Національна академія наук України
Інститут кібернетики імені В. М. Глушкова

На правах рукопису

СТЕЦЮК Петро Іванович

УДК 519.8

СУБГРАДІЄНТНІ МЕТОДИ З ПЕРЕТВОРЕННЯМ ПРОСТОРУ ДЛЯ МІНІМІЗАЦІЇ НЕГЛАДКИХ ОПУКЛИХ ФУНКЦІЙ

01.05.01 — теоретичні основи інформатики та кібернетики
(математична кібернетика)

Автореферат дисертації на здобуття наукового ступеня
кандидата фізико-математичних наук

Київ 1996

Дисертацією є рукопис.

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

Науковий керівник: член-кореспондент НАН України,
доктор фізико-математичних наук,
професор ШОР Н. З.

Офіційні опоненти: доктор фізико-математичних наук
ГУПАЛ А. М.,

доктор фізико-математичних наук,
професор ДАНІЛІН Ю. М.

Провідна організація: Київський національний університет
імені Тараса Шевченка.

Захист відбудеться «25» жовтня 1996 р. о 11 год. на засіданні спеціалізованої вченої ради Д 01.39.02 при Інституті кібернетики імені В. М. Глушкова НАН України за адресою:

252022 Київ 22, проспект Академіка Глушкова, 40.

З дисертацією можна ознайомитися в науково-технічному архіві інституту.

Автореферат розісланий «18» вересня 1996 р.

Учений секретар
спеціалізованої вченої ради СИНЯВСЬКИЙ В. Ф.

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ.

Актуальність теми. Проблеми мінімізації негладких опуклих функцій займають центральне місце серед алгоритмічних проблем оптимізації. Реалізація схем декомпозиції при розв'язуванні структурованих задач оптимізації великої розмірності, знаходження двоїстих оцінок в багатоекстремальних і дискретних оптимізаційних задачах, аналіз стійкості динамічних систем, пошук раціональних стратегій за умови невизначеності або конфлікту – все це призводить до необхідності мінімізації негладких опуклих функцій. Широта та важливість застосування стимулювали дослідження в цьому напрямі у всьому світі.

Одними з найбільш ефективних методів для мінімізації негладких опуклих функцій є запропоновані Н.З.Шором r-алгоритми (клас субградієнтних методів з розтягом простору в напрямі різниці двох послідовних субградієнтів). Багаторічна практика використання $r(\alpha)$-алгоритму підтверджує ефективність цього класу методів для задач з різноманітною структурою поверхностей рівня $f(x)$, в тому числі і дуже овражною. Результати роботи $r(\alpha)$-алгоритму не поступаються кращим із методів в швидкості збіжності та субоптимальним методам першого порядку – за числом обчислень $f(x)$ і $\partial f(x)$. При цьому його відрізняє невелика кількість допоміжних арифметичних операцій на ітерації. Однак обгрунтування збіжності r-алгоритмів проведено недосить повно, і результати про збіжність конкретних методів цього класу не поширюються на загальний випадок задач мінімізації негладких опуклих функцій.

У звязку з цим актуальним є дослідження можливостей побудови ефективних субградієнтних методів з перетворенням простору (субградієнтних методів змінної метрики), для яких можна було б обгрунтувати збіжність при мінімізації довільної функції із класу опуклих функцій. В. являється це можна зробити і, більш того, якщо взяти за основу ідею, висловлену Н.З.Шором в 1969 р., яка полягає у використанні лінійних неортогональних перетворень простору для поліпшення умов субградієнтного процесу в перетвореному просторі аргументів, то можна побудувати ряд сімейств практично ефективних методів для мінімізації негладких опуклих функцій. Для того, щоб звузити широкий клас субградієнтних методів змінної метрики, розглянемо тільки такі методи, які використовують монотонне зменшення об'єму локалізаторів множини екстремумів. На користь цієї характеристики свідчить той факт, що на ній засновані багато "яскравих" результатів в опуклому програмуванні.

Мета роботи. Аналіз причин, що утрудняють обгрунтування r-алгоритмів при мінімізації негладких опуклих функцій. Дослідження можливостей побудови ефективних методів субградієнтного типу з неретворенням простору, розробка конкретних варіантів таких методів з обгрунтуванням їх збіжності для всього класу опуклих функцій.

Наукова новизна роботи. Побудований приклад функції, який пояснює перешкоди при обгрунтуванні r-алгоритмів для негладких опуклих функцій. Знайдена форма апроксимуючого множину экстремумів еліпсоїда, за допомогою якого можна обгрунтовувати збіжність близьких до r-алгоритмів методів з розтягом простору. Побудована конструкція методу центрів тяжіння простих тіл з монотонним зменшенням їх об'єму, яка дозволяє конструювати методи змінної метрики з обгрунтуванням їх збіжності для задачі опуклого програмування. Досліджені два простих однорангових оператори перетворення простору, за допомогою яких можна обгрунтовувати методи змінної метрики для мінімізації опуклих функцій, використовуючи зовнішню апроксимацію множини екстремумів еліпсоїдами спеціального типу. На їх основі побудовані і обгрунтовані ряд субградієнтних методів змінної метрики. Показана ефективність цих методів при мінімізації овражних негладких опуклих функцій.

Практична цінність роботи полягає в тому, що запропоновані методи програмно реалізовані і можуть бути використані при розв'язуванні задач опуклого програмування, зокрема для знаходження допустимої точки системи опуклих нерівностей.

Апробація роботи. Результати роботи неодноразово обговорювались на семінарах відділу "Методи розв'язування складних задач оптимізації" Інституту кібернетики ім. В.М.Глушкова НАН України та на семінарах "Теорія оптимальних рішень" Наукової ради НАН України по проблемі "Кібернетика".

Структура та обсяг роботи. Дисертація складається з вступу, трьох розділів, висновків та списку літератури. Обсяг дисертації – 135 с., в тому числі основний текст – 129 с., бібліографія – 52 найменування.

Публікації. За темою дисертації автор опублікував 5 робіт.

ЗМІСТ РОБОТИ

У **вступі** обгрунтовується актуальність теми дисертації, показаний зв'язок методів з перетворенням простору аргументів (методів змінної метрики) з рядом результатів в області математичного програмування, наведений конспективний огляд головних результатів дисертації. Алгоритмічно методи з перетворенням простору змінних можна розглядати в двох формах: B-форма (коректується матриця B, що задає обернений оператор до оператору перетворення простору аргументів) та H-форма (коректується симетрична матриця $H = BB^T$). Вибрана орієнтація на методи в B-формі враховуючи, що вони дозволяють дати просту інтерпретацію процесу в перетвореному просторі аргументів.

Додатково до загальновідомих позначень опуклого аналізу в робс ' використані такі: R^n – евклідовий простір розмірністю n з скалярним добутком $(x, y), x \in R^n, y \in R^n$; $f(x)$ – опукла функція векторного аргументу $x \in X$; $\partial f(x)$ – субградієнт $f(x)$; X – прямий простір аргументів (змінних); $Y = AX$ – перетворений лінійним оператором простір аргументів (змінних). Лінійний оператор A із R^n в R^n задається матрицею A розміром $n \times n$. Обернений до A оператор задається матрицею $B = A^{-1}$. A^T, B^T – транспонування матриць, x, y – елементи (точки) із X і Y відповідно. Для позначень напрямів в X і Y частіше всього використовуються p, g, ξ, η.

У розділі 1 аналізуються як проблеми, що виникають при обгрунтуванні збіжності r-алгоритмів для мінімізації негладких опуклих функцій, так і можливі засоби подолання цих проблем. У §1.1 наведен й приклад опуклої кусочно-лінійної функції, при мінімізації якої можливе "зациклювання" $r_\mu(\alpha)$-алгоритму (ідеалізований варіант r-алгоритмів) в неоптимальній точці. Він полягає в тому, що для негладких функцій точки з лінійно-залежною множиною субградієнтів можуть служити "ловушками" для мінімізуючої послідовності $r_\mu(\alpha)$-алгоритму, і показує, що обгрунтувати збіжність $r_\mu(\alpha)$-алгоритму для всього класу опуклих функцій неможливо. Для того щоб це зробити, потрібно модифікувати $r_\mu(\alpha)$-алгоритм, і деякі з таких модифікацій обговорюються. Їх зміст зводиться до того, щоб напрям субградієнта замінити на найкоротший вектор до опуклої оболонки накопичених в точці субградієнтів, що дозволяє відрізнити точку мінімуму від "ловушки". Це призводить до незначної модифікації $r_\mu(\alpha)$-алгоритму із збереженням його основних властивостей. Однак в обчислювальному плані такі модифікації можна вважати ідеалізованими, так як для опуклих функцій вони вимагають досить точної процедури спуску в заданому напрямі.

У §1.2 проаналізовані причини, які перешкоджають обгрунтуванню збіжності r-алгоритмів ($r_\mu(\alpha)$-алгоритм, граничний варіант r-алгоритмів та $r(\alpha)$-алгоритм) для всього класу опуклих функцій. Основна причина полягає в тому, що r-алгоритми не мають досить конструктивного критерію зупинки для негладких опуклих функцій. Підтвердженням цьому є матеріал §1.1. Якщо субградієнтні методи з розтягом простору доповнити критерієм зупинки, який був би однаково хорошим як для гладких, так і для негладких опуклих функцій, то багатьох проблем можна уникнути. В зв'язку з цим розглянутий підхід заснований н зовнішній апроксимації множини екстремумів еліпсоїдом з монотонним зменшенням його об'єму. Це призводить до дуже простої геометричної інтерпретації оператора розтягу простору $R_\alpha(\xi)$, зв'язаної з перетворенням еліпсоїду спеціального типу $ell(y^k, \xi_k, a, b)$ в $Y_k = R^n$ в кулю. Тут точка

y^k задає центр еліпсоїда в Y_k, ξ_k – напрям в Y_k, в якому довжина півосі еліпсоїда рівна a, b – довжини його півосей в $(n-1)$ ортогональних до ξ_k напрямах. Щоб перетворити $ell(y^k, \xi_k, a, b)$ в кулю радіусом b досить застосувати операцію розтягу простору в напрямі ξ_k з коефіцієнтом $\alpha_k = \frac{b}{a}$. Якщо $b > a$, це буде розтягом простору аргументів, якщо $b < a$, – його стисненням. При цьому до простору субградієнтів буде застосований оператор $R_{\beta_k}(\xi_k)$ і відповідно в першому випадку одержимо його стиснення, в другому – його розтяг.

Якщо такі еліпсоїди використовувати для локалізації множини екстремумів так, щоб гарантувати зменшення їх об'єму від ітерації до ітерації, тоді для задач опуклого програмування в рамках методів відсічень з розтягом простору автоматично одержимо метод, збіжний зі швидкістю геометричної прогресії по послідовності рекордів функції, яка мінімізується. Певний спосіб побудови цих еліпсоїдів буде визначати конкретний алгоритм. Для методів з розтягом простору така інтерпретація є досить наглядною і дозволяє зв'язати для них кроковий множник у напрямі руху і коефіцієнт розтягу простору з такими характеристиками опуклих функцій, як геометрія ліній рівня $f(x)$ і геометрія субдиференціала в екстремумі.

У цьому ж параграфі розглянуто ще один апроксимуючий множину екстремумів еліпсоїд $\mathrm{ell}(x_0, a, b, c)$, який використовує перетворення простору, близьке до того, що має місце в r-алгоритмах. Ним є еліпсоїд спеціального типу, що містить опукле тіло в R^n, одержане внаслідок перерізу кулі $S(x_0, R) = \{x : \|x - x_0\| \leq R\}$ і двох півпросторів $P(x_0, g_1) = \{x : (x - x_0, g_1) \leq 0\}$ і $P(x_0, g_2) = \{x : (x - x_0, g_2) \leq 0\}$, таких, що $(g_1, g_2) < 0$. Об'єм еліпсоїда $\mathrm{ell}(x_0, a, b, c)$ менший, ніж об'єм кулі $S(x_0, R)$ в $\sqrt{1 - (\frac{g_1}{\|g_1\|}, \frac{g_2}{\|g_2\|})^2}$ раз.

Перетворення простору аргументів, що переводить еліпсоїд $\mathrm{ell}(x_0, a, b, c)$ в кулю, задається за допомогою послідовного використання двох операторів розтягу простору:

$$T(\xi_1, \xi_2) = R_{\alpha_1}\left(\frac{\xi_1+\xi_2}{\|\xi_1+\xi_2\|}\right) R_{\alpha_2}\left(\frac{\xi_1-\xi_2}{\|\xi_1-\xi_2\|}\right), \tag{1}$$

де $\|\xi_1\| = \|\xi_2\| = 1$; $(\xi_1, \xi_2)^2 < 1$; $\alpha_1 = \frac{1}{\sqrt{1-(\xi_1,\xi_2)}}$; $\alpha_2 = \frac{1}{\sqrt{1+(\xi_1,\xi_2)}}$; $R_\alpha(\xi) = I + (\alpha - 1)\xi\xi^T$ – оператор розтягу простору в напрямі ξ ($\|\xi\| = 1$) з коефіцієнтом α.

Обернений до $T(\xi_1, \xi_2)$ оператор (для перетворення простору субградієнтів) має вигляд

$$T^{-1}(\xi_1, \xi_2) = R_{\beta_1}\left(\frac{\xi_1+\xi_2}{\|\xi_1+\xi_2\|}\right) R_{\beta_2}\left(\frac{\xi_1-\xi_2}{\|\xi_1-\xi_2\|}\right), \tag{2}$$

де $\beta_1 = \frac{1}{\alpha_1} = \sqrt{1 - (\xi_1, \xi_2)}$; $\beta_2 = \frac{1}{\alpha_2} = \sqrt{1 + (\xi_1, \xi_2)}$.

Для оператора (1)–(2) справедлива наступна лема (тут і далі нумерація тверджень, винесених в автореферат, співпадає з нумерацією в тексті дисертації).

Лема 1. Нехай B_k – невироджена матриця розміром $n \times n$, p_1, p_2 – n-мірні вектори, такі, що $\left(\frac{B_k^T p_1}{\|B_k^T p_1\|}, \frac{B_k^T p_2}{\|B_k^T p_2\|}\right)^2 < 1$, і нехай

$$B_{k+1} = B_k R_{\beta_1}\left(\frac{\xi_1+\xi_2}{\|\xi_1+\xi_2\|}\right) R_{\beta_2}\left(\frac{\xi_1-\xi_2}{\|\xi_1-\xi_2\|}\right), \qquad (3)$$

де $\xi_1 = \frac{B_k^T p_1}{\|B_k^T p_1\|}$; $\xi_2 = \frac{B_k^T p_2}{\|B_k^T p_2\|}$; $\beta_1 = \sqrt{1-(\xi_1,\xi_2)}$; $\beta_2 = \sqrt{1+(\xi_1,\xi_2)}$. Тоді матриця B_{k+1} – невироджена і, крім того, виконуються умови. а) $\det(B_{k+1}) = \det(B_k)\sqrt{1-(\xi_1,\xi_2)^2}$; б) $(B_{k+1}^T p_1, B_{k+1}^T p_2) = 0$. □

Лема 1 має просту геометричну інтерпретацію. Умова а) відображає зменшення об'єму еліпсоїда типу $\mathrm{ell}(x_0, a, b, c)$ по відношенню до об'єму кулі $S(x_0, c)$, який у випадку, коли $(B_k^T p_1, B_k^T p_2) < 0$, можна використовувати для зовнішньої апроксимації множини екстремумів. Пояснимо геометричний зміст умови б). Нехай $g_f(x)$ – субградієнт $f(x)$. Тоді $g_{\varphi_k}(y) = B_k^T g_f(x)$ – субградієнт функції $\varphi_k(y) = f(B_k y)$ в перетвореному за допомогою лінійного оператора просторі $Y = A_k X$, де $A_k = B_k^{-1}$. Тому, якщо p_1 і p_2 – субградієнти $f(x)$ в точці x_k, то перетворення простору $Y = A_{k+1}X$, $A_{k+1} = B_{k+1}^{-1}$ (B_{k+1} обчислена згідно з (3)), ортогоналізує їх прообрази $B_{k+1}^T p_1$ і $B_{k+1}^T p_2$ для функції $\varphi_{k+1}(y) = f(B_{k+1}y)$. Умова $\left(\frac{B_k^T p_1}{\|B_k^T p_1\|}, \frac{B_k^T p_2}{\|B_k^T p_2\|}\right)^2 < 1$ потрібна для того, щоб ортогоналізація була можливою.

Використання оператора (1) для перетворення простору аргументів вимагає в два рази більше арифметичних операцій, ніж розтяг простору в напрямі різниці двох послідовних субградієнтів. Однак по відношенню до останнього він має ряд переваг: по-перше, враховується кут між векторами, що забезпечує змінний коефіцієнт розтягу простору; по-друге, якщо кут між векторами тупий, то можна гарантувати зменшення об'єму еліпсоїда, який локалізує множину екстремумів, що для методів змінної метрики на основі оператора розтягу простору дозволяє обгрунтовувати збіжність по послідовності рекордів функції, яка мінімізується.

У §1.3 на основі оператора (1) побудовані та обгрунтовані два субградієнтних методи з розтягом простору для знаходження точки мінімуму опуклої функції за відомим оптимальним значенням функції. Для них характерно використання незначної інформації про відсікаючі множину екстремумів гіперплощини: перший використовує не більше двох, другий – не більше трьох гіперплощин. Перший з методів побудований по типу r-алгоритмів, і для перетворення простору використовує два послідовних субградієнти, другий з яких одержаний згідно з класичним фейєровським кроком у напрямі першого антисубградієнта. Його можна рахувати певним варіантом r-алгоритмів із змінним коефіцієнтом розтягу простору та класичним фейєровським регулюванням крокового множника.

Другий метод використовує два послідовних субградієнти та вектор агрегатного типу, який є опуклою комбінацією обчислених раніше субградієнтів. Він також, як і перший, використовує класичний фейєровський крок в напрямі останнього обчисленого субградієнта, однак для вибору напрямів розтягу використовує останній субградієнт та вектор агрегатного типу. Використання останнього обгрунтовано наступною лемою:

Лема 3. Нехай p_1, p_2, p_3 – вектори в R^n, такі, що $\|p_1\| = \|p_2\| = \|p_3\| = 1$, $(p_1, p_2) = 0$, і умови $(p_1, p_3) \geq 0$ та $(p_2, p_3) \geq 0$ не виконуються одночасно. Тоді серед опуклих комбінацій векторів p_1 і p_2 вектор

$$p = \begin{cases} p_1, & \text{якщо } (p_1, p_3) < 0 \text{ і } (p_2, p_3) \geq 0, \\ p_2, & \text{якщо } (p_1, p_3) \geq 0 \text{ і } (p_2, p_3) < 0, \\ \lambda_1 p_1 + \lambda_2 p_2, & \text{якщо } (p_1, p_3) < 0 \text{ і } (p_2, p_3) < 0 \end{cases}$$

доставляє мінімум скалярному добутку (p, p_3). Тут $\lambda_1 = -\frac{(p_1,p_3)}{\sqrt{(p_1,p_3)^2+(p_2,p_3)^2}}$, $\lambda_2 = -\frac{(p_2,p_3)}{\sqrt{(p_1,p_3)^2+(p_2,p_3)^2}}$. $\square$

При використанні агрегатного вектора автоматично вирішується проблема його "поновлення", що є непростою проблемою в методах агрегатного типу.

Обгрунтування збіжності розглянутих методів підкріплено відповідними теоремами, які засновані на монотонному зменшенні об'єму еліпсоїда для локалізації множини екстремумів. Проведені тестові експерименти, що показують ефективність цих методів при мінімізації овражних опуклих гладких і негладких функцій. Аналіз роботи методів проведений в порівнянні з результатами роботи $r(\alpha)$-алгоритму. Другий з методів не поступається $r(\alpha)$-алгоритму, а в ряді випадків його переважає. Обговорюються можливі способи поліпшення методів за умови використання більш значної кількості відсікаючих множину екстремумів гіперплощин.

У розділі 2 обговорюється можливість побудови для задачі опуклого програмування методів першого порядку, для яких можна було б розумно оцінити складність обробки інформації (ОІ-трудомісткість) - обсяг допоміжної роботи (в арифметичних операціях), необхідний при реалізації методу. В рамках методів відсічень з використанням лінійних перетворень простору розглянута загальна конструкція – метод центрів тяжіння простих тіл (МЦТПТ), котра для задач опуклого програмування дозволяє обгрунтовувати алгоритми, збіжні з швидкістю геометричної прогресії за послідовністю рекордів функції, яка мінімізується. Вона заснована на зовнішній апроксимації множини екстремумів простими опуклими тілами (еліпсоїд, паралелепіпед, симплекс) з достатнім зменшенням їх об'єму і регуляризацією цих тіл (перетворення еліпсоїда в кулю, паралелепіпеда – в куб, неправильного симплекса – в правильний) за допомогою лінійних перетворень простору.

Для задачі опуклого програмування

$$f_0(x) \to \min,$$

$$f_i(x) \le 0, \quad i = 1, \dots, l$$

($f_i(x), i = 0, \dots, l$ – опуклі функції векторного аргументу $x \in X$, $X = R^n$) МЦТПТ генерує ітераційну послідовність $\{x_k\}_{k=0}^{\infty}$ таким чином:

перед початком обчислень маємо: x_0 – центр тяжіння L_0 ($X^* \subset L_0$, X^* – множина екстремумів, L_0 – простий локалізатор в R^n одного з типів – куля, куб або правильний симплекс); $q = 1 - \frac{1}{e}$ – константа, яка характеризує необхідне зменшення об'єму на ітерації МЦТПТ; m_0 – максимальне число накопичених субградієнтів ($m_0 \le 2n$); $B_0 = I_n$ – одинична матриця розміром $n * n$.

Крок 1. Обчислимо згідно з певними правилами (внутрішній алгоритм) ($x_i, f(x_i), \partial f(x_i), \partial \varphi_k(y_i) = B_k^T \partial f(x_i), i = 1, 2, \dots, m), m \le m_0$. Тут $f(x)$ одна із функцій $f_i(x), i = 0, \dots, l$. Якщо одержані достатні умови екстремуму, то кінець роботи алгоритма.

Крок 2. У просторі аргументів $Y_k = B_k^{-1} X$ відносно центру тяжіння (точки $y_k = B_k^{-1} x_k$) локалізатора L_k (одне із тіл – куля, куб або правильний симплекс) вибираємо черговий локалізатор L'_{k+1} (одне із тіл – еліпсоїд, паралелепіпед або неправильний симплекс), такий, щоб $Y_k^* = B_k^{-1} X^* \subset L'_{k+1}$ і $q^* = \text{vol}(L'_{k+1})/\text{vol}(L_k)$ було мінімальним. Якщо $q^* \ge q$, то переходимо до кроку 1.

Крок 3. Обчислимо чергове наближення

$$x_{k+1} = x_k + \Gamma \sum_{i=1}^{m} \alpha_i \partial \varphi_k(y_i),$$

яке задає образ в X центру тяжіння L'_{k+1} в Y_k. Тут скаляри $\alpha_i, i = 1, 2, \dots, m$ задають в Y_k центр тяжіння L'_{k+1} відносно центру тяжіння L_k.

Крок 4. $B_{k+1} = B_k T_{k+1}^{-1}$ (регуляризація L'_{k+1}). T_{k+1} – невироджена матриця розміром $n * n$ для перетворення L'_{k+1} в L_{k+1} (еліпсоїда – в кулю, паралелепіпеда – в куб, неправильного симплекса – в правильний), така, що $\det(T_{k+1}^{-1}) = q^*$.

Крок 5. $\partial \varphi_{k+1}(y_i) = (T_{k+1}^{-1})^T \partial \varphi_k(y_i), i = 1, 2, \dots, m$ (перерахунок субградієнтів при переході в черговий перетворений простір аргументів).

Крок 6. Приймемо $k = k + 1$ і перейдем до кроку 2.

ОІ-трудомісткість ітерації МЦТПТ визначається ОІ-трудомісткістю кроків 1 – 6, де крок 1 визначає внутрішній алгоритм, кроки 2 – 6 – зовнішній. Крок 3 вимагає $O(n^2)$ арифметичних операцій, кроки 4 і 5 – $O(n^3)$ в самому найгіршому випадку (при використанні еліпсоїдів, паралелепіпедів або симплексів самого загального виду). Тому, якщо забезпечити легкий вибір зовнішнім алгоритмом чергового локалізатора (крок 2) і для внутрішнього алгоритму ОІ-трудомісткість порядку $O(n^3)$, одержимо загальну ОІ-трудомісткість ітерації МЦТПТ не більше $O(n^3)$. Якщо при цьому для МЦТПТ на кожній ітерації зможемо гарантувати зменшення об'єму в $(1 - \frac{1}{e})$ раз, то він буде реалізовувати теоретично оптимальну оцінку за послідовністю рекордів функції, що мінімізується, і його загальна ОІ-трудомісткість буде – $O(n^4)$ арифметичних операцій. Використовуючи спеціальні локалізатори одного із зазначених типів і враховуючи, що робота внутрішнього алгоритму потрібна не на кожній ітерації, для МЦТПТ можна забезпечити середню ОІ-трудомісткість ітерації – $O(n^2)$ арифметичних операцій.

Загальна схема МЦТПТ дозволяє конкретизуючи внутрішній і зовнішній алгоритми одержувати методи для розв'язування задачі опуклого програмування збіжні за послідовністю рекордів функції, яка мінімізується. Так, якщо орієнтуватися на вибір локалізатора одного й того ж типу, то одержимо метод описаних еліпсоїдів (МОЕ), метод описаних паралелепіпедів (МОП) або ж метод описаних симплексів (МОС). Методи в рамках схеми МЦТПТ можуть застосовуватися для більш широкого класу задач, ніж задачі опуклого програмування, наприклад, таких як задача знаходження сідлових точок опукло-ввігнутих функцій, окремі випадки задач розв'язування варіаційних нерівностей, а також спеціальні класи задач лінійної і нелінійної додатковості.

У §2.2 обговорюються можливі схеми побудови зовнішнього і внутрішнього алгоритмів. Для зовнішнього алгоритму головна увага приділена використанню простих локалізаторів типу еліпсоїдів, що зумовлено мінімумом інформації про локалізатори і тим, щоб забезпечити ОІ-трудомісткість ітерації – $O(n^2)$ арифметичних операцій. Для внутрішнього алгоритму увага приділена побудові нижньої оцінки та евристичним способам побудови ефективних відсікаючих гіперплощин з наближеним виконанням достатніх умов оптимальності по типу ε-субградієнтних методів – $0 \in \partial_\varepsilon f(x)$, де ε розраховується на основі одержаної в процесі обчислень інформації.

§2.3 присвячений ряду проблем для методів опуклого програмування, котрі просто розв'язуються в рамках МЦТПТ. Це такі: "відсів" лишніх відсікаючих гіперплощин; перевірка несумісності системи обмежень задачі опуклого програмування; проблема вибору крокового множника в напрямі руху для субградієнтних методів; простота розпаралелювання процесу на рівні змістовних підзадач; проблема "рестарта" для методів з лінійними перетвореннями простору та ін.

Найбільш змістовним з точки зору практичних методів є розділ 3. Тут досліджені два простих однорангових оператори перетворення простору, за допомогою яких можна обгрунтовувати методи змінної метрики, використовуючи зовнішню апроксимацію множини екстремумів еліпсоїдами спеціального типу з монотонним зменшенням їх об'єму. Особливість цих операторів (використовують ортогональні напрями) дає можливість забезпечити просту та наглядну геометричну інтерпретацію процесів в перетвореному просторі змінних. Для знаходження точки мінімуму опуклої функції за відомим оптимальним значенням функції на основі цих операторів побудовано ряд субградієнтних методів змінної метрики з використанням класичного фейєровського кроку в напрямі антисубградієнта: два методи по типу r-алгоритмів – за допомогою однорангового еліпсоїдального оператора; сімейство методів ортогонального субградієнтного спуску в перетвореному просторі аргументів – за допомогою однорангового доортогоналізуючого оператора. Перетворення простору в цих методах направлене на зменшення овражності функцій, що забезпечує їх ефективність при мінімізації негладких опуклих функцій.

У §3.1 проаналізовані існуючі методи фейєровського типу для задачі

$$\min f(x), \tag{4}$$

де $f(x)$ – опукла функція векторного аргументу $x \in X$, $X = R^n$. Припускається, що множина екстремумів задачі (4) – X^* непуста і відоме значення мінімуму $f(x)$: $f^* = f(x^*)$, $x^* \in X^*$. Зазначений ряд проблем для цих методів. Так самий простий з таких методів використовує класичний фейєровський крок в напрямі антисубградієнта і за кінцеве число ітерацій дозволяє знайти ε_f-розв'язок задачі (4), а саме точку $x_k \in X^*_{\varepsilon_f}$, для якої виконується $f(x_k) - f^* \le \varepsilon_f$. Однак він неефективний на практиці і підвищити його ефективність можна за допомогою лінійних перетворень простору. Розглянуті основні співвідношення, необхідні для реалізації фейєровських процесів змінної метрики.

У §3.2 розглянуто одноранговий еліпсоїдальний оператор перетворення простору аргументів (лінійний оператор із R^n в R^n), який в матричній формі має вигляд

$$T_1(\xi,\eta) = I - \tfrac{1}{1-(\xi,\eta)^2}\left(\left(1-\sqrt{1-(\xi,\eta)^2}\right)\eta - (\xi,\eta)\xi\right)\eta^T . \tag{5}$$

Тут I – одинична матриця розміром $n * n$; $\xi, \eta \in R^n$ – вектори, такі, що $\|\xi\| = 1$, $\|\eta\| = 1$ і їх скалярний добуток задовольняє умові $(\xi,\eta)^2 \ne 1$. Обернений до (5) оператор $T_1^{-1}(\xi,\eta)$ (для перетворення простору субградієнтів) має вигляд

$$T_1^{-1}(\xi,\eta) = I + \frac{1}{\sqrt{1-(\xi,\eta)^2}}\left(\left(1-\sqrt{1-(\xi,\eta)^2}\right)\eta - (\xi,\eta)\xi\right)\eta^T. \qquad (6)$$

Для одорангового еліпсоїдального оператора справедлива лема.

Лема 5. Нехай B_k – невироджена матриця розміром $n*n$, p_1, p_2 – n-мірні вектори, такі, що $\left(\frac{B_k^T p_1}{\|B_k^T p_1\|}, \frac{B_k^T p_2}{\|B_k^T p_2\|}\right)^2 < 1$. І нехай $B_{k+1} = B_k T_1^{-1}(\xi,\eta)$, де $\xi = \frac{B_k^T p_1}{\|B_k^T p_1\|}$ та $\eta = \frac{B_k^T p_2}{\|B_k^T p_2\|}$.

Тоді матриця B_{k+1} – невироджена і, крім того, для неї виконані наступні умови: а)$\det(B_{k+1}) = \det(B_k)\sqrt{1-(\xi,\eta)^2}$, б)$(B_{k+1}^T p_1, B_{k+1}^T p_2) = 0$. □

Використовуючи зменшення об'єму еліпсоїда, який локалізує множину екстремумів, однранговий еліпсоїдальний оператор (5) дозволяє обгрунтовувати методи для розв'язування задачі (4) по тому ж принципу, що й оператор (1). Однак для B-форми методів оператор (5) більш раціональний, так як його використання вимагає в два рази менше арифметичних операцій. Для знаходження ε_f-розв'язку задачі (4) аналогічно методам §1.3 побудовані і обгрунтовані два фейєровських методи змінної метрики: метод, який використовує два послідовних субградієнти, та метод, який використовує два послідовних субградієнти і вектор агрегатного типу. При цьому проблема поповлення агрегатного вектора розв'язується автоматично, аналізуючи кути між двома останніми субградієнтами і попереднім агрегатним вектором.

Для знаходження ε_f-розв'язку задачі (4) метод, який використовує два послідовних субградієнти і вектор агрегатного типу, має вигляд:

перед початком процесу маємо $\varepsilon_f > 0$, $x_0 \in R^n$. Тоді, якщо $f(x_0) - f^* \leq \varepsilon_f$, то x_0 – ε_f-розв'язок і кінець роботи алгоритму. Інакше приймем $h_0 = \frac{f(x_0)-f^*}{\|\partial f(x_0)\|}$, $\xi_0 = \frac{\partial f(x_0)}{\|\partial f(x_0)\|} \in R^n$, $p_0 = 0 \in R^n$, $B_0 = I$ – одинична матриця розміром $n*n$.

Нехай на k-й ітерації олержані $x_k \in R^n$, h_k, $\xi_k \in R^n$, $p_k \in R^n$, B_k – матриця $n*n$.

1.Обчислимо чергове наближення

$$x_{k+1} = x_k - h_k B_k \xi_k. \qquad (7)$$

2.Обчислимо $f(x_{k+1})$, $\partial f(x_{k+1})$. Тоді, якщо $f(x_{k+1}) - f^* \leq \varepsilon_f$, то x_{k+1} – ε_f-розв'язок і кінець роботи алгоритму. В протилежному випадку приймем

$$\xi_{k+1} = \frac{B_k^T \partial f(x_{k+1})}{\|B_k^T \partial f(x_{k+1})\|}, \quad h_{k+1} = \frac{f(x_{k+1})-f^*}{\|B_k^T \partial f(x_{k+1})\|}. \qquad (8)$$

3.Обчислимо

$$\lambda_1 = -\frac{(p_k,\xi_{k+1})}{\sqrt{(p_k,\xi_{k+1})^2+(\xi_k,\xi_{k+1})^2}}, \quad \lambda_2 = -\frac{(\xi_k,\xi_{k+1})}{\sqrt{(p_k,\xi_{k+1})^2+(\xi_k,\xi_{k+1})^2}}$$

і приймем

$$p_{k+1} = \begin{cases} \lambda_1 p_k + \lambda_2 \xi_k, & \text{якщо } \lambda_1 > 0 \text{ и } \lambda_2 > 0, \\ p_k, & \text{якщо } \lambda_1 > 0 \text{ і } \lambda_2 \le 0, \\ \xi_k, & \text{якщо } \lambda_1 \le 0 \text{ і } \lambda_2 > 0, \\ 0, & \text{якщо } \lambda_1 \le 0 \text{ і } \lambda_2 \le 0. \end{cases} \tag{9}$$

4.Якщо $(p_{k+1}, \xi_{k+1}) \ge 0$, приймем $B_{k+1} = B_k$ і перейдем до п. 5. Інакше обчислимо

$$\eta = \left(\frac{1}{\sqrt{1-(p_{k+1},\xi_{k+1})^2}} - 1 \right) \xi_{k+1} - \frac{(p_{k+1},\xi_{k+1})}{\sqrt{1-(p_{k+1},\xi_{k+1})^2}} p_{k+1}, \tag{10}$$

$$B_{k+1} = B_k \left(I + \eta \xi_{k+1}^T \right), \quad h_{k+1} = \frac{h_{k+1}}{\sqrt{1-(p_{k+1},\xi_{k+1})^2}}, \tag{11}$$

$$p_{k+1} = \frac{1}{\sqrt{1-(p_{k+1},\xi_{k+1})^2}} \left(p_{k+1} - (p_{k+1},\xi_{k+1}) \xi_{k+1} \right). \tag{12}$$

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

Для наведеного методу скінцеву збіжність до ε_f-розв'язку задачі (4) з любої стартової точки гарантує теорема (такий самий результат має місце і для першого методу).

Теорема 7. Нехай на кожній ітерації методу (7)–(12) $\|B_k\| \le c_1$ і $\|\partial f(x_k)\| \le c_2$. Тоді методом (7)–(12) розв'язується задача (4) з точністю ε_f за функціоналом, не більше ніж за K ітерацій, де $K = \left] \left(\frac{c_1 c_2 \|x_0 - x^*\|}{\varepsilon_f} \right)^2 \right[+ 1$. □

Показано, що H-форми наведених методів співпадають з H-формами їх аналогів із §1.3, в силу чого тестові експерименти з §1.3 мають місце і для методів §3.2. На тестових задачах перевірена стійкість методів по відношенню до знаходження дуже точних ε_f-розв'язків. Обговорюються можливі шляхи побудови і обгрунтування ряду інших методів змінної метрики на базі однорангового еліпсоїдального оператора.

У §3.3 розглянутий доортогоналізуючий однорангoвий оператор перетворення простору аргументів (лінійний оператор із R^n в R^n), який в матричному вигляді буде таким:

$$T_\lambda(p_m, p) = I + \frac{p_m - p}{\|p_m - p\|^2} \left(\frac{1}{\lambda} p_m + p \right)^T, \tag{13}$$

де I – одинична матриця розміром $n*n$, λ – скаляр, такий, що $\lambda(\lambda+1) \ne 0$. Тут $p = \sum_{i=1}^{m-1} (p_m, p_i) p_i$, а набір векторів

$$P = \{p_1, p_2, ..., p_m\},\ \|p_i\| = 1,\ i = 1, \ldots, m,\ m \le n, \tag{14}$$

в R^n, що задовольняє таким умовам:

$$(p_i, p_j) = 0,\ i \ne j,\ i = 1, 2, \ldots, m-1,\ j = 1, 2, \ldots, m-1, \tag{15}$$

$$\|p_m - \sum_{i=1}^{m-1}(p_m,p_i)p_i\|^2 = 1 - \sum_{i=1}^{m-1}(p_m,p_i)^2 > 0. \quad (16)$$

Обернений до (13) оператор

$$T_\lambda^{-1}(p_m,p) = I - \frac{p_m - p}{\|p_m - p\|^2}\left(\frac{1}{\lambda+1}p_m + \frac{\lambda}{\lambda+1}p\right)^T. \quad (17)$$

Для доортогоналізуючого однорангового оператора має місце лема.

Лема 8. Нехай набір векторів (14) задовольняє (15) і (16); $\tilde{P} = \{\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_m\}$ – набір векторів, які одержані з (14); $\tilde{p}_i = (T_\lambda^{-1}(p_m,p))^T p_i$ для всіх $i = 1,2,\ldots,m$ при $\lambda(\lambda+1) \neq 0$. Тоді ні один з векторів $\tilde{p}_i$ тотожно не дорівнює нульовому вектору $(\tilde{p}_i, \tilde{p}_j) = 0$, $i \neq j$, $i = 1,2,\ldots,m$, $j = 1,2,\ldots,m$. Крім того,

$$\tilde{p}_i = p_i, \quad i = 1,2,\ldots,m-1, \quad \tilde{p}_m = \frac{\lambda}{\lambda+1}(p_m - p).$$

Лема 8 означає, що набір m векторів в R^n, з яких перші $m-1$ векторів взаємно ортогональні, а останній не ортогональний попереднім, переводиться за допомогою оператора (13) в набір взаємно ортогональних векторів (фактично доортогоналізовується m-й вектор, чим зумовлена назва оператора). Оператор (13) допускає однорангову корекцію матриці B_{k+1} при переході в черговий перетворений простір аргументів і забезпечує досить простий механізм роботи з спеціальним конусом. Його використання дозволяє будувати конструктивні методи змінної метрики з досить наглядною геометричною інтерпретацією в перетвореному просторі аргументів. Деякі з таких методів обговорюються в дисертації.

Для знаходження ε_f-розв'язку задачі (4) за допомогою оператора (13) побудоване сімейство методів ортогонального (в перетвореному просторі аргументів) субградієнтного спуску з класичним фейєровським кроком в напрямі антисубградієнта – $ORTGF(x_0, \varepsilon_f, \lambda, \varepsilon_K, \varepsilon_R, m_0)$, де параметри ε_K і ε_R задають правила формування тупого конуса, який переводиться в ортогональний, а параметр m_0 – максимальне число утворюючих конус субградієнтів. Метод $ORTGF(x_0, \varepsilon_f, \lambda, \varepsilon_K, \varepsilon_R, m_0)$.

Перед початком процесу $B_0 = I_n$, $P_0 = \emptyset$ (P_k – множина утворюючих конус субградієнтів).

Нехай на k-й ітерації одержані x_k, B_k, P_k. Обчислимо $f(x_k)$ і $\partial f(x_k)$. Якщо $f(x_k) - f^* \leq \varepsilon_f$, то x_k – ε_f-розв'язок і кінець роботи методу. Інакше переходимо до $(k+1)$-ї ітерації.

1.Приймем

$$\xi_k = \frac{B_k^T \partial f(x_k)}{\|B_k^T \partial f(x_k)\|}, \quad h_k = \frac{f(x_k) - f^*}{\|B_k^T \partial f(x_k)\|}.$$

2.Сформуємо множину

$$\tilde{P}_k = \{p_i \quad P_k : (p_i, \xi_k) < -\varepsilon_K\},$$

зберігаючи порядок черговості p_i в P_k.

3.Якщо $size(P_k) = 0$, то приймем

$$B_{k+1} = B_k, \quad \xi_{k+1} = \xi_k, \quad h_{k+1} = h_k$$

і перейдем до п. 4. Інакше обчислюємо вектор

$$\tilde{p}_k = \sum_{p_i \in P_k} (p_i, \xi_k) p_i$$

і перераховуємо параметри

$$B_{k+1} = B_k(I - \eta_1 \eta_2^T), \text{ де } \eta_1 = \frac{\xi_k - \tilde{p}_k}{\|\xi_k - \tilde{p}_k\|^2}, \quad \eta_2 = \frac{1}{\lambda+1}\xi_k + \frac{\lambda}{\lambda+1}\tilde{p}_k;$$

$$\xi_{k+1} = \frac{\frac{\lambda}{\lambda+1}(\xi_k - \tilde{p}_k)}{\|\frac{\lambda}{\lambda+1}(\xi_k - \tilde{p}_k)\|}, \quad h_{k+1} = \frac{h_k}{\|\frac{\lambda}{\lambda+1}(\xi_k - \tilde{p}_k)\|}.$$

4.Обчислимо чергове наближення

$$x_{k+1} = x_k - h_{k+1} B_{k+1} \xi_{k+1}.$$

5.Сформуємо чергову множину

$$P_{k+1} = \{p_i \in \tilde{P}_k : |(p_i, \xi_{k+1})| < \varepsilon_\Pi\} \cup \xi_{k+1},$$

зберігаючи порядок черговості p_i в $\tilde{P}_k$; ξ_{k+1} буде останнім вектором в P_{k+1}. Якщо $size(P_{k+1}) > m_0$, тоді

$$P_{k+1} = P_{k+1} \setminus \{p_1 \in P_{k+1}\},$$

де p_1 – перший вектор в P_{k+1}.

6.Переходимо до наступної ітерації з x_{k+1}, B_{k+1}, P_{k+1}.

Для методу $ORTGF$ при любому значенні параметра λ справедлив теорема.

<u>Теорема 8</u>. Для k-ї ітерації методу $ORTGF$

$$(A_k(x_k - x^*), \xi_k) \geq \frac{f(x_k) - f^*}{\|B_k^T \partial f(x_k)\|} \geq 0,$$

де $A_k = B_k^{-1}$. Коли $P_k \neq \emptyset$, то $\|p_i\| = 1$ для всіх $p_i \in P_k$; $(p_i, p_j) = 0$ для $p_i, p_j \in P_k$, таких, що $i \neq j$; крім того, справедливі нерівності

$$(A_k(x_k - x^*), p_i) \geq 0, \quad p_i \in P_k,$$

$$(A_k(x_k - x^*), \tilde{p}_k) \leq 0.$$

В окремому випадку справедлива теорема, яка обгрунтовує збіжність методу $ORTGF$ по зменшенню об'єму еліпсоїда для локалізації множини екстремумів.

Теорема 10. Нехай на кожній ітерації методу $ORTGF$ при $\lambda = -\frac{1}{2}$ $\|B_k\| \le c_1$, $\|\partial f(x_k)\| \le c_2$. Тоді методом $ORTGF$ розв'язується задача (4) з точністю ε_f, не більше, як за K ітерацій, де $K =]\left(\frac{c_1 c_2 \|x_0 - x^*\|}{\varepsilon_f}\right)^2[+1$.

Проведені числові експерименти з $ORTGF$ при $\lambda = -0.5$ і $\lambda = 1$. для ряду гладких і негладких задач, що в цілому демонструють стійку роботу цих методів.

У §3.4 підведені підсумки розділу 3, які полягають у наступному. Використання розглянутих операторів перетворення простору, спрямоване на зменшення овражності функції, дозволяє значно поліпшити ефективність субградієнтних (ε-субградієнтних) методів. Це підтверджують тестування запропонованих субградієнтних методів фейеровського типу із змінною метрикою. Так, якщо за велику ітерацію рахувати n ітерацій одного з цих методів, що за обсягом обчислень можна порівняти з однією ітерацією методу Ньютона, то знаходження розв'язку з досить високою точністю в тестових задачах (в тому числі і з овражною структурою поверхностей рівня) вимагало порівняно незначної кількості великих ітерацій. Для задачі опуклого програмування використання операторів (5) і (1) в поєднанні з оператором розтягу простору дозволяє побудувати серію ефективних методів змінної метрики з обгрунтуванням їх збіжності, використовуючи монотонне зменшення об'єму еліпсоїдів для локалізації множини екстремумів. При цьому для методів з B-формі можна дати просту геометричну інтерпретацію процесів у перетвореному просторі аргументів.

У висновках наведені основні наукові результати роботи.

ОСНОВНІ РЕЗУЛЬТАТИ РОБОТИ

1. Побудований приклад опуклої кусочно-лінійної функції, для якої можливе "зациклювання" $r_\mu(\alpha)$-алгоритму. Обговорюються модифікації $r_\mu(\alpha)$-алгоритму, що дозволяють обійти таку ситуацію при мінімізації негладких опуклих функцій.

2. Аналізуються головні причини, які затрудняють обгрунтування збіжності r-алгоритмів. Обговорюється підхід, заснований на зовнішній апроксимації множини екстремумів еліпсоїдами спеціального типу, який дозволяє будувати ефективні методи з розтягом простору і обгрунтовувати їх збіжність для мінімізації негладких опуклих функцій. Для перетворення одного з таких апроксимуючих еліпсоїдів в кулю використовується перетворення простору, близьке до того, що має місце в r-алгоритмах.

3. Побудовані два субградієнтних методи з розтягом простору для знаходження точки мінімуму опуклої функції з відомим значенням функції в

оптимумі. Обгрунтування їх збіжності засновано на монотонному зменшенні об'єму еліпсоїда для локалізації множини экстремумів. Проведене чисельне порівняння цих методів з $r(\alpha)$-алгоритмом.

4. Розглянута загальна конструкція методів центрів тяжіння простих тіл (МЦТПТ), яка для задач опуклого програмування дозволяє будувати методи першого порядку, економні за складністю обробки інформації і збіжні за послідовністю рекордів функції, що мінімізується. МЦТПТ заснована на зовнішній апроксимації множини екстремумів простими геометричними тілами з гарантованим зменшенням їх об'єму і регуляризацією цих тіл за допомогою лінійних неортогональних перетворень простору.

5. Вивчені властивості двох класів лінійних неортогональних операторів перетворення простору: однорангового еліпсоїдального і доортогоналізуючого однорангового операторів. Для знаходження точки мінімуму опуклої функції за відомим оптимальним значенням функції на їх основі побудовані та обгрунтовані прості субградієнтні методи змінної метрики, які використовують класичний фейєровський крок в напрямі антисубградієнта: два методи по типу r-алгоритмів на основі однорангового еліпсоїдального оператора і сімейство методів ортогонального субградієнтного спуску на основі доортогоналізуючого однорангового оператора. Наведені тестові експерименти, які підтверджують їх ефективність при мінімізації овражних негладких функцій.

6. Запропоновані в дисертаційній роботі алгоритми програмно реалізовані, проведено їх тестування. Вони можуть бути використані при розв'язуванні задач опуклого програмування, зокрема, для знаходження допустимої точки системи опуклих нерівностей.

Основні положення дисертації опубліковані в таких працях:

1. Стецюк П.И. К вопросу сходимости r-алгоритмов // Кибернетика и сист. анализ. – 1995. – N6. – С. 173-177.

2. Стецюк П.И. К обоснованию сходимости алгоритмов с растяжением пространства // Теория оптимальных решений. - Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 1995. – С.4-8.

3. Стецюк П.И. Об одной схеме методов отсечений.– Киев, 1995. – 34 с. – (Препр. / НАН Украины. Ин-т кибернетики им. В.М.Глушкова; 95-25).

4. Стецюк П.И. r-алгоритмы и эллипсоиды // Кибернетика и сист. анализ. – 1996. – N1. – С. 113-134.

5. Стецюк П.И. Классические фейеровские методы с преобразованием пространства // Методы решения экстремальных задач. - Киев: Ин-т кибернетики им. В.М.Глушкова НАН Украины, 1996. – С.3-9.

Стецюк П.И. Субградиентные методы с преобразованием пространства для минимизации негладких выпуклых функций. Рукопись.

Диссертация на соискание ученой степени кандидата физико-математических наук по специальности 01.05.01 – теоретические основы информатики и кибернетики (математическая кибернетика), Институт кибернетики имени В.М.Глушкова НАН Украины, Киев, 1996.

Исследуются эффективные методы субградиентного типа с преобразованием пространства, использующие внешнюю аппроксимацию множества экстремумов простыми телами с монотонным уменьшением их объема. Для нахождения точки минимума выпуклой функции при известном ее оптимальном значении построен и обоснован ряд субградиентных методов переменной метрики. Практическая эффективность предложенных методов проверена для тестовых задач минимизации негладких выпуклых функций.

Stetsyuk P.I. The subgradient-type methods with space transformations for minimization of nonsmooth convex functions.

The dissertation for the degree of a candidate of physics and mathematics on speaciality 01.05.01 – theoretical basis of informatics and cybernetics (mathematical cybernetics). V.M.Glushkov Institute of Cybernetics, NAS of Ukraine, Kiev, 1996.

The effective subgradient-type methods with space transformations using the outer approximation of the extremum set by simple bodies with monotonic decreasing of their volume was investigated. For finding of the optimal point of convex function with its known optimal value several variable metrix subgradient-type methods were constructed and proved. Practical effectivness of these methods was tested for nonsmooth convex function minimization problem.

Ключові слова:

негладка опукла функція, r-алгоритми, методи змінної метрики, зовнішня апроксимація множини екстремумів, екстремальні еліпсоїди, збіжність за послідовністю рекордів функції, яка мінімізується, лінійний оператор перетворення простору.

Підп. до друку 11.07.96. Формат 60×84/16. Папір офісний. Офс. друк. Ум. друк. арк. 0,93. Ум. фарбо-відб. 1,05. Обл.-вид. арк. 1,0. Тираж 100 прим. Зам. 342.

Редакційно-видавничий відділ з поліграфічною дільницею
Інституту кібернетики імені В. М. Глушкова НАН України
252022 Київ 22, проспект Академіка Глушкова, 40