів (Н.З. Шор, Г.І. Горбач). У цих задачах кількість замовлень досягала 10 000, число станів перевищувало 50, номенклатура складала близько 1000 найменувань видів продукції. Запропоновані алгоритми були використані при плануванні завантаження прокатних станів у "Союзглавметалі" Держпочстачу СРСР зі значним економічним ефектом. Програми використовувалися при створенні системи автоматизованого планування виробництва та розподілу замовлень у чорній металургії, для задачі розподілу коксівного вугілля між коксохімічними заводами (Н.З. Шор і В.І. Гершович). Разом з ДержНДІЦА МінЦА СРСР розроблено комплекс економіко-математичних моделей та алгоритмів оптимізації для систем поточного і перспективного планування цивільної авіації. Ці алгоритми були програмно реалізовані, були проведені розрахунки по перспективах розвитку парку цивільної авіації (Н.З. Шор, Г.М. Юн, М.Г. Журбенко, С.К. Андрусенко).

Методи недиференційовної оптимізації зіграли велику роль при створенні пакетів прикладних програм (ППП). Теоретичні дослідження в області побудови автоматизованих програмних засобів почалися ще в 50-х роках з робіт В.М. Глушкова і привели до розробки складних діалогових пакетів, призначених для розв'язання різних класів задач оптимізації. Програми з використанням субградієнтних методів включені в ППП ПЛАНЕР, ДИСПРО та ДИСНЕЛ, розроблені в Інституті кібернетики у 80-і роки для машин серії ЄС ЕОМ. У цих пакетах був реалізований широкий спектр методів оптимізації для розв'язання і дослідження широкого кола питань оптимального планування, проектування й керування, розміщення та реконструкції виробництва, проектування технічних пристроїв і машин, складання розкладів виконання робіт в умовах обмеженості ресурсів. У [25] подано огляд розроблених в Інституті кібернетики методів оптимізації, що стали центральними при реалізації зазначених пакетів прикладних програм. У [26] описані математичні моделі прикладних задач і системне математичне забезпечення ППП ПЛАНЕР. Більшість з цих задач в удосконаленій формі були включені в ППП ДИСНЕЛ. У [27] описано призначення, класи розв'язуваних задач, системне й алгоритмічне забезпечення для ППП ДИСПРО-3. Тут методи негладкої оптимізації застосовувалися головним чином при розв'язанні "оціночних" задач в методі гілок та границь для спеціальних класів дискретних задач.

В останні десять років методи недиференційовної оптимізації також знайшли ряд важливих практичних застосувань. Для розв'язання спеціальних класів задач двохетапного стохастичного програмування з простою та фіксованою рекурсією було розроблено програми Shor1 (автори – Н.З. Шор і О.П. Лиховид) та Shor2 (автори – Н.З. Шор і М.Г. Журбенко). Вони базуються на використанні схеми декомпозиції за змінними та розв'язанні координуючої негладкої задачі за допомогою r-алгоритму з адаптивним регулюванням кроку. Програми Shor1 і Shor2 включені до складу системи моделювання для задач стохастичного лінійного програмування SLP–IOR[2], розробленої в Інституті дослідження операцій, Цюріх, Швейцарія. Було розроблено також програмне забезпечення для ряду задач оптимального проектування й маршрутизації в мережах з урахуванням можливого виходу з ладу окремих компонентів мережі та зміни вимог до потоків. Його опис включений у монографію [24]. Методи негладкої оптимізації використовувалися для розв'язання ряду задач проектування енергетичних установок [28]. Ці роботи були виконані Ю.П. Лаптіним та М.Г. Журбенком, а результати були впроваджені в Харківському ЦКБ "Енергопрогрес". r-алгоритми та його модифікації використані О.М. Кисельовою при розв'язанні неперервних задач оптимальної розбивки множин n-мірного евклідова простору, що зводяться до нескінченновимірних задач математичного програмування з булевими змінними [29] .

3.5.5 Лагранжеві оцінки в багатоекстремальних квадратичних задачах

Багато задач поліноміальної та дискретної оптимізації зводяться до багатоекстремальних задач нелінійного програмування квадратичного типу. Для розв'язання цих задач можна використовувати лагранжеві оцінки в комбінації з методом гілок та границь. Лагранжеві оцінки можна покращувати, додаючи в модель так звані функціонально надлишкові квадратичні обмеження. Задачі знаходження лагранжевих оцінок, як правило, зводяться до мінімізації негладких опуклих функцій і можуть успішно розв'язуватися сучасними методами недиференційовної оптимізації. У даному розділі описано цей підхід і наведено приклади його використання для задач поліноміального типу та деяких екстремальних задач на графах.

Розглянемо оптимізаційну задачу квадратичного типу в такому формулюванні: потрібно знайти

$$q^* = \inf_{x \in E^n} Q_0(x) \tag{15}$$

при обмеженнях

$$Q_i(x) = 0, \quad i \in I = \{1,\dots,m\}. \tag{16}$$

Тут $Q_\nu(x)$ – квадратичні функції

$$Q_\nu(x) = (K_\nu x, x) + (b_\nu, x) + c_\nu,$$

де K_ν – симетричні матриці розміру $n \times n$, $b_\nu \in E^n$, c_ν – скаляри, $\nu \in \{0 \bigcup I\}$. Задача (15), (16) є окремим випадком задачі (12)–(13), коли $X \equiv E^n$, а всі функції $f_\nu(x), \nu = 0,1,\dots,m$ є квадратичними (частина з них може бути і лінійними).

У загальному випадку задача (15), (16) є багатоекстремальною й відноситься

[2] http://www.ior.uzh.ch/Pages/English/Research/StochOpt/SLP_Users_Guide.pdf

до класу NP-складних задач. Оцінки знизу для q^* можна одержати шляхом такої лагранжевої релаксації. Нехай $u = (u_1, \ldots, u_m) \in E^m$ – вектор множників Лагранжа задачі (15), (16). Розглянемо функцію Лагранжа

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

і функцію

$$\psi(u) = \inf_{x \in E^n} L(x,u) = \inf_{x \in E^n} [(K(u)x, x) + (b(u), x) + c(u)],$$

де

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

Нехай $\Omega^+ = \{u : \lambda_n[K(u)] > 0\}$ – підмножина u, для яких матриця $K(u)$ додатньо визначена; Ω^0 – підмножина u, для яких $\lambda_n[K(u)] = 0$, тут $\lambda_1(K) \geq \lambda_2(K) \geq \ldots \geq \lambda_n(K)$ – власні числа симетричної матриці K розміру $n \times n$, упорядковані в порядку незростання. Область визначення $dom\psi$ функції $\psi(u)$ складається з Ω^+ та підмножини точок $u \in \Omega^0$, для яких має розв'язок система рівнянь

$$2K(u)x + b(u) = 0. \tag{17}$$

Для інших точок $\psi(u) = -\infty$. Якщо $dom\psi \neq \varnothing$, то існує нетривіальна оцінка знизу

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

для величини q^* (умова $\psi^* = +\infty$ означає, що система (16) несумісна). Оцінку ψ^* з будь-якою заданою точністю можна знайти за поліноміальний час за допомогою методів мінімізації опуклих недиференційовних функцій, наприклад, методу еліпсоїдів або r-алгоритмів [22]. Схема використання $r(\alpha)$-алгоритму з адаптивним регулюванням кроку докладно описана у роботі [30].

Якщо ψ^* досягається на $u^* \in \Omega^+$, то

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

де $x(u^*)$ –- розв’язок системи (17) при $u = u^*$. Інакше ψ^* досягається на границі області Ω^+, при цьому може існувати так званий "розрив двоїстості"

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

Один із способів зменшення розриву Δ^* пов'язаний з уведенням функціонально надлишкових обмежень (при цьому може збільшитися і кількість змінних). Функціонально надлишковими обмеженнями назвемо обмеження, додавання яких залишає множину оптимальних розв’язків початкової задачі незмінною. Однак при цьому змінюється функція Лагранжа, що може в деяких випадках зменшити розрив між оптимальним значенням q^* цільової функції та лагранжевою (двоїстою) оцінкою ψ^*. Якщо додати у задачу (15), (16) квадратичні функціонально надлишкові обмеження

$$Q_{m+1}(x)=0,\dots,Q_{m+r}(x)=0, r\ge 1,$$

то нова задача матиме вигляд: знайти

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

при обмеженнях

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

Їй відповідає довший вектор множників Лагранжа

$$U=\{\{u\},u_{m+1},\dots,u_{m+r}\},$$

а функція Лагранжа для неї буде мати такий вигляд:

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

Оскільки $L(x,u)=L_1(x,(\{u\},0,\dots,0))$, $\psi_1(\{u\},0,\dots,0)=\psi(u)$, то

$$\psi_1^*=\sup_{U\in dom\psi_1}\psi_1(U)\ge\sup_{u\in dom\psi}\psi(u)=\psi^*,$$

тобто функціонально надлишкові обмеження можуть покращити лагранжеві оцінки.

Обмеження, що є лінійними комбінаціями вже існуючих обмежень, не впливають на точність оцінки ψ_1^*. Внесок таких обмежень у функцію Лагранжа еквівалентний лише певній зміні множників Лагранжа при існуючих обмеженнях. Однак, додавання функціонально надлишкових обмежень, що є нетривіальними наслідками умов задачі, у ряді випадків приводить до того, що нижня оцінка ψ_1^* навіть може стати точною для q^*. Прикладами функціонально надлишкових обмежень можуть служити такі:

a) квадратичні наслідки лінійних обмежень: наприклад, квадратичне обмеження у формі $(b_i^Tx+c_i)(b_j^Tx+c_j)\ge 0$ є наслідком із двох лінійних обмежень-нерівностей: $b_i^Tx+c_i\ge 0$ і $b_j^Tx+c_j\ge 0$;

b) квадратичні обмеження, що характеризують неоднозначність представлення добутку трьох, чотирьох, або більшого числа змінних задачі. Як правило, вони мають місце при зведенні поліноміальної задачі до квадратичної. Наприклад, нехай маємо змінні x_1, $x_2=x_1^2$ та $x_3=x_1^3$. Тоді квадратичне обмеження у формі рівності $x_2^2-x_1x_3=0$ є наслідком неоднозначного представлення x_1^4, а саме $x_1^4=(x_1^2)^2=(x_1^3)(x_1)$;

c) квадратичні обмеження, що є наслідками булевості або бінарності змінних задачі. Наприклад, для бінарних змінних $x_i^2=1$, $x_j^2=1$ та $x_k^2=1$ завжди виконується квадратична нерівність: $x_ix_j+x_ix_k+x_jx_k\ge -1$.

Докладну інформацію про зазначені сімейства функціонально надлишкових обмежень можна знайти в монографіях [21, 22], а також у роботах [31, 32, 33, 34, 35]. Монографії [21, 22] містять інформацію про обмеження a) та b**)** та приклади їх використання в поліноміальних і дискретних задачах. У роботах [31, 32] для екстремальних задач на графах (максимальна стійка множина вершин графа, максимальний розріз графа тощо) досліджені різні сімейства функціонально надлишкових квадратичних обмежень, що враховують специфічні особливості кожної конкретної задачі для того або іншого сімейства графів. У роботах [33, 34] для бінарних і булевих квадратичних задач розглядаються загальні схеми генерації функціонально надлишкових обмежень у формі квадратичних рівностей, що є наслідками булевості або бінарності змінних. У роботі [35] розглянуті правила для генерації сімейств функціонально надлишкових квадратичних обмежень у формі нерівностей, що є наслідком використання непарної кількості бінарних змінних. Їх можна використовувати як у бінарних, так і в булевих квадратичних задачах.

Історично ідея функціонально надлишкових обмежень походить з робіт Н.З. Шора [36, 37], де описано спосіб одержання оцінок знизу для цільової функції в поліноміальних задачах математичного програмування. Задача мінімізації поліноміальної функції зводиться до квадратичної оптимізаційної задачі, для якої будуються лагранжеві оцінки. При побудові необхідних і достатніх умов точності лагранжевих оцінок для окремих поліноміальних задач використовувалися функціонально надлишкові обмеження.

Однією з таких є задача знаходження глобального мінімуму поліноміальної функції $P(x)$ від однієї або декількох перемінних. Ця задача спеціальним образом зводиться до квадратичних багатоекстремальних задач (на мінімум) з багатьма функціонально надлишковими квадратичними обмеженнями. У роботі [22] доведено, що лагранжева оцінка для таких квадратичних задач збігається зі значенням p^* (значення полінома $P(x)$ в точці глобального мінімуму) тоді й тільки тоді, коли поліном $\overline{P}(x) = P(x) - p^*$ може бути представлений як сума квадратів інших поліномів. Ці результати мають відношення до класичних робіт Д. Гільберта з розкладу невід'ємних поліноміальних форм у суму квадратів. Розроблений метод дає можливість не тільки довести існування такої декомпозиції (якщо вона існує), але і знайти одне з можливих представлень полінома $\overline{P}(x)$ у вигляді суми квадратів інших поліномів. При цьому одержуємо також значення глобального мінімуму полінома $P(x)$.

Лагранжеві оцінки та покращення цих оцінок за допомогою функціонально надлишкових обмежень знайшли застосування й для ряду екстремальних задач на графах, які можна описати за допомогою квадратичних моделей [21, 22, 31, 32]. Багато комбінаторних оптимізаційних задач формулюються у вигляді булевих задач лінійного програмування, а відповідні двоїсті оцінки можуть бути отримані лагранжевою релаксацією цих ЛП-задач. Але в ряді випадків формулювання таких комбінаторних задач у вигляді нелінійних задач квадратичного типу дає точніші двоїсті оцінки. Велика кількість прикладів такого роду міститься в монографії [22].

У [22, 31, 32] наведено результати застосування техніки функціонально надлишкових обмежень для дослідження різних класів NP-складних комбінаторних булевих задач, сформульованих у формі неопуклих квадратичних задач оптимізації. Умова булевості змінної $x \in \{0,1\}$ представляється квадратичною рівністю $x^2 - x = 0$. Це такі екстремальні задачі на графах, як задача знаходження максимальної зваженої незалежної множини вершин графа, задача знаходження максимального розрізу графа, оптимальна бісекція графа, мінімальне розбиття графа на k частин з фіксованим числом вершин у кожній частині тощо. Дуже цікавими виявилися результати, пов'язані з задачею знаходження зваженої максимальної незалежної множини вершин графа. Тут лагранжеві оцінки тісно пов'язані з відомими числами Ловаса $\vartheta(G,w)$ та $\vartheta'(G,w)$ [15]. Ці результати містяться в монографії Н.З. Шора [22].

Нехай $G=(V,E)$ – зважений неорієнтований граф з множиною вершин V і множиною ребер E, вага кожної вершини $i \in V$ задається додатнім числом w_i. Підмножина вершин $S \subseteq V$ називається стійкою (або незалежною) множиною графа G, якщо для будь-яких $i, j \in S$ ребро (i, j) не належить E. Зважене число стійкості графа G визначається як $\alpha(G,w) = \max \sum_{i \in S} w_i$, де $S \subseteq V$ – стійка множина. Множина S^*, на якій досягається $\alpha(G,w)$, називається максимальною зваженою стійкою (або незалежною) множиною. Задача знаходження $\alpha(G,w)$ є NP-складною навіть в окремому випадку, коли усі ваги ребер дорівнюють одиниці [38]. Обчислення верхніх оцінок, що досить добре апроксимують зверху $\alpha(G,w)$, має як практичний, так і теоретичний інтерес. Н.З. Шор запропонував три верхніх оцінки для $\alpha(G,w)$.

Першу оцінку умовимося позначати $\psi_1^*(G,w)$. Вона пов'язана з формулюванням задачі про максимальну зважену стійку множину графа у вигляді такої квадратичної булевої задачі: знайти

$$\alpha(G,w) = \max \sum_{i \in V} w_i x_i \tag{18}$$

при обмеженнях

$$x_i x_j = 0 \quad \forall (i,j) \in E, \tag{19}$$

$$x_i^2 - x_i = 0 \quad \forall i \in V. \tag{20}$$

Тут булева змінна $x_i \in \{0,1\}$ дорівнює одиниці, якщо вершина $i \in V$ включається в стійку множину, і нулеві у противному випадку. Булеві змінні для усіх вершин описані квадратичними обмеженнями-рівностями (20). Квадратичні обмеження (19) означають, що якщо дві вершини зв'язані ребром у графі G, то вони обидві не можуть одночасно належати стійкій множині.

Друга оцінка (умовимося позначати її $\psi_2^*(G,w)$) відповідає квадратичній задачі, у якій до обмежень (19)–(20) додано таке сімейство функціонально надлишкових обмежень

$$x_i x_j \ge 0 \quad \forall (i,j) \notin E. \tag{21}$$

Вони є наслідками нерівностей $x_i = x_i^2 \ge 0$ для усіх $i \in V$. У [22] доведено, що $\psi_1^*(G,w) = \vartheta(G,w)$ і $\psi_2^*(G,w) = \vartheta'(G,w)$, де $\vartheta(G,w)$ і $\vartheta'(G,w)$ – відомі числа Ловаса [15]. Тобто оцінки Шора мають такі ж властивості як і числа Ловаса. Наприклад, $\psi_1^*(G,w) = \alpha(G,w)$, якщо граф G належить сімейству перфектних графів.

Найбільш цікавою виявилася третя оцінка Шора (умовимося позначати її $\psi_3^*(G,w)$). Вона пов'язана з квадратичною задачею, у якій до обмежень (19)–(21) додано таке сімейство функціонально надлишкових обмежень

$$x_i x_k + x_j x_k \le x_k, \forall (i,j) \in E, k \ne i, j. \tag{22}$$

Наявність обмежень (22) додає оцінці $\psi_3^*(G,w)$ ряд чудових властивостей, що пов'язані зі спеціальними сімействами графів. Так, наприклад, у роботі [22] показано, що верхня оцінка $\psi_3^*(G,w)$ є точною для $\alpha(G)$, коли граф G є t-перфектним, або h-перфектним. Недавно виявлені нові властивості для оцінки $\psi_3^*(G,w)$ [39]. Вони пов'язані з тим, що з квадратичних обмежень (19), (20) і (22) випливають правильні лінійні нерівності для p-колеса в графі (відома підструктура в графі, англійський термін для якої p-wheel). Це дає можливість розширити сімейство графів, для яких задача знаходження $\alpha(G,w)$ не є NP-складною, а може бути вирішена за поліноміальний час. Як окремий випадок, сюди попадає і відоме сімейство W-перфектних графів, для якого цей факт доведено на основі методу еліпсоїдів [15].

Техніка лагранжевих оцінок знайшла застосування і в деяких інших задачах оптимізації. Її можна вважати альтернативою використанню методів внутрішніх точок для розв'язання задач напіввизначеного програмування (semidefinite programming). Дійсно, багато задач напіввизначеного програмування доцільно розглядати, як окремий випадок задач недиференційовної оптимізації та застосовувати для розв'язання ефективні методи мінімізації негладких опуклих функцій. Адже умова невід'ємної визначеності деякої симетричної $(n\times n)$-матриці X еквівалентна тому, що мінімальне власне число цієї матриці $\lambda_n(X)\geq 0$. Але $\lambda_n(X)$ – увігнута недиференційовна функція елементів матриці, тобто якщо елементи матриці $X(u)=\{x_{ij}\}_{i,j=1}^n$ є лінійними функціями від вектора невідомих параметрів $u\in R^m$, то умова $X(u)\succeq 0$ еквівалентна опуклому негладкому обмеженню $\varphi(u)=-\lambda_n(X(u))\leq 0$.

Тестові експерименти, пов'язані з розв'язанням задач напіввизначеного програмування, що виникають при знаходженні лагранжевих оцінок у ряді екстремальних задач на графах, – таких як максимальний розріз графа, оптимальна бісекція графа та мінімальна зважена по ребрах розрізу розбивка вершин графа на k частин із заданим числом вершин у кожній частині, можна знайти в [22, 32]. Експерименти проводилися з використанням однієї з версій $r(\alpha)$-алгоритму [30]. Результати показали досить високу швидкість збіжності, стійкість обчислень поблизу границі області визначення функції, що мінімізується. Це означає можливість одержання лагранжевих оцінок з високою точністю за допомогою методів недиференційовної оптимізації.

На закінчення відмітимо, що субградієнтні методи недиференційовної оптимізації, розроблені в Інституті кібернетики під керівництвом академіка Наума Зуселевича Шора, вплинули на розвиток теорії і практики для багатьох напрямків математичного програмування. Вони визнані провідними спеціалістами світового наукового співтовариства в області оптимізації. Результати Наума Зуселевича Шора та його учнів були високо оцінені і в колишньому СРСР, і в Україні. За плідну роботу з розробки чисельних методів недиференційовної оптимізації та їх застовувань Н.З.Шор отримав Державні премії СРСР (1981) і України (1973, 1993, 2000) в області науки і техніки, премію імені В.М. Глушкова Національної академії наук України (1987), премію імені В.С. Михалевича Національної академії наук України (1999).

В даний час розвиток методів недиференційовної оптимізаціїі та їх застосувань активно продовжується. Ще в далекому 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". Не дивлячись на те, що пройшло вже 25 років з того часу, проблема чіткого обгрунтування *r*-алгоритмів залишається актуальною й донині. Як крок в цьому напрямі можна відмітити роботу [40], де для перетворення спеціального еліпсоїда в кулю використовується перетворення простору, близьке до використаного в *r*-алгоритмах. Однак, розтяг простору тут використовується в напрямі різниці двох нормованих субградієнтів. Близьким до напряму різниці двох субградієнтів він буде тільки тоді, коли норми субградієнтів близькі, а коли норми субградієнтів сильно відрізняються, це буде суттєво відрізнятися від розтягу простору в напрямі різниці двох субградієнтів. Подібне перетворення простору використовується також в розроблених М.Г.Журбенком нових модифікаціях ε-субградієнтних алгоритмів з розтягом простору [24]. Дослідження в даних напрямках активно продовжуються ...

Література

1. *Поляк Б.Т.* Введение в оптимизацию. – М.: Наука, 1983. – 384с.

2. *Шор Н.З.* Методы минимизации недиференцируемых функций и их приложения. – К.: Наук. думка, 1979. – 199с.

3. *Шор Н.З.* Применение метода градиентного спуска для решения сетевой транспортной задачи. – В кн.: Материалы научн. семинара по теор. и прикл. вопр. кибернетики и исследования операций: Науч. совет по кибернетике АН УССР. Киев, 1962. – Вып. 1. – С. 9–17.

4. *Шор Н.З., Гамбурд П.Р.* Некоторые вопросы сходимости обобщенного градиєнтного спуска // Кибернетика, 1971. – № 6. – С. 82–84.

5. *Михалевич В.С., Трубин В.А., Шор Н.З.* Оптимизационные задачи производственно-транспортного планирования. Модели, методы, алгоритмы. – М.: Наука, 1986. – 260 с.

6. *Ермольев Ю.М., Шор Н.З.* Метод случайного поиска для задач двухэтапного стохастического программирования и его обобщение // Кибернетика, 1968. – № 1. – С. 90–92.

7. *Еремин И.И.* О методе ``штрафов'' в выпуклом программировании. // Кибернетика, 1966. – № 4. – С. 63–67.

8. *Поляк Б.Т.* Один общий метод решения екстремальных задач // ДАН СССР, 1967. – 174. – № 1. – С. 33–36.

9. *Balinski M.L., Wolfe P. (eds).* Nondifferentiable optimization. Math. Programming Study, 3. – North-Holland, Amsterdam, 1975. – 178 p.

10. *Шор Н.З., Билецкий В.И.* Метод растяжения пространства для ускорения

сходимости в задачах овражного типа. – В кн. Тр. семинара Науч. совета АН УССР по кибернетике "Теория оптимальных решений", Киев, 1969. – № 2. – С. 3–18.

11. *Шор Н.З.* Использование операций растяжения пространства в задачах минимизации выпуклых функций // Кибернетика, 1970. – № 1. – С. 6–12.

12. *Юдин Д.Б., Немировский А.С.* Информационная сложность и эффективные методы решения выпуклых экстремальных задач // Экономика и мат. Методы, 1976. – Вып.2. – № 12. – С. 357–369.

13. *Шор Н.З.* Метод отсечения с растяжением пространства для решения задач выпуклого программирования // Кибернетика, 1977. – № 1. – С. 94–95.

14. *Хачиян Л.Г.* Полиномиальные алгоритмы в линейном программировании // ДАН СССР, 1979. – 244. – № 5. – С. 1093–1096.

15. *Grotschel M., Lovasz L., Schrijver A.* Geometric Algorithms and Combinatorial Optimization. – Berlin: Springer-Verlag, 1988. – 362 p.

16. *Bachem A., Grotschel M., Korte B. (eds).* Mathematical Programming: The State of Art, Bonn, 1982. – Springer-Verlag. 1983. – 655 p.

17. *Шор Н.З., Журбенко Н.Г.* Метод минимизации, использующий операцию растяжения пространства в направлении разности двух последовательных градиентов // Кибернетика, 1971. – № 3. – С. 51–59.

18. *Журбенко Н.Г.* Построение семейства методов сопряженных направлений на основе использования оператора растяжения пространства // Теория и приложения методов оптимизации. – К.: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 1998. – С. 12–18.

19. *Журбенко Н.Г.* Квазиньютоновские алгоритмы минимизации на основе использования оператора растяжения пространства // Теория оптимальных решений. – К.: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 1999. – С. 45–50.

20. *Шор Н.З.* Монотонные модификации r-алгоритмов и их приложения // Кибернетика и системный анализ, 2002. – № 6. – С. 74–95.

21. *Шор Н.З., Стеценко С.И.* Квадратичные экстремальные задачи и недиференцируемая оптимизация. – К.: Наук. думка, 1989. – 208 с.

22. *Shor N.Z.* Nondifferentiable Optimization and Polynomial Problems. – Boston; Dordrecht; London: Kluwer Academic Publishers, 1998. – 412 p.

23. *Шор Н.З., Соломон Д.И.* Декомпозиционные методы в дробно-линейном программировании. – Кишинев: Штиинца, – 1989. – 204 с.

24. *Шор Н.З., Сергієнко І.В., Шило В.П., та ін.* Задачі оптимального проектування надійних мереж. – К.: Наук. думка, 2005. – 230 с.

25. *Михалевич В.С., Сергиенко И.В., Шор Н.З.* Исследование методов решения оптимизационных задач и их приложения // Кибернетика, 1981. – № 4. – С. 89–113.

26. *Михалевич В.С., Сергиенко И.В., Трубин В.А., Шор Н.З. и др.* Пакет прикладных программ для решения задач производственно-транспортного планирования большой размерности (ПЛАНЕР) // Кибернетика, 1983. – № 3. – С. 57–71.

27. *Михалевич В.С., Сергиенко И.В., Шор Н.З. и др.* Пакет программ ДИСПРО-3: назначение, классы решаемых задач, системное и алгоритмическое обеспечение // Кибернетика, 1985. – № 1. – С. 56–71.

28. *Лаптин Ю.П., Журбенко Н.Г., Левин М.М., Волковицкая П.И.* Использование средств оптимизации в системе автоматизированного

проектирования энергетических котлоагрегатов КРОКУС // Энергетика и электрофикация, 2003. – № 7. – С. 41–51.

29. *Киселева Е.М., Шор Н.З*. Непрерывные задачи оптимального разбиения множеств: теория, алгоритмы, приложения. – К.: Наук. думка, 2005. – 564 с.

30. *Шор Н.З., Стецюк П.И.* Использование модификации r-алгоритма для нахождения глобального минимума полиномиальных функций // Кибернетика и системный анализ, 1997. – № 4. – С. 28–49.

31. *Шор Н.З.* Роль избыточных ограничений в улучшении двойственных оценок для полиномиальных оптимизационных задач // Кибернетика и системный анализ, 1998. – № 4. – С. 106–121.

32. *Shor N.Z., Stetsyuk P.I.* Lagrangian bounds in multiextremal polynomial and discrete optimization problems // J. of Global Optimization, 2002. – N 23. – P. 1–41.

33. *Стецюк П.И.* О функционально избыточных ограничениях для булевых оптимизационных задач квадратичного типа // Кибернетика и системный анализ, 2005. – № 6. – С. 168–172.

34. *Стецюк П.И., Пардалос П.М.* Об уточнении лагранжевых двойственных оценок в бинарных и булевых квадратичных задачах // Теория оптимальных решений. – К.: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 2006. – С. 145–153.

35. *Стецюк П.И.* Новые модели квадратичного типа для задачи о максимальном взвешенном разрезе графа // Кибернетика и системный анализ, 2006. – № 1. – С. 63–75.

36. *Шор Н.З.* Об одном подходе к получению глобальных экстремумов в полиномиальных задачах математического программирования // Кибернетика, 1987. – № 5. – С. 102–106.

37. *Шор Н.З.* Об одном классе оценок глобального минимума полиномиальных функций // Кибернетика, 1987. – № 6. – С. 9–11.

38. *Гэри М., Джонсон Д.* Вычислительные машины и труднорешаемые задачи. – М.: Мир, 1982. – 416 с.

39. *Стецюк П.И., Чумаков Б.М.* О свойствах одной верхней оценки Н.З.Шора для взвешенного числа устойчивости графа // Праці міжнародного симпозіума "Питання оптимізації обчислень (ПОО-XXXIII)". – К.: Ін-т кібернетики ім. В.М.Глушкова НАН України, 2007. – С. 271–271.

40. *Стецюк П.И.* r-алгоритмы и эллипсоиды // Кибернетика и систем. анализ, 1996. – № 1. – С. 113–134.

Підготував: Стецюк Петро Іванович, канд. фіз.-мат. наук, виконувач обов'язків завідувача відділу Інституту кібернетики ім. В.М. Глушкова НАН України,

тел: (044)-526-21-68, e-mail: stetsyuk@d120.icyb.kiev.ua