Одеський національний університет імені І. І. Мечникова
Інститут математики, економіки і механіки
Кафедра оптимального керування і економічної кібернетики

Дипломна робота

магістра

на тему: **«Назва роботи українською мовою»**

«Title in English»

Виконав: студент денної форми навчання
спеціальності 113 Прикладна математика
Прізвище Ім'я Батькович

Керівник: к. ф.-м. н. Страхов Є. М.
Рецензент: к. т. н., доц. Климчук О. А.

Рекомендовано до захисту:	Захищено на засіданні ЕК № ______
Протокол засідання кафедри	Протокол № ____ від «____» ____ р.
№ ___ від «____» _______ р.	Оцінка _______ / _______ / _______
Завідувач кафедри	Голова ЕК
__________ ______________	__________ _________________

Одеса – 2017 р.

ЗМІСТ

ВСТУП

Повштухом до росту зацікавленості теорією та практикою математичного програмування стало відкриття в 1947 році обчислювального методу розв’язування завдань лінійного програмування. Цей чисельний метод був названий симплекс методом.

Одночасно з підвищенням інтересу до лінійного програмування, популярність здобували й нелінійні задачі. У 1951 році була опублікована праця Куна і Таккера, у якій було викладено необхідні та достатні умови оптимальності розв’язку нелінійних задач. Саме ця праця стала фундаментом до багатьох наступних робіт нелінійного програмування.

Починаючи з 1954 року почала з’являтися велика кількість праць, які присвячені квадратичному програмуванню. Але у більшості робіт були майже однакові обчислювальні алгоритми.

На сучасному етапі розвитку науки та комп’ютерних технологій теорія чисельних методів нелінійного програмування є досить розвиненою. Проте ще досить важко дати чіткі рекомендації щодо застосування того чи іншого методу. Тому теорія методів оптимізації продовжує розвиватися, з’являються нові методи, які мають свої переваги у порівнянні з попередніми.

У літературі описані багатокрокові методи, а саме - двокроковий метод спряжених градієнтів. Цей метод вважається досить ефективним для задач великої розмірності. Метод спряжених градієнтів має перевагу перед однокроковими градієнтрими методами, бо він у більшій мірі враховує геометричні властивості цільової функції. Опираючись на цю інформацію, можна піти далі і розглянути трикрокові, чотирикрокові, п’ятикрокові і т.д. методи.

У роботі розглянуто чотирикроковий метод, дано чисельну порівняльну характеристику у порівнянні з трикроковим методом.

Мета роботи – дослідити чотирикроковий метод мінімізації функцій.

Відповідно до поставленої мети в роботі вирішуються такі конкретні **завдання**:

1) розглянути теоретичні засади чотирьохкрокового методу;
2) перевірити його доцільність на прикладах.

Об'єкт дослідження – пошук мінімуму функцій.

Предмет – чотирикроковий метод мінімізації функцій.

РОЗДІЛ 1

МЕТОДИ НЕЛІНІЙНОГО ПРОГРАМУВАННЯ БЕЗ ОБМЕЖЕНЬ

1.1. Постановка задачі нелінійного програмування

У загальному вигляді математична модель задачі нелінійного програмування формулюється наступним чином: знайти вектор , $X = (x_1, x_2, \ldots, x_n)$ при якому б цільова функція набувала свого екстремального (максимального чи мінімального) значення (можливо лише в тому випадку коли цільова функція неперервна, а допустима множина розв'язків замкнена, непуста і обмежена):

$$\max(\min)\Phi = \phi(x_1, x_2, \ldots, x_n) \tag{1.1.1}$$

при наступних обмеженнях:

$$\begin{cases} g_1(x_1, x_2, \ldots, x_n)\{\leq, =, \geq\}b_1 \\ g_2(x_1, x_2, \ldots, x_n)\{\leq, =, \geq\}b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ g_m(x_1, x_2, \ldots, x_n)\{\leq, =, \geq\}b_m \end{cases} \tag{1.1.2}$$

$$x_j \geq 0, \ (j = \overline{1,n}) \tag{1.1.3}$$

де x_j $(j = \overline{1,n})$ - змінні
$\phi(x_1, x_2, \ldots, x_n)$ - нелінійна функція від n змінних,
$g_i(x_1, x_2, \ldots, x_n)$ $(i = \overline{1,m})$ - обмеження
b_i $(i = \overline{1,m})$ - фіксовані значення
m - кількість обмежень
n - кількість параметрів

У випадку відсутності обмежень маємо задачу безумовної оптимізації.

З курсу вищої математики відомо, що в точці екстремуму (мінімуму, максимуму) нелінійної функції всі її часткові похідні дорівнюють нулю. Отже, для знаходження екстремуму нелінійної функції n змінних необхідно визначити її часткові похідні за всіма змінними і прирівняти їх до нуля. Розв'язок отриманої системи n рівнянь з n невідомими дасть значення змінних, при яких має досягаться екстремум функції.

Слід зазначити, що точний розв'язок системи рівнянь, в загальному випадку системи нелінійних рівнянь, являє собою досить складне завдання. Тому для пошуку екстремуму нелінійної функції часто використовуються інші методи, зокрема градієнтні методи.

Задачі безумовної мінімізації на практиці зустрічаються рідко, однак методи їхнього розв'язку є основою розв'язку більшості практичних задач умовної оптимізації.

1.2. Загальна характеристика методів спуску

Всі методи розв'язку задачі безумовної оптимізації полягають у побудові послідовністі точок $\{x^n\}$ так, щоб послідовність функцій $\phi(x^n)$ була спадною (тобто спуск уздовж функції). На k-му кроці $(k > 0)$ визначається вектор s^k, в напрямку якого функція $\phi(x^k)$ зменшується. У цьому напрямку робиться крок величиною β_k і отримується нова точка $x^{k+1} = x^k + \beta_k s^k$, в якій $\phi(x^{k+1}) < \phi(x^k)$. Послідовність $\{x^n\}_{n\in\mathbb{N}}$, що задовольняє цій умові, називається релаксаційною послідовністю, а відповідні методи – методами спуску.

Методи спуску поділяються на методи із застосуванням інформації про похідні функції і без використання такої. Різні методи спуску відрізняються вибором напрямку і величиною кроку. Як правило, для знаходження β_k використовується процедура одновимірного пошуку. Щодо побудови напрямків спуску розроблено різні підходи, які визначають подальшу класифікацію методів. Зокрема, ті методи, які при побудові напрямків спуску використовують лише інформацію поточної ітерації, називають однокроковими. Якщо використовується додатково інформація попереднього кроку, то двокроковими, і т.д. Класичним двокроковим методом є метод спряжених градієнтів.

Методи безумовної оптимізації також можна поділити на:

- **Методи прямого пошуку**

У методах прямого пошуку мінімуму цільової функції (або методах нульового порядку) використовується інформація лише про значення функції. Багато з цих методів не мають строгого теоретичного обґрунтування і побудовані на основі евристичних міркувань. Тому питання збіжності методів прямого пошуку ще мало вивчені, а оцінки швидкості збіжності зазвичай відсутні. Разом із цим дані методи ідейно пов'язані з методами першого і другого порядків, що в ряді випадків дозволяє оцінювати ефективність алгоритмів прямого пошуку стосовно мінімізації деяких класів функцій. Поширеним способом оцінки ефективності методів прямого пошуку є обчислювальні експерименти та порівняльний аналіз методів за результатами таких експериментів.

До методів нульового порядку належать методи, які не використовують похідні для вибору напрямку спуска: метод Гауса, метод Хука Дживса, метод обертових напрямків (Розенброка), метод деформованого багатогранника (пошук по симплексу), метод Пауелла.

- **Методи першого порядку**

Методи 1-го порядку використовують інформацію про похідну функції. Якщо обмежена знизу цільова функція $\phi(x), x \in \mathbb{R}^n$ є диференційованою на множині $\mathbb{R}^n$, то алгоритм пошуку точки x^* її мінімуму можна побудувати, використовуючи інформацію, принаймні, про градієнт цієї функції. Такі методи називаються градієнтними. Градієнтні методи безумовної оптимізації використовують тільки перші похідні цільової функції і є методами лінійної апроксимації на кожному кроці, тому, що цільова функція на кожному кроці замінюється дотичною гіперплощиною до її графіку в поточній точці. У всіх цих методах передбачається, що $\phi'(x)$ існують і неперервні. Градієнтні методи розрізняються тільки способом визначення β_k і s^k.

До методів першого порядку належать методи: найшвидшого спуску (Коші) та спряжених градієнтів.

- **Методи 2-го порядку (Ньютонівські методи)**

Коли цільова функція $\phi(x)$ двічі диференційована в $\mathbb{R}^n$, то ефективність процесу пошуку точки x^* її мінімуму можна підвищити, використовуючи інформацію не тільки про градієнт цієї функції, а й про її матрицю Гессе $H(x)$. Напрям пошуку, що відповідає найшвидшому спуску, пов'язаний з лінійною апроксимацією цільової функції. Методи, які використовують інформацію про другі похідні, виникли із квадратичної апроксимації цільової функції $\phi(x)$, яку можна отримати при розкладанні функції в ряд Тейлора 2-го порядку. Мінімум $\phi(x)$ (якщо він існує) досягається там же, де і мінімум квадратичної форми.

Якщо матриця Гессе цільової функції, обчислена в точці x^k, є додатно визначеною, то точка x^* єдина і може бути знайдена з умови, що градієнт функції дорівнює нульовому вектору. Алгоритм оптимізації, в якому напрям пошуку формулюється з цього співвідношення, називається методом Ньютона.

В задачах знаходження мінімуму довільної квадратичної функції із додатною матрицею других похідних метод Ньютона дає рішення за одну ітерацію незалежно від вибору початкової точки.

До методів 2-го порядку належать: метод Ньютона-Рафсона, модифікації методу Ньютона.

- **Методи змінної метрики**

Серед алгоритмів багатовимірної мінімізації виділяють групу алгоритмів, що поєднують переваги методів Ньютона та найшвидшого спуску. Дані алгоритми прийнято відносити до так званим квазіньютонівським методам. Особливістю цих алгоритмів є те, що при їх використанні немає необхідності обертати й обчислювати матрицю Гессе цільової функції $\phi(x)$ і у цей же час вдається зберегти високу швидкість збіжності алгоритмів, що притаманна методам Ньютона та його модифікаціям. У цих методах обернена матриця Гессе апроксимується іншою матрицею. Метрика змінюється на кожній ітерації, і тому методи так само називаються методами зі змінною метрикою.

До методів змінної метрики належать наступні методи: Пірсона, Девідона Флетчера Пауелла, Бройдена Флетчера Шенно, Пауелла і Мак Корміка, і інші.

- **Методи випадкового пошуку**

Методи випадкового пошуку реалізують ітеративний процес руху оптимізаційних змінних в просторі з використанням випадкових напрямків. Одна з переваг цих методів – достатня простота, методи володіють великим спектром можливих напрямків руху.

Можливі два алгоритми пошуку. Алгоритм пошуку з лінійною стратегією: визначений напрямок, в якому цільова функція зменшується, не змінюється до тих пір, поки він не призведе до збільшення цільової функції.

Стратегія лінійного пошуку хороша, коли далеко до оптимуму. Поблизу оптимуму більш доцільна нелінійна стратегія, при якій випадкова зміна напрямків не залежить від результату.

1.3. Метод спряжених градієнтів

Розглянемо задачу мінімізації функції багатьох змінних без обмежень

$$\phi(x) \to min,\ x \in \mathbb{R}^n, \tag{1.3.1}$$

де

$\phi(x) : \mathbb{R}^n \to \mathbb{R}^1$ - неперервно диференційовна функція

Позначимо її градієнт через $\phi'(x)$.

Метод спряжених градієнтів має такий загальний вигляд:

$$x^{k+1} = x^k + \beta_k s^k,\ k = 0, 1, \ldots$$

$$S = \begin{cases} -\phi'(x^k) & , k = 0 \\ -\phi'(x^k) + \gamma_1^{k-1} s^{k-1} & , k = 1, 2, \ldots \end{cases} \tag{1.3.2}$$

де $x^0, x^1, \ldots, x^k, \ldots$ - послідовні наближення,
$s^0, s^1, \ldots, s^k, \ldots$ - напрямки спуску,
β_k - величина кроку вздовж напрямку спуску,
γ_{k-1} - числовий параметр

Одним із можливих варіантів вибору кроку β_k є розв'язування задачі одновимірної мінімізації:

$$\beta_k : \min_{\beta \geq 0} \phi(x^k + \beta s^k) \tag{1.3.3}$$

Практичні дослідження доводять, що якість розв’язку задачі (1.3.1) та швидкість збіжності алгоритму суттєво залежать від якості розв’язку одновимірної задачі (1.3.3).

Різновиди методу (1.3.2) визначаються способом обчислення параметру γ_{k-1}, зокрема

- $\gamma_{k-1} = \dfrac{\|\phi'(x^k)\|^2}{\|\phi'(x^{k-1})\|^2}$ (метод Флетчера - Рівза)
- $\gamma_{k-1} = \dfrac{(\phi'(x^k), \phi'(x^k) - \phi'(x^{k-1}))}{\|\phi'(x^{k-1})\|^2}$ (метод Полака - Ріб’єра)

1.4. Трикроковий метод для задачі багатовимірної оптимізації

Для задачі нелінійного програмування (1.3.1) трикроковий метод має такий алгоритм[6]:

$$x^{k+1} = x^k + \beta_k s^k, \ k = 0, 1, \ldots$$

$$S = \begin{cases} -\phi'(x^k) & , k = 0 \\ -\phi'(x^k) + \gamma_1^{k-1} s^{k-1} & , k = 1 \\ -\phi'(x^k) + \gamma_1^{k-1} s^{k-1} + \gamma_2^{k-2} s^{k-2} & , k = 2, \ldots \end{cases} \tag{1.4.1}$$

де $x^0, x^1, \ldots, x^k, \ldots$ - послідовні наближення

$s^0, s^1, \ldots, s^k, \ldots$ - напрямки спуску

$\beta_k, \gamma_{k-1}, \gamma_{k-2}$ - числові параметри

РОЗДІЛ 2

ЧОТИРЬОХКРОКОВИЙ МЕТОД МІНІМІЗАЦІЇ ФУНКЦІЙ БАГАТЬОХ ЗМІННИХ БЕЗ ОБМЕЖЕНЬ

Для розв'язання задачі мінімізації функції багатьох змінних без обмежень (1.3.1) розглянемо чотирьохкроковий метод з таким алгоритмом:

$$x^{k+1} = x^k + \beta_k s^k,\ k = 0, 1, \ldots$$

$$S = \begin{cases} -\phi'(x^k) & ,k=0 \\ -\phi'(x^k) + \gamma_1^{k-1} s^{k-1} & ,k=1 \\ -\phi'(x^k) + \gamma_1^{k-1} s^{k-1} + \gamma_2^{k-2} s^{k-2} & ,k=2 \\ -\phi'(x^k) + \gamma_1^{k-1} s^{k-1} + \gamma_2^{k-2} s^{k-2} + \gamma_3^{k-3} s^{k-3} & ,k=3,\ldots \end{cases} \tag{2.0.1}$$

де $x^0, x^1, \ldots, x^k, \ldots$ - послідовні наближення
$s^0, s^1, \ldots, s^k, \ldots$ - напрямки спуску
$\beta_k, \gamma_j^m (j = \overline{1,3})$ - числові параметри

Параметр β_k будемо визначати з умови (1.3.3).

Означення 2.1. Вектори s' і s'' називаються спряженими (відносно матриці A), якщо вони відмінні від нуля і $(As', s'') = 0$.
Вектори $s^0, s^1, ..., s^m$ називаються взаємно спряженими (відносно матриці A), якщо всі вони відмінні від нуля і $(As', s'') = 0, i \neq j, 0 \leq i,j \geq m$. Матриця A вважається симетричною і додатньо визначеною $(A > 0)$.

2.1. Теоретичні положення методу

Розглянемо деякі властивості методу при умові, що функція $\phi(x)$ є квадратичною.

$$\phi(x) = \frac{1}{2}(Ax,x) + (b,x) + c \tag{2.1.1}$$

Побудуємо систему взаємно спряжених напрямків за правилом (2.0.1).

$$0 = (s^k, As^{k-1}) = -(\phi'(x^k), As^{k-1}) + \gamma_1^{k-1}(s^{k-1}, As^{k-1}) +$$

$$+\gamma_2^{k-2} \underbrace{(s^{k-2}, As^{k-1})}_{0} + \gamma_3^{k-3} \underbrace{(s^{k-3}, As^{k-1})}_{0} \Longrightarrow$$

$$\gamma_1^{k-1} = \frac{(\phi'(x^k), As^{k-1})}{(s^{k-1}, As^{k-1})} \tag{2.1.2}$$

Внаслідок того, що матриця A додатньо визначена, знаменник у (2.1.2) не дорівнює нулеві. Аналогічно отримаємо:

$$0 = (s^k, As^{k-2}) = -(\phi'(x^k), As^{k-2}) + \gamma_1^{k-1} \underbrace{(s^{k-1}, As^{k-2})}_{0} +$$

$$+\gamma_2^{k-2}(s^{k-2}, As^{k-2}) + \gamma_3^{k-3} \underbrace{(s^{k-3}, As^{k-2})}_{0} \Longrightarrow$$

$$\gamma_2^{k-2} = \frac{(\phi'(x^k), As^{k-2})}{(s^{k-2}, As^{k-2})} \tag{2.1.3}$$

$$0 = (s^k, As^{k-3}) = -(\phi'(x^k), As^{k-3}) + \gamma_1^{k-1} \underbrace{(s^{k-1}, As^{k-3})}_{0} +$$

$$+\gamma_2^{k-2} \underbrace{(s^{k-2}, As^{k-3})}_{0} + \gamma_3^{k-3}(s^{k-3}, As^{k-3}) \Longrightarrow$$

$$\gamma_3^{k-3} = \frac{(\phi'(x^k), As^{k-3})}{(s^{k-3}, As^{k-3})} \tag{2.1.4}$$

Теорема 2.1. *Для диференційовної функції $\phi(x)$ послідовність $\{x^k\}$, що побудована за (2.0.1), (2.1.2), (2.1.3), (2.1.4), така, що*

$$(\phi'(x^{k+1}), s^k) = 0,\ k = 0, 1, \dots \tag{2.1.5}$$

Доведення: Враховуючи (2.0.1) отримаємо $Ax^{k+1} = Ax^k + \beta_k As^k$. Так як $\phi'(x) = Ax + b$, то маємо

$$\phi'(x^{k+1}) = \phi'(x) + \beta_k As^k \tag{2.1.6}$$

З (2.0.3) отримаємо, що

$$\begin{aligned} &\frac{d}{d\beta}\phi(x^k+\beta s^k)\Big|_{\beta=\beta_k}=0 \qquad &, \ \beta_k>0 \\ &\frac{d}{d\beta}\phi(x^k+\beta s^k)\Big|_{\beta<0} \ \geq 0 \qquad &, \ \beta_k=0 \end{aligned}$$

Якщо $\beta_k > 0$, то

$$0=\frac{d}{d\beta}\phi(x^k+\beta s^k)\Big|_{\beta=\beta_k}=(\phi'(x^k+\beta_k s^k),s^k)=(\phi'(x^{k+1}),s^k)$$

Отже, отримали, що $(\phi'(x^{k+1}), s^k) = 0, \ k = 0, 1, \ldots$

Застосуємо метод індукції для доведення того, що співвідношення (2.1.5) справедливе і при $\beta_k = 0$:

1) $0 \leq \dfrac{d}{d\beta}\phi(x^0+\beta s^0)\Big|_{\beta=0}=(\phi'(x^1),s^0)=(\phi'(x^0),-\phi'(x^0))=$
   $=-\|\phi'(x^0)\|^2 \Rightarrow (\phi'(x^1),s^0)=0$
2) припустимо, що $(\phi'(x^k), s^{k-1}) = 0$
3) доведемо, що $(\phi'(x^{k+1}), s^k) = 0, \beta_k = 0$
   Так, як $x^{k+1} = x^k$, то враховуючи (2.1.6) маємо:

$$\begin{gathered} 0\leq\frac{d}{d\beta}\phi(x^k+\beta s^k)\Big|_{\beta=0}=(\phi'(x^{k+1}),s^k)=(\phi'(x^k),s^k)= \\ =(\phi'(x^k),-\phi'(x^k)+\gamma_1^{k-1}s^{k-1}+\gamma_2^{k-2}s^{k-2}+\gamma_3^{k-3}s^{k-3})= \\ =-(\phi'(x^k),\phi'(x^k))+\gamma_1^{k-1}(\phi'(x^k),s^{k-1})+ \\ +\gamma_2^{k-2}(\phi'(x^k),s^{k-2})+\gamma_3^{k-3}(\phi'(x^k),s^{k-3}) \end{gathered}$$

Враховуючи (2.1.6) і припущення індукції маємо:

$$\begin{gathered} (\phi'(x^k),s^{k-2})=(\phi'(x^{k-1})+\beta_{k-1}As^{k-1},s^{k-2})=(\phi'(x^{k-1}),s^{k-2})+ \\ +\beta_{k-1}(As^{k-1},s^{k-2})=(\phi'(x^{k-1}),s^{k-2})=0 \\ (\phi'(x^k),s^{k-3})=(\phi'(x^{k-1})+\beta_{k-1}As^{k-1},s^{k-3})=(\phi'(x^{k-1}),s^{k-3})+ \\ +\beta_{k-1}(As^{k-1},s^{k-3})=(\phi'(x^{k-1}),s^{k-3})=(\phi'(x^{k-2})+\beta_{k-2}As^{k-2},s^{k-3})= \end{gathered}$$

$$= (\phi'(x^{k-2}), s^{k-3}) + \beta_{k-2}(As^{k-2}, s^{k-3}) = (\phi'(x^{k-2}), s^{k-3}) = 0$$

Отже, отримали:

$$0 \leq (\phi'(x^{k+1}), s^k) = -\|\phi'(x^k)\|^2 \leq 0$$

Таким чином довели, що $(\phi'(x^{k+1}), s^k) = 0$

$\square$

Теорема 2.2. *Вектори* $\phi'(x^k)$ *і* $\phi'(x^{k+1})$ *ортогональні,* $k = 0, 1, \ldots$

Доведення: Відомо, що квадратична функція (2.1.1) досягає мінімального значення при

$$\beta_k = -\frac{(\phi'(x^k), s^k)}{(As^k, s^k)} \tag{2.1.7}$$

Тоді враховуючи (2.1.6) та (2.1.7), отримаємо:

$$(\phi'(x^{k+1}), \phi'(x^k)) = (\phi'(x^k) + \beta_k As^k, \phi'(x^k)) =$$

$$= (\phi'(x^k) - \frac{(\phi'(x^k), s^k)}{(As^k, s^k)} As^k, \phi'(x^k)) =$$

$$= (\phi'(x^k), \phi'(x^k)) - \frac{(\phi'(x^k), s^k)}{(As^k, s^k)} (As^k, \phi'(x^k))$$

Розглянємо $(\phi'(x^k), s^k)$.

$$(\phi'(x^k), s^k) = -(\phi'(x^k), \phi'(x^k)) + \gamma_1^{k-1} \underbrace{(\phi'(x^k), s^{k-1})}_{0} +$$

$$+\gamma_2^{k-2} \underbrace{(\phi'(x^k), s^{k-2})}_{0} + \gamma_3^{k-3} \underbrace{(\phi'(x^k), s^{k-3})}_{0} =$$

$$= -(\phi'(x^k), \phi'(x^k)) \tag{2.1.8}$$

Далі,

$$(As^k, s^k) = (As^k, -\phi'(x^k) + \gamma_1^{k-1} s^{k-1} + \gamma_2^{k-2} s^{k-2} + \gamma_3^{k-3} s^{k-3}) =$$

$$= -(As^k, \phi'(x^k)) + \gamma_1^{k-1}(As^k, s^{k-1}) + \gamma_2^{k-2}(As^k, s^{k-2}) + \gamma_3^{k-3}(As^k, s^{k-3}) =$$
$$= -(As^k, \phi'(x^k)) \qquad (2.1.9)$$

З урахуванням (2.1.8) і (2.1.9) отримаємо:

$$(\phi'(x^{k+1}), \phi'(x^k)) = (\phi'(x^k), \phi'(x^k)) - \frac{-(\phi'(x^k),\phi'(x^k))}{-(As^k,\phi'(x^k))}(As^k, \phi'(x^k)) = 0$$

□

Теорема 2.3. *Нехай* $x^0 \in \mathbb{R}^n$, *точки* $x^1, x^2, \ldots, x^{n-1}$ *і вектори* $s^0, s^1, \ldots, s^{n-1}$ *отримані за формулами (2.0.1), (2.1.2), (2.1.3), (2.1.4) і* $\phi'(x^k) \neq 0$ $(k = \overline{0,n-1})$, *тоді вектори* $s^0, s^1, \ldots, s^{n-1}$ *взаємно спряжені, а градієнти* $\phi'(x^0), \phi'(x^1), \ldots, \phi'(x^{n-1})$ *взаємно ортогональні.*

Доведення: Теорему будемо доводити методом математичної індукції.

1) $\phi'(x^0)$ i $\phi'(x^1)$ - ортогональні внаслідок теореми 2.2
   $s^0 \neq 0$ - за умовою теореми
   $s^1 \neq 0$, так як $s^1 = -\phi'(x^1) - \gamma_0\phi'(x^0) = 0$, а це неможливо, враховуючи ортогональність $\phi'(x^0)$ i $\phi'(x^1)$
   спряженість s^0 i s^1 отримаємо з (2.1.2), (2.1.3), (2.1.4)
2) Припустимо, що $k \leq n-1$
   вектори $s^0, s^1, \ldots, s^{k-1}$ - взаємно спряжені
   градієнти $\phi'(x^0), \phi'(x^1), \ldots, \phi'(x^{k-1})$ - взаємно ортогональні
3) За теоремою 2.2 $(\phi'(x^k), \phi'(x^{k-1})) = 0$
   при $i \leq k-2$, використовуючи (2.1.6), (2.0.2) та індукцію маємо:

$$(\phi'(x^k), \phi'(x^i)) = (\phi'(x^{k-1}), \phi'(x^i)) + \beta_{k-1}(As^{k-1}, \phi'(x^i)) =$$

$$= \beta_{k-1}(As^{k-1}, -s^k + \gamma_1^{k-1}s^{k-1} + \gamma_2^{k-2}s^{k-2} + \gamma_3^{k-3}s^{k-3}) = 0$$

   Взаємна ортогональність векторів $\phi'(x^0), \phi'(x^1), \ldots, \phi'(x^{k-1})$ доведена. Вектор $s^k \neq 0$, інакше вектори $\phi'(x^0), \ldots, \phi'(x^k)$ були б лінійнозалежними (враховуючи (2.0.1)), а це суперечить їх взаємній ортогональності. Доведемо, що вектори $s^0, \ldots, s^k$ взаємно спряжені.

За (2.1.2) $(s^k, As^{k-1}) = 0$, враховуючи (2.1.7) маємо:

$$\beta_i = -\frac{(\phi'(x^i), s^i)}{(As^i, s^i)} = -\frac{(\phi'(x^i), -\phi'(x^i) - \gamma_1^{k-1}\phi'(x^i) - \ldots)}{(As^i, s^i)} =$$

$$= \frac{(\phi'(x^i),\phi'(x^i))}{(As^i, s^i)}$$

з цього випливає, що $\beta_i \neq 0,\ i \leq k$, тоді з (2.1.6) отримаємо:

$$As^i = \frac{(\phi'(x^{i+1} - \phi'(x^i), s^i)}{\beta_i} \tag{2.1.10}$$

При $i \leq k-2$, використовуючи (2.0.1), індукцію і (2.1.10), та доведену взаємну ортогональність градієнтів, отримаємо:

$$(s^k, As^i) = (-\phi'(x^k) + \gamma_1^{k-1}s^{k-1} + \gamma_2^{k-2}s^{k-2} + \gamma_3^{k-3}s^{k-3}, As^i) =$$

$$= -\left(\phi'(x^k), \frac{(\phi'(x^{i+1} - \phi'(x^i), s^i)}{\beta_i}\right) = 0$$

□

Отже розглянутий чотирьохкроковий метод (2.0.1), (2.1.2), (2.1.3), (2.1.4) належить до методів спряжених напрямків.

Тепер сформулюємо чотирьохкроковий метод для мінімізації неквадратичних функцій.
Для цього перетворимо формули (2.1.2), (2.1.3), (2.1.4) так, щоб до них не входила матриця A.

$$\gamma_1^{k-1} = \frac{(\phi'(x^k), As^{k-1})}{(s^{k-1}, As^{k-1})} = \frac{(\phi'(x^k), \phi'(x^k) - \phi'(x^{k-1}))}{(s^{k-1},\phi'(x^k) - \phi'(x^{k-1}))} =$$

$$= \frac{(\phi'(x^k), \phi'(x^k) - \phi'(x^{k-1}))}{(-\phi'(x^{k-1}) - \gamma_2^{k-2}\phi'(x^i) - \ldots,\phi'(x^k) - \phi'(x^{k-1}))} =$$

$$= \frac{(\phi'(x^k), \phi'(x^k) - \phi'(x^{k-1}))}{\|\phi'(x^{k-1})\|^2} \tag{2.1.11}$$

далі отримаємо:

$$\gamma_2^{k-2} = \frac{(\phi'(x^k), As^{k-2})}{(s^{k-2}, As^{k-2})} = \frac{(\phi'(x^k), \phi'(x^{k-1}) - \phi'(x^{k-2}))}{\|\phi'(x^{k-2})\|^2} \tag{2.1.12}$$

$$\gamma_3^{k-3} = \frac{(\phi'(x^k), As^{k-3})}{(s^{k-3}, As^{k-3})} = \frac{(\phi'(x^k), \phi'(x^{k-2}) - \phi'(x^{k-3}))}{\|\phi'(x^{k-3})\|^2} \tag{2.1.13}$$

Отже, для неквадратичних функцій, чотирьохкроковий метод має вигляд (2.0.1), (2.1.11), (2.1.12), (2.1.13)

2.2. Обчислювальний експеримент

Зробимо аналіз роботи описаного чотирикрокового методу порівняно із трикроковим методом. Для порівняння використаємо такі показники, як точність отриманного розв'язку та кількість проведених ітерацій. Під кількістю ітерацій будемо розуміти кількість отриманих послідовних наближень до розв'язку задачі. Зауважимо, що кількість обчислень значень функції та її градієнту у чотирикроковому методі не змінюється порівняно із трикроковим.

Для розрахунків використовувалася мова програмування Python. Обчислення проводилися із точністю $\varepsilon = 10^{-6}, 10^{-8}$. Критерій зупинки процесу обчислень полягав у одночасному виконанні трьох умов:

$$\phi(x^{k-1}) - \phi(x^k) < \varepsilon(1 + |\phi(x^k)|),$$

$$\|x^{k-1} - x^k\| < \sqrt{\varepsilon}(1 + \|x^k\|),$$

$$\|\phi'(x^k)\| \leq \sqrt[3]{\varepsilon}(1 + |\phi(x^k)|)$$

Для більш ефективного порівняння результати обчислень для кожної початкової точки зведемо у таблиці та проілюструємо гравічно. На графіках вісь абсцис показує кількість ітерацій, вісь ординат - логарифм відхилення значення цільової функції у i-й точці від оптимального. Пунктирною лінією позначений процес мінімізіції трикроковим методом, суцільною - чотирикроковим.

Задача 2.1.

$$\phi(x) = (-x_1 + 1)^2 + 100\left(-x_1^2 + x_2\right)^2$$

Точний розв’язок задачі: x* = [1, 1] f* = 0

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв’язок	$f(x^*)$	Кількість ітерацій
10^{-6}	[-1.2, 1]	24.2	4 кроковий	[1.00027 1.00053]	0.0	93
			3 кроковий	[1.00069 1.00137]	0.0	29
	[1.0, -1.2]	484.0	4 кроковий	[0.99995 0.99991]	0.0	81
			3 кроковий	[0.99994 0.99988]	0.0	21
	[0.0, 0.0]	1.0	4 кроковий	[1.00015 1.00029]	0.0	42
			3 кроковий	[1.00014 1.00029]	0.0	24
	[-1.0, -1.0]	404.0	4 кроковий	[0.99998 0.99995]	0.0	59
			3 кроковий	[0.99992 0.99985]	0.0	25

Табл. 2.1

a)

b)

c)

d)

Рис. 2.1. Correlation signal peaks: a) numerical experiment, b) registered correlation signals, c) intensity distribution of correlation signals in numerical experiment, d) correlation signals intensity distribution for DCRAW processed data.

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
10^{-8}	[-1.2, 1]	24.2	4 кроковий	[0.99999 0.99999]	0.0	98
			3 кроковий	[0.99997 0.99994]	0.0	35
	[1.0, -1.2]	484.0	4 кроковий	[1. 0.99999]	0.0	32
			3 кроковий	[1. 1.]	0.0	25
	[0.0, 0.0]	1.0	4 кроковий	[0.99993 0.99985]	0.0	24
			3 кроковий	[1. 1.]	0.0	19
	[-1.0, -1.0]	404.0	4 кроковий	[0.99998 0.99996]	0.0	20
			3 кроковий	[1. 1.]	0.0	16

Табл. 2.2

Задача 2.2.

$$\phi(x) = (x_1 + 10x_2)^2 + 10\,(x_1 - x_4)^4 + (x_2 - 2x_3)^4 + 5\,(x_3 - x_4)^2$$

Точний розв’язок задачі: x* = [0, 0, 0, 0] f* = 0

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
10^{-6}	[3, -1, 0, 1]	215	4 кроковий	[0.04999 -0.00496 0.03379 0.03432]	3e-05	36
			3 кроковий	[0.0008 -0.00008 -0.00348 -0.00348]	0.0	36
	[1, 1, 1, 1]	122	4 кроковий	[-0.06942 0.00694 -0.04062 -0.04114]	7e-05	43
			3 кроковий	[-0.04808 0.00482 -0.02532 -0.02525]	1e-05	21
	[-1.0, 1.0, -1.0, 1.0]	342.0	4 кроковий	[-0.08631 0.00862 -0.04842 -0.04857]	0.00014	38
			3 кроковий	[0.00402 -0.00041 -0.01603 -0.01611]	0.0	29
	[0.0, 2.0, -1.0, 1.0]	686.0	4 кроковий	[0.00241 -0.00024 0.00156 0.00154]	0.0	44
			3 кроковий	[-0.03611 0.00363 -0.00961 -0.00963]	1e-05	27

Табл. 2.3

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв’язок	$f(x^*)$	Кількість ітерацій
10^{-8}	[3, -1, 0, 1]	215	4 кроковий	[-0.02558 0.00256 -0.01165 -0.01175]	0.0	56
			3 кроковий	[0.00019 -0.00002 -0.00241 -0.00241]	0.0	34
	[1, 1, 1, 1]	122	4 кроковий	[0.00081 -0.00008 -0.00795 -0.00794]	0.0	106
			3 кроковий	[-0.02441 0.00244 -0.01503 -0.01505]	0.0	37
	[-1.0, 1.0, -1.0, 1.0]	342.0	4 кроковий	[0.00905 -0.0009 -0.00682 -0.00667]	0.0	83
			3 кроковий	[-0.00833 0.00083 -0.00384 -0.00385]	0.0	44
	[0.0, 2.0, -1.0, 1.0]	686.0	4 кроковий	[0.00071 -0.00007 0.00963 0.00962]	0.0	40
			3 кроковий	[0.00041 -0.00004 -0.00281 -0.00281]	0.0	62

Табл. 2.4

Задача 2.3.

$$\phi(x) = \left(x_1 + x_2^2 - 7\right)^2 + \left(x_1^2 + x_2 - 11\right)^2$$

Точний розв’язок задачі: x* = [3, 2] f* = 0

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв’язок	$f(x^*)$	Кількість ітерацій
10^{-6}	[1, 1]	106	4 кроковий	[3. 2.]	0.0	11
			3 кроковий	[3. 2.]	0.0	9
	[1.0, 4.0]	136.0	4 кроковий	[3. 2.]	0.0	17
			3 кроковий	[3.00011 2.00001]	0.0	9
	[0.0, 0.0]	170.0	4 кроковий	[2.99991 1.99991]	0.0	30
			3 кроковий	[2.99999 2.00016]	0.0	19
	[2.5, 2.5]	8.125	4 кроковий	[3. 2.]	0.0	13
			3 кроковий	[2.99998 2.00004]	0.0	7

Табл. 2.5

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
10^{-8}	[1, 1]	106	4 кроковий	[3. 2.]	0.0	12
			3 кроковий	[3. 2.]	0.0	10
	[1.0, 4.0]	136.0	4 кроковий	[3. 2.]	0.0	18
			3 кроковий	[2.99998 2.00003]	0.0	12
	[0.0, 0.0]	170.0	4 кроковий	[2.99998 2.]	0.0	35
			3 кроковий	[2.99999 1.99998]	0.0	23
	[2.5, 2.5]	8.125	4 кроковий	[3. 2.]	0.0	14
			3 кроковий	[3.00001 2.00001]	0.0	9

Табл. 2.6

Задача 2.4.

$$\phi(x) = \left(x_1^2 + 12x_2 - 1\right)^2 + \left(49x_1^2 + 84x_1 + 49x_2^2 + 2324x_2 - 681\right)^2$$

Точний розв’язок задачі: x* = [0.28581, 0.27936] f* = 5.9225

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
10^{-6}	[1, 1]	3330769	4 кроковий	[0.28582 0.27933]	5.92256	88
			3 кроковий	[0.28582 0.27933]	5.92256	26
	[0.0, 0.0]	463762.0	4 кроковий	[0.28583 0.27933]	5.92256	70
			3 кроковий	[0.2858 0.27933]	5.92256	22
	[-5.0, -7.0]	188873649.0	4 кроковий	[0.28581 0.27933]	5.92256	91
			3 кроковий	[0.28581 0.27933]	5.92256	69
	[0.2, 0.3]	1556.9665	4 кроковий	[0.28583 0.27933]	5.92256	41
			3 кроковий	[0.28587 0.27932]	5.92256	21

Табл. 2.7

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
10^{-8}	[1, 1]	3330769	4 кроковий	[0.28581 0.27933]	5.92256	91
			3 кроковий	[0.28582 0.27933]	5.92256	27
	[0.0, 0.0]	463762.0	4 кроковий	[0.28581 0.27933]	5.92256	73
			3 кроковий	[0.28582 0.27933]	5.92256	25
	[-5.0, -7.0]	188873649.0	4 кроковий	[0.28581 0.27933]	5.92256	152
			3 кроковий	[0.28581 0.27933]	5.92256	64
	[0.2, 0.3]	1556.9665	4 кроковий	[0.28581 0.27933]	5.92256	15
			3 кроковий	[0.28581 0.27933]	5.92256	12

Табл. 2.8

Задача 2.5.

$$\phi(x) = (-x_1 + 1)^2 + (-x_2 + 1)^2 + 100\left(x_3 - \left(\frac{x_1}{2} + \frac{x_2}{2}\right)^2\right)^2$$

Точний розв’язок задачі: x* = [1, 1, 1] f* = 0

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
10^{-6}	[-1.2, 2, 0]	8.4	4 кроковий	[0.99756 0.99819 0.99575]	1e-05	67
			3 кроковий	[1.00002 0.99978 0.99979]	0.0	31
	[0.0, 0.0, 0.0]	2.0	4 кроковий	[1.00009 1.00009 1.00018]	0.0	51
			3 кроковий	[1.00047 1.00047 1.00091]	0.0	23
	[0.0, 1.0, -1.2]	211.25	4 кроковий	[0.99745 1.00094 0.99839]	1e-05	66
			3 кроковий	[1.00443 0.99998 1.00444]	2e-05	53
	[2.3, 1.0, -0.3]	915.24062	4 кроковий	[1.0009 0.9995 1.00042]	0.0	51
			3 кроковий	[1.00005 0.99997 1.00002]	0.0	24

Табл. 2.9

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
10^{-8}	[-1.2, 2, 0]	8.4	4 кроковий	[1.00002 0.99975 0.99977]	0.0	96
			3 кроковий	[1. 1. 1.]	0.0	33
	[0.0, 0.0, 0.0]	2.0	4 кроковий	[1. 1. 1.]	0.0	53
			3 кроковий	[1.0004 1.0004 1.0008]	0.0	26
	[0.0, 1.0, -1.2]	211.25	4 кроковий	[1.00002 0.99996 0.99998]	0.0	75
			3 кроковий	[1. 1. 1.]	0.0	57
	[2.3, 1.0, -0.3]	915.24062	4 кроковий	[1.00093 1.00004 1.00097]	0.0	63
			3 кроковий	[1.00004 0.99996 1.]	0.0	25

Табл. 2.10

Задача 2.6.

$$\phi(x) = 1352.99688810716 x_1^2 - 70000.0 + \frac{319.284802043423 x_1^2 + 1.67624521072797 x_2^2 +}{0.00525 x_4^2 + 1}$$

Точний розв’язок задачі: x* = [0, 0, 0, 1] f* = -70000

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок
10^{-6}	[2.7, 90, 1500, 10]	28580.05033	4 кроковий	[0.13676 -15.05046 1495.27438 331.41988]
			3 кроковий	[-0.00089 25.86293 1488.12225 473.17638]
	[2.0, 140.0, 1707.0, 31.0]	-8899.54441	4 кроковий	[0.00941 -28.76126 1702.29196 688.04944]
			3 кроковий	[0.00421 -35.77071 1702.12319 709.41251]
	[1.0, 1.0, 1.0, 1.0]	-65340.91918	4 кроковий	[0. 0.00001 0.00561 1.00689]
			3 кроковий	[-0. 0.00001 0.0015 1.00691]
	[-1.0, 0.0, 1.0, -1.0]	-65346.95247	4 кроковий	[-0.00001 0. 0.00001 -1.00519]
			3 кроковий	[0. 0. 0. -1.00519]

Табл. 2.11

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок
10^{-8}	[2.7, 90, 1500, 10]	28580.05033	4 кроковий	[-0.00743 4.95678 1492.21697 388.69679]
			3 кроковий	[-0.00011 -8.10017 1403.31712 696.40143]
	[2.0, 140.0, 1707.0, 31.0]	-8899.54441	4 кроковий	[-0.00397 -32.08949 1702.15048 590.31179]
			3 кроковий	[-0.00018 -3.70682 1679.4614 1104.81388]
	[1.0, 1.0, 1.0, 1.0]	-65340.91918	4 кроковий	[0. 0.00001 0.00561 1.00689]
			3 кроковий	[-0. 0.00001 0.0015 1.00691]
	[-1.0, 0.0, 1.0, -1.0]	-65346.95247	4 кроковий	[-0.00001 0. 0.00001 -1.00519]
			3 кроковий	[0. 0. 0. -1.00519]

Табл. 2.12

Задача 2.7.

$$\phi(x) =$$

Точний розв'язок задачі:

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
height						

Табл. 2.13

10^{-6}

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
height						

Табл. 2.14

10^{-8}

Задача 2.8.

$$\phi(x) =$$

Точний розв'язок задачі:

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
height						

Табл. 2.15

10^{-6}

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
height						

Табл. 2.16

10^{-8}

Задача 2.9.

$$\phi(x) =$$

Точний розв'язок задачі:

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
height						

Табл. 2.17

10^{-6}

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
height						

Табл. 2.18

10^{-8}

Задача 2.10.

$$\phi(x) =$$

Точний розв'язок задачі:

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
height						

Табл. 2.19

10^{-6}

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
height						

Табл. 2.20

10^{-8}

Задача 2.11.

$$\phi(x) =$$

Точний розв’язок задачі:

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв’язок	$f(x^*)$	Кількість ітерацій
height						

Табл. 2.21

10^{-6}

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
height						

Табл. 2.22

10^{-8}

Задача 2.12.

$$\phi(x) =$$

Точний розв'язок задачі:

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
height						

Табл. 2.23

10^{-6}

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв’язок	$f(x^*)$	Кількість ітерацій
height						

Табл. 2.24

10^{-8}

Задача 2.13.

$$\phi(x) =$$

Точний розв'язок задачі:

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
height						

Табл. 2.25

10^{-6}

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
height						

Табл. 2.26

10^{-8}

Задача 2.14.

$$\phi(x) =$$

Точний розв'язок задачі:

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
height						

Табл. 2.27

10^{-6}

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
height						

Табл. 2.28

10^{-8}

ВИСНОВКИ

У роботі розглянуто чотирьохкроковий метод для задач безумовної мінімізації. Також обгрунтована його належність до методів спряжених напрямків і проведено обчислювальні експерименти для тестових функцій.

Результати експериментів вказують на неефективніть запропонованого методу у порівнянні з трикроковим методом. Трикроковий алгоритм дає розв’язок при суттєво меншій кількості кроків.

Слід зазначити, що кількість обчислень функції та градієнту в обох методах однакова.

Зважаючи на отримані результати, використання запропонованого методу не доцільне.

СПИСОК ЛІТЕРАТУРИ

1. Абрамов О. В. Параметрический синтез стохастических систем с учетом требований надежности / О. В. Абрамов. — М.: Наука, 1992. — 175 с.
2. Алберт А. Регрессия, псевдоинверсия и рекуррентное оценивание / А. Алберт. — М.: Наука, 1977. — 224 с.
3. Афанасьев В. Н. Математическая теория конструирования систем управления / Афанасьев Б. Н., Колмановский В. Б., Носов В. Р. — М.: Высшая школа, 2003. — 614 с.
4. Пантаев М. Ю. Матанализ с человеческим лицом, или Как выжить после предельного перехода: Полный курс математического анализа. Т. 1. / М. Ю. Пантаев. — М.: ЛЕНАНД, 2015. — 368 с.
5. Пантаев М. Ю. Матанализ с человеческим лицом, или Как выжить после предельного перехода: Полный курс математического анализа. Т. 2. / М. Ю. Пантаев. — М.: Книжный дом «ЛИБРОКОМ», 2015. — 416 с.