

МЕТОДИЧНІ ВКАЗІВКИ
до виконання лабораторних робіт
«МЕТОДИ МІНІМІЗАЦІЇ ФУНКЦІЙ БАГАТЬОХ ЗМІННИХ»
з курсу «Методи оптимізації»
для студентів денної та заочної форми навчання
спеціальності 122 «Комп'ютерні науки»

Затверджено
редакційно-видавничою
радою університету,
протокол № 2 від 23.06.2023 р.

Харків
НТУ «ХП»
2023

Методичні вказівки до виконання лабораторних робіт «Методи мінімізації функцій багатьох змінних» з курсу «Методи оптимізації» для студентів денної та заочної форм навчання спеціальності 122 «Комп'ютерні науки» / уклад. Ю. А. Плаксій. – Харків: НТУ «ХПІ». – 54 с.

Укладач Ю. А. Плаксій

Рецензент В. Б. Успенський

Кафедра комп'ютерного моделювання процесів і систем

ЗМІСТ

1. МЕТОДИ МІНІМІЗАЦІЇ ФУНКЦІЙ БАГАТЬОХ ЗМІННИХ	5
1.1. Постановка задачі оптимізації. Основні поняття, припущення і визначення.	5
1.2. Вимоги до виконання лабораторних робіт	11
1.3. Порядок виконання лабораторних робіт.	11
1.4. Вимоги до представлення результатів лабораторних робіт	12
<i>Контрольні питання.</i>	12
2. МЕТОДИ БЕЗУМОВНОЇ МІНІМІЗАЦІЇ ФУНКЦІЙ БАГАТЬОХ ЗМІННИХ	14
2.1. Градієнтний метод з різними метриками	14
2.2. Лабораторна робота "Мінімізація функції двох змінних градієнтним методом з різними метриками"	17
<i>Контрольні питання.</i>	18
2.3. Метод спряжених градієнтів Флетчера-Рівса	20
2.4. Лабораторна робота "Мінімізація функції двох змінних методом Флетчера-Рівса"	23
<i>Контрольні питання.</i>	23
2.5. Метод Пауелла	25
2.6. Лабораторна робота "Мінімізація функції двох змінних методом Пауелла"	27
<i>Контрольні питання.</i>	27
2.7. Метод Ньютона-Рафсона	29
2.8. Лабораторна робота "Мінімізація функції двох змінних методом Ньютона-Рафсона"	31
<i>Контрольні питання.</i>	31
2.9. Метод Хука-Дживса.	32
2.10. Лабораторна робота "Мінімізація функції двох змінних методом Хука-Дживса"	33
<i>Контрольні питання.</i>	34
2.11. Метод деформованого багатогранника.	35
2.12. Лабораторна робота "Мінімізація функції двох змінних методом деформованого багатогранника"	39
<i>Контрольні питання.</i>	39
3. МЕТОДИ УМОВНОЇ МІНІМІЗАЦІЇ ФУНКЦІЙ БАГАТЬОХ ЗМІННИХ	40
3.1. Метод зовнішніх штрафних функцій.	40
3.2. Лабораторна робота "Мінімізація функції двох змінних методом зовнішніх штрафних функцій"	41
<i>Контрольні питання.</i>	42
3.3 Метод проекції градієнта	43

3.4. Лабораторна робота "Мінімізація функції двох змінних методом проекції градієнта Розена"	46
<i>Контрольні питання</i>	46
3.5. Метод можливих (допустимих) напрямків Зойтендейка ..	47
3.6. Лабораторна робота "Мінімізація функції двох змінних методом допустимих напрямків Зойтендейка "	49
<i>Контрольні питання</i>	49
СПИСОК ЛІТЕРАТУРИ	50
Додаток А. ЦІЛЬОВІ ФУНКЦІЇ І ОБМЕЖЕННЯ ..	51
Додаток Б. ПРИКЛАД ОФОРМЛЕННЯ ТИТУЛЬНОЇ СТОРІНКИ ЗВІТУ	53

ПОЗНАЧЕННЯ

$f(x)$ - скалярна функція векторного аргументу x (цільова функція);

$\arg f(x)$ - аргумент функції $f(x)$;

$x = (x_1, x_2, \dots, x_n)^T$ - вектор-стовпчик змінних;

$x = (x_1, x_2, \dots, x_n)^T$ - точка мінімуму цільової функції;

$f(x)$ - мінімальне значення цільової функції;

$\frac{\partial f(x)}{\partial x_k}$ - частинна похідна цільової функції по змінній x_k ;

R^n - n -мірний простір арифметичних векторів з дійсними координатами;

$U_\delta(x_k)$ - δ -окіл точки x_k ($x: \|x - x_k\| \leq \delta$);

$f'(x)$ - градієнт функції $f(x)$;

f'_k - градієнт функції $f(x)$ в точці x_k ;

$H(x)$ - матриця других похідних (матриця Гессе) функції $f(x)$;

H_k - матриця Гессе функції $f(x)$ в точці x_k ;

E - одинична матриця;

σ - точність знаходження точки мінімуму x ;

$\|x_k\|$ - евклідова норма вектора x_k ;

$\|p_k\|_q = \left(\sum_{i=1}^n |p_{ki}|^q \right)^{\frac{1}{q}}$, $q \in (1, 2, \infty)$ - q -норма вектора $p = (p_1, p_2, \dots, p_n)$ в просторі R^n ;

$p_k, p_n > 0$ - скалярний добуток векторів p_k і p_n ;

$[a, b]$ - відрізок;

(a, b) - інтервал;

$a \in A$ - елемент a належить множині A ;

$A \subset B$ - множина A міститься в множині B ;

$\{x_k\}_0^n$ - послідовність елементів x_k , $k=0, 1, \dots, n$;

$f(x) = o(g(x))$ - функція $f(x)$ є нескінченно малою в порівнянні з функцією $g(x)$;

ПК - персональний комп'ютер.

1. МЕТОДИ МІНІМІЗАЦІЇ ФУНКЦІЙ БАГАТЬОХ ЗМІННИХ

1.1. Постановка задачі оптимізації. Основні поняття, припущення і визначення

Задача оптимізації - це задача пошуку екстремуму функції або функціонала на заданій множині допустимих розв'язків.

Нехай задана неперервна скалярна функція $f(x)$ векторного аргументу $x=(x_1, x_2, \dots, x_n)$, де $x \in X$, X - деяка множина з простору R^n , $X \subset R^n$. Екстремум може бути мінімумом або максимумом, але в подальшому будемо розглядати тільки задачі мінімізації функції, оскільки пошук максимуму функції $f(x)$ зводиться до пошуку мінімуму функції $-f(x)$, бо

$$\max f(x) = \min(-f(x)).$$

Функція $f(x)$, для якої розглядається задача мінімізації, називається *цільовою функцією*.

Якщо допустима множина розв'язків $X=R^n$, то задача

$$\min f(x) = f(x), \quad x \in R^n$$

називається *задачею безумовної мінімізації*. Якщо допустима множина розв'язків є підмножиною простору R^n : $X \subset R^n$, то задача

$$\min f(x) = f(x), \quad x \in X$$

називається *задачею умовної мінімізації* або задачею мінімізації в умовах обмежень.

Точка мінімуму цільової функції визначається як

$$x = \arg f(x).$$

❖ *Визначення.*

Неперервна функція $f(x)$ називається *опуклою*, якщо

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2)$$

для будь-якого $x_1, x_2 \in R^n$ і будь-якого $\lambda \in (0,1)$.

В теорії чисельних методів оптимізації функцій багатьох змінних і в відповідних алгоритмах пошуку доволі часто використовуються поняття градієнту $f'(x)$ цільової функції і матриці других похідних $H(x)$ цільової функції.

Вектор градієнту складається з частинних похідних цільової функції по змінним x_1, x_2, \dots, x_n :

$$f'(x) = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \dots, \frac{\partial f(x)}{\partial x_n} \right).$$

Матриця других похідних (матриця Гессе) має вигляд:

$$H(x) = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_n \partial x_n} \end{pmatrix}.$$

Матриця Гессе є симетричною матрицею, елементи матриці $H(x)$ відносно головної діагоналі є однаковими: $h_{ij} = h_{ji}$.

Серед функцій багатьох змінних особливе місце займають *квадратичні функції*, які є найбільш досліджені і для яких розроблені спеціальні методи відшукування мінімуму. Квадратична функція n змінних може бути представлена в стандартному вигляді

$$f(x) = a + \langle b, x \rangle + \frac{1}{2} \langle Hx, x \rangle, \quad (1.1)$$

де a - скаляр, b - вектор коефіцієнтів при змінних x_1, x_2, \dots, x_n , H - матриця Гессе. Градієнт квадратичної функції визначається за формулою:

$$f'(x) = b + Hx. \quad (1.2)$$

❖ *Визначення.*

Точка x^* називається точкою *локального мінімуму* функції $f(x)$, якщо існує таке $\delta > 0$, що має місце нерівність $f(x) \leq f(x)$ для всіх $x \in U_\delta(x)$, тобто для всіх x , які задовольняють умові $\|x - x^*\| < \delta$.

Точка x^* називається точкою *строого локального мінімуму* функції $f(x)$, якщо існує таке $\delta > 0$, що має місце нерівність $f(x) < f(x)$ для $x \in U_\delta(x)$, тобто для всіх x , які задовольняють умові $\|x - x^*\| < \delta$.

Точка x^* називається точкою *глобального мінімуму* функції $f(x)$, якщо має місце нерівність $f(x) \leq f(x)$ для всіх $x \in R^n$. В цьому випадку функція $f(x)$ може мати декілька точок мінімуму x_1, x_2, \dots, x_k з однаковим значенням цільової функції $f(x_1) = f(x_2) = \dots = f(x_k)$.

Точка x^* називається точкою *строого глобального мінімуму* функції $f(x)$, якщо має місце нерівність $f(x) < f(x)$ для всіх $x \in R^n$.

При дослідженні оптимізаційних задач завжди виникають два питання.

1. *Питання аналізу «в статистиці».* Як визначити, чи є деяка точка x оптимальним розв'язком задачі мінімізації?

❖ *Критерії оптимальності.*

❖ *Умови оптимальності першого порядку.*

Твердження 1. Якщо функція $f(x)$ є опуклою, то вона в будь-якій точці локального мінімуму досягає свого найменшого значення.

Твердження 2. Якщо диференційована функція $f(x)$ має в точці x екстремум (максимум або мінімум), то $f'(x)=0$.

Твердження 3. Якщо функція $f(x)$ опукла і диференційована в точці x і $f'(x)=0$, то x - точка глобального мінімуму.

❖ *Умови оптимальності другого порядку.*

Твердження 1. Якщо функція $f(x)$ двічі диференційована і в точці x градієнт $f'(x)=0$ і матриця Гессе є невід'ємно визначеною: $H(x) \geq 0$, то точка x є точкою локального мінімуму.

Твердження 2. Якщо функція $f(x)$ двічі диференційована і в точці x градієнт $f'(x)=0$ і матриця Гессе є позитивно визначеною: $f''(x) > 0$, то точка x є точкою строгого локального мінімуму.

2. *Питання аналізу «в динаміці».* Якщо x не є точкою оптимуму, то виникає питання: яка послідовність дій призводить до отримання оптимального розв'язку? Аналіз «в динаміці» завжди пов'язаний з побудуванням за деяким правилом послідовності точок пошуку x_0, x_1, \dots, x_n , яка за певних умов збігається до шуканої точки мінімуму.

❖ **Визначення.**

Послідовність точок x_0, x_1, \dots, x_n називається *релаксаційною*, якщо для кожної точки x_k виконується умова $f(x_k) \leq f(x_{k-1})$. Якщо мінімум функції $f(x)$ існує, то релаксаційна послідовність збігається до точки мінімуму x .

Методи пошуку мінімуму функції $f(x)$, які засновані на побудуванні релаксаційної послідовності точок x_0, x_1, \dots, x_n , називаються *релаксаційними методами*.

Основною задачею мінімізації є побудова релаксаційної послідовності і знаходження точки мінімуму x із заданою точністю ϵ .

Лінії рівня цільової функції визначаються рівнянням

$$f(x) = C_k,$$

де числа C_k задаються формулою $C_k = C_0 + k \Delta C_k$. Зазвичай числа C_0 і ΔC_k задаються в залежності від вигляду рівняння поверхні цільової функції. Лінії рівня - це лінії перетину (замкнуті або незамкнуті) поверхні $z=f(x)$ площиною $z=C_k$.

Найчастіше лінії рівня використовуються для візуалізації процесу пошуку і дослідження методів мінімізації функцій двох змінних. Для цього вони проектуються на площину двох змінних з проставлянням відповідних проекцій

на цю площину точок пошуку. Приклад побудування ліній рівня для функції двох змінних наведений на рис.1.1.

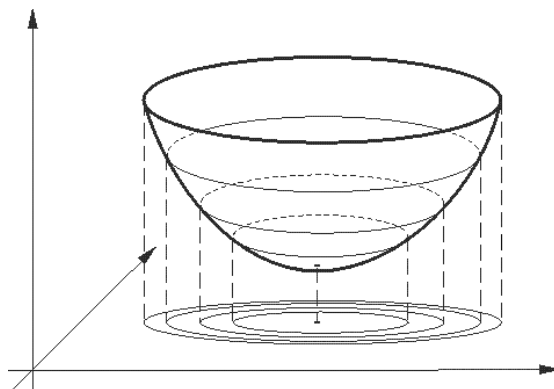


Рисунок 1.1 – Лінії рівня квадратичної функції двох змінних

❖ **Загальна структура релаксаційних методів пошуку.**

Релаксаційна послідовність будується, починаючи з деякої початкової точки x_0 , за формулою

$$x_{k+1} = x_k + \alpha_k p_k, \quad (1.3)$$

де p_k – напрямок пошуку в точці x_k , α_k – величина кроку в напрямку p_k . Для цього на кожному кроці обчислень треба знати вектори напрямків пошуку p_0, p_1, \dots, p_n і величини кроків пошуку в цих напрямків $\alpha_0, \alpha_1, \dots, \alpha_n$. Всі релаксаційні методи поділяються в залежності від вибору напрямку p_k і способу визначення величини кроку α_k . Якщо напрямок пошуку p_k визначений, то задача визначення величини кроку α_k – це задача одновимірної мінімізації

$$f(x_k + \alpha p_k) = f(x_k + \alpha_k p_k) = f(x_{k+1}). \quad (1.4)$$

В залежності від способів розв’язку цієї задачі розрізняють *методи найшвидшого спуску* і *методи з дробленням кроку*.

В *методах найшвидшого спуску* для розв’язання задачі (1.4) можна застосувати чисельні методи одновимірної мінімізації, або методи пошуку кореня α_k нелінійного рівняння

$$\frac{d}{d\alpha} f(x_k + \alpha p_k) = 0. \quad (1.5)$$

В *методах з дробленням кроку* на кожній ітерації задається деяке α_{k0} , знаходиться точка

$$x_{k+1}^0 = x_k + \alpha_{k0} p_k \quad (1.6)$$

і перевіряється умова

$$f(x_{k+1}^0) < f(x_k). \quad (1.7)$$

Якщо умова (1.7) виконується, то

$$x_{k+1} = x_{k+1}^0,$$

інакше треба замість α_{k0} взяти α_{k0}^r , де $r \in (0, 1)$ і повторити процедуру (1.6) - (1.7).

➤ **Зауваження.**

Якщо в ході дроблення кроку α_k отримано, що поточне значення кроку $\alpha_{k0} \leq \sigma$, де σ - задана точність в задачі мінімізації, то це означає, що можна покласти $x = x_k$, тобто процес пошуку мінімуму завершено в точці x_k .

Розглянемо формування релаксаційної послідовності для мінімізації квадратичної функції $f(x) = a + \langle b, x \rangle + \frac{1}{2} \langle Hx, x \rangle$, де H - матриця Гессе. В цьому випадку величина кроку α_k визначається на основі розв'язку задачі одновимірної мінімізації:

$$f(x_k + \alpha p_k) = f(x_k + \alpha_k p_k) = f(x_{k+1}).$$

В результаті отримаємо, що:

$$\alpha_k = - \frac{f'_k p_k}{\langle p_k, p_k \rangle} \quad (1.8)$$

Таким чином, для квадратичної цільової функції $f(x)$ для знаходження величини кроку на кожній ітерації замість розв'язання задачі одновимірної мінімізації (1.4) можна використовувати формулу (1.8).

❖ **Поняття про швидкість збіжності методу.**

Швидкість збіжності методу - це відношення норми $\|x_{k+1} - x\|$ до норми $\|x_k - x\|$.

❖ **Визначення.**

Якщо $\|x_{k+1} - x\| \leq \rho \|x_k - x\|$, де $\rho \in (0, 1)$, то швидкість збіжності *лінійна*. Це відповідає швидкості спадання геометричної прогресії із знаменником ρ .

В цьому випадку маємо:

$$\begin{aligned} \|x_1 - x\| &\leq \rho \|x_0 - x\|, \\ \|x_2 - x\| &\leq \rho \|x_1 - x\| \leq \rho^2 \|x_0 - x\|, \\ \|x_3 - x\| &\leq \rho \|x_2 - x\| \leq \rho^2 \|x_1 - x\| \leq \rho^3 \|x_0 - x\|, \dots, \\ \|x_{k+1} - x\| &\leq \rho^{k+1} \|x_0 - x\|. \end{aligned} \quad (1.9)$$

Оскільки $\|x_{k+1} - x\|$ інтерпретується як відстань між точкою x_{k+1} і точкою мінімуму x , то умова (1.9) означає, що ця відстань в порівнянні між відстанню початкової точки пошуку x_0 до шуканої точки мінімуму x зменшилася в ρ^{k+1} разів.

Якщо $\|x_{k+1} - x\| \leq \rho_k \|x_k - x\|$, де чисельна послідовність $\{\rho_k\}_0^n$ збігається до нуля: $\{\rho_k\} \rightarrow 0$, $\rho_k \in (0, 1)$, то має місце *надлінійна* швидкість збіжності.

Якщо $\|x_{k+1} - x\| \leq C \|x_k - x\|^2$, то відповідний метод мінімізації має *квадратичну швидкість збіжності*.

Введемо в розгляд *q-норму вектора* ρ в просторі R^n , інтерпретуючи її як відстань вектора ρ до "нульового" вектора $\rho_0 = (0, 0, \dots, 0)$. Представимо *q-норму* вектора ρ у вигляді:

$$\|\rho\|_q = \left(\sum_{i=1}^n |\rho_i|^q \right)^{\frac{1}{q}}, \text{ де } q \in (1, 2, \infty). \quad (1.10)$$

В формулі (1.10) величина q може приймати тільки 3 значення: $q=1$, $q=2$, $q \rightarrow \infty$.

Розглянемо більш детально ці норми для випадку двомірного простору ($\rho_k \in R^2$) і отримаємо рівняння відповідних одиничних сфер в просторі R^2 . Рівняння одиничних сфер мають вигляд:

$$\|\rho\|_q = 1, \quad q \in (1, 2, \infty). \quad (1.11)$$

Для випадку $q=1$ з формули (1.10) в умовах (1.11) маємо наступне рівняння одиничної сфери:

$$\|\rho\|_1 = \sum_{i=1}^2 |\rho_i| = |\rho_1| + |\rho_2| = 1. \quad (1.12)$$

Графічне зображення одиничної сфери в просторі R^2 у випадку норми $\|\rho\|_1$ представлено на рис.1.2 (позначене як $\|\cdot\|_1$). Така норма називається *октаедричною*.

Випадок $q=2$ відповідає Евклідовій нормі. Для цього випадку з формули (1.10) отримаємо, що рівняння одиничної сфери має вигляд:

$$\|\rho\|_2 = \left(\sum_{i=1}^2 |\rho_i|^2 \right)^{\frac{1}{2}} = (\rho_1^2 + \rho_2^2)^{\frac{1}{2}} = 1,$$

тобто

$$\rho_1^2 + \rho_2^2 = 1. \quad (1.13)$$

Зображення одиничної сфери у випадку норми $\|\rho\|_2$ представлено на рис.1.2 (позначене як $\|\cdot\|_2$). Це - двомірна сфера, отже, така норма називається *сферичною*.

У випадку $q \rightarrow \infty$ для одиничної сфери маємо рівняння:

$$\left(\sum_{i=1}^2 |\rho_i|^q \right)^{\frac{1}{q}} = 1.$$

Введемо $|\rho_i|$, ($i=1,2$). Тоді рівняння одиничної сфери можна переписати у вигляді

$$\left(\left| \frac{\rho_1}{M} \right|^q + \left| \frac{\rho_2}{M} \right|^q \right)^{\frac{1}{q}} = 1.$$

Оскільки при $q \rightarrow \infty$

$$\left| \frac{\rho_i}{M} \right|^q = \begin{cases} 1, & \text{якщо } \rho_i = M \\ 0, & \text{якщо } \rho_i < M \end{cases},$$

то рівняння одиничної сфери в \mathbb{R}^2 у випадку $q \rightarrow \infty$ перетворюється до вигляду:

$$\max(\rho_1, \rho_2) = 1. \quad (1.14)$$

На рис.1.2 представлено зображення одиничної сфери у випадку $\|\cdot\|_\infty$ (позначене як $\|\cdot\|_\infty$). У зв'язку з зображенням одиничної сфери у вигляді двомірного куба ця норма називається *кубічною*.

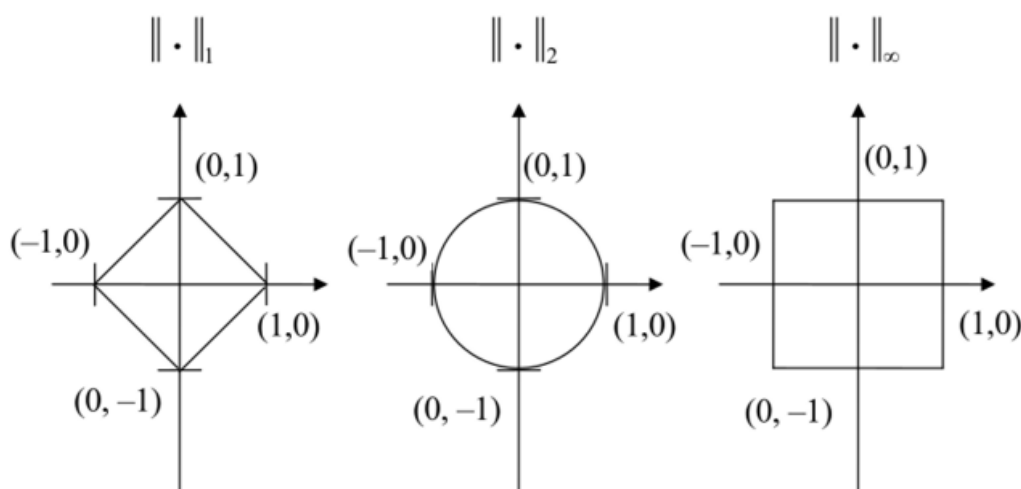


Рисунок 1.2 – Вигляд одиничної сфери в \mathbb{R}^2 для різних норм

1.2. Вимоги до виконання лабораторних робіт

Мета виконання лабораторних робіт - вивчення низки методів безумовної і умовної мінімізації функцій багатьох змінних та отримання навичок практичної роботи з ними на прикладі чисельного розв'язання задачі мінімізації цільової функції двох змінних на ПК.

1.3. Порядок виконання лабораторних робіт

Для виконання лабораторних робіт необхідно:

- 1) отримати від викладача завдання на лабораторні роботи і дані з таблиць A1 і A2 Додатку А;
- 2) ознайомитись з необхідним теоретичним матеріалом для виконання лабораторної роботи з конкретного методу мінімізації;
- 2) ознайомитись з алгоритмом методу і скласти блок-схему алгоритму;

3) самостійно скласти комп'ютерну програму мовою програмування високого рівня (C++, C#, Python, Java, ...) для реалізації алгоритму на ПК та провести необхідні обчислення;

4) отримати точку мінімуму x і мінімальне значення цільової функції $f(x)$;

5) за результатами роботи програми зробити висновки про ефективність методу для мінімізації конкретної цільової функції. Визначити найкращий метод, що характеризується найменшим числом обчислювання цільової функції.

1.4. Вимоги до представлення результатів лабораторних робіт

За результатами виконання лабораторних робіт треба оформити єдиний звіт в електронному вигляді, який включає титульний аркуш, постановку задачі з наведенням початкових даних для кожного алгоритма, навести блок-схеми алгоритмів, представити отримані результати обчислень на кожній ітерації в табличному вигляді, зробити висновки щодо інтерпретації отриманих результатів і ефективності застосування метода.

Контрольні питання

1. Сформулюйте задачу безумовної мінімізації функції багатьох змінних.
2. Сформулюйте задачу умовної мінімізації функції багатьох змінних.
3. Дайте визначення опуклої функції.
4. Запишіть градієнт функції $f(x) = 12x_1^2 + x_2^2 - 4x_1x_2 - 2x_2 + 5$ в точці $x = (-1, -2)$.
5. Запишіть матрицю Гессе функції $f(x) = x_1^2 + 10x_2^2 - 3x_1x_2 - 4x_1 + 2x_2$ в точці $x = (2, -1)$.
6. Запишіть квадратичну функцію $f(x) = 6x_1^2 + x_2^2 + 2x_1x_2 - 2x_1 - x_2$ в стандартному вигляді $f(x) = a + \langle b, x \rangle + \frac{1}{2} \langle Hx, x \rangle$.
7. Запишіть формулу для вектора градієнта квадратичної функції через матрицю Гессе.
8. Дайте визначення точки локального мінімуму функції багатьох змінних.
9. Дайте визначення точки строгого локального мінімуму функції багатьох змінних.

10. Дайте визначення точки глобального мінімуму функції багатьох змінних.
11. Дайте визначення точки строгого глобального мінімуму функції багатьох змінних.
12. Наведіть умови оптимальності першого порядку.
13. Наведіть умови оптимальності другого порядку.
14. Яка послідовність точок пошуку називається релаксаційною?
15. Запишіть рівняння ліній рівня функції $f(x) = 2x_1^2 - 4x_2^2 + 2x_1x_2 + x_1$.
16. Дайте визначення швидкості збіжності методу пошуку.
17. Дайте визначення q -норми вектора P_k .
18. Який вигляд має одинична сфера в просторі R^2 у випадку октаедричної норми?
19. Який вигляд має одинична сфера в просторі R^2 у випадку сферичної норми?
20. Який вигляд має одинична сфера в просторі R^2 у випадку кубічної норми?
21. Запишіть аналітичну формулу для визначення величини кроку у випадку мінімізації квадратичної функції.
22. Опишіть схему дроблення кроку.

2. МЕТОДИ БЕЗУМОВНОЇ МІНІМІЗАЦІЇ ФУНКЦІЙ БАГАТЬОХ ЗМІННИХ

2.1. Градієнтний метод з різними метриками

Розглянемо основи градієнтного методу. Нехай функція $f(x)$ є опуклою в околі $U_\delta(x_k)$ точки x_k і диференційованою n разів. Розкладемо функцію $f(x)$ в околі точки x_k в ряд Тейлора:

$$f(x_k + \alpha \Delta x) = f(x_k) + \alpha \langle f'(x_k), \Delta x \rangle + o(\|\Delta x\|), \quad (2.1)$$

$f'(x_k)$ – градієнт функції $f(x)$ в точці x_k . В умовах малої величини $\|\Delta x\|$ і $\alpha > 0$ розкладення (2.1) можна переписати як наближену рівність:

$$f(x_k + \alpha \Delta x) = f(x_k) + \alpha \langle f'(x_k), \Delta x \rangle. \quad (2.2)$$

Для того щоб в точці $x_{k+1} = x_k + \alpha \Delta x$ значення функції $f(x)$ було меншим за значення в точці x_k , необхідно, щоб виконувалась умова:

$$f(x_k + \alpha \Delta x) - f(x_k) < 0, \quad (2.3)$$

тобто необхідно забезпечити виконання нерівності

$$\langle f'(x_k), \Delta x \rangle < 0 \quad (2.4)$$

Оскільки скалярний добуток $\langle f'(x_k), \Delta x \rangle$ можна представити у вигляді

$$\langle f'(x_k), \Delta x \rangle = \|f'(x_k)\| \cdot \|\Delta x\| \cdot \cos \theta,$$

де θ – кут між векторами $f'(x_k)$ і Δx , то очевидно, що найменше значення $f(x_{k+1})$ забезпечує $\cos \theta = -1$, тобто $\theta = \pi$. Це означає, що вектор Δx має бути протилежним до вектора $f'(x_k)$. Таким чином "найшвидший спуск" має місце, коли

$$\Delta x = -f'(x_k). \quad (2.5)$$

Це означає, що напрямок пошуку p_k в поточній точці x_k співпадає з напрямком антиградієнта в цій точці:

$$p_k = -f'(x_k).$$

Отже, релаксаційна послідовність точок пошуку в градієнтному методі задається формулою

$$x_{k+1} = x_k - \alpha f'_k, \quad (2.6)$$

де позначено $f'_k = f'(x_k)$.

Модифікації градієнтного методу пов'язані зі способом знаходження величини кроку в напрямку антиградієнта.

Градієнтний метод найшвидшого спуску.

Для визначення кроку α на k -й ітерації розглянемо задачу

$$f(x_k - \alpha f'_k) = f(x_k - \alpha f'_k) = f(x_{k+1}).$$

Її розв'язок α_k задовольняє рівнянню

$$\frac{d}{d\alpha} f(x_k - \alpha f'_k) = 0.$$

Виконуючи диференціювання, отримаємо рівняння для отримання α_k

$$f'(x_k - \alpha_k f'_k), f'_k \geq 0, \quad (2.7)$$

Оскільки $x_{k+1} = x_k - \alpha_k f'_k$, то рівняння (2.7) означає, що в градієнтному методі найшвидшого спуску градієнти в сусідніх точках пошуку ортогональні:

$$f'_{k+1}, f'_k \geq 0,$$

тобто пошук мінімуму ведеться в ортогональних напрямках $\rho_k = -f'_k$, $\rho_{k+1} = -f'_{k+1}$ (рис. 2.1).

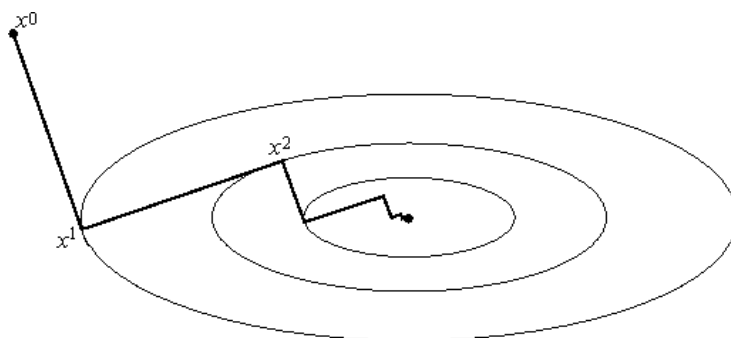


Рисунок 2.1 – Напрямки пошуку в градієнтному методі мінімізації функції двох змінних

Градієнтний метод з дробленням кроку.

Алгоритм градієнтного методу з дробленням кроку на k -й ітерації складається з наступних кроків.

1) Визначити поточний вектор градієнта цільової функції $f'_k = f'(x_k)$.

2) Перевірити умову закінчення пошуку

$$\|f'_k\| \leq \sigma. \quad (2.8)$$

Якщо умова (2.8) виконується, то $x = x_k$.

В іншому випадку треба задати новий напрямок пошуку $\rho_k = -f'_k$ і деяке початкове значення величини кроку $\alpha = \alpha^0$ в напрямку ρ_k .

3) Обчислити *можливу* наступну точку релаксаційної послідовності

$$x_{k+1}^0 = x_k - \alpha f'_k$$

і значення цільової функції в цій точці:

$$f(x_{k+1}^0) = f(x_k - \alpha f'_k).$$

4) Перевірити виконання умови (2.3) з урахуванням (2.2), (2.5) у вигляді:

$$f(x_{k+1}^0) - f(x_k) \leq -\epsilon \alpha \|f'_k\|^2, \quad (2.9)$$

де $\epsilon \in (0, 1)$.

Якщо умова (2.9) виконана, покласти $x_{k+1} = x_{k+1}^0$, $\alpha_k = \alpha$. Обчислити градієнт в точці x_{k+1} :

$$f'_{k+1} = f'(x_k - \alpha_k f'_k).$$

Перейти до пункту 2), змінивши в (2.8) k на $k+1$.

Якщо умова (2.9) не виконана, треба подрібнити крок, поклавши $\alpha = \alpha^0 \cdot r$, де $r \in (0, 1)$ - постійне число, і повторити пункти 1) -4).

➤ **Зауваження.**

1. На початку кожної ітерації початкове значення величини кроку α^0 задається однаковим.

2. Градієнтний метод з дробленням кроку має лінійну швидкість збіжності.

Далі буде показано, що градієнтний метод, заснований на формуванні релаксаційної послідовності точок пошуку у вигляді (2.6), відповідає випадку застосування *сферичної норми* для вектору градієнта $f'(x)$.

Градiєнтний метод з різними метриками (нормами).

Напрямок пошуку p_k в градієнтному методі цілком залежить від вибраної норми вектора p_k (або пов'язаної з нею *спряженою нормою* вектора градієнта f'_k).

Спряжена q' -норма вектора градієнта f'_k до q -норми вектора p_k (1.10) має вигляд:

$$\|f'\|_{q'} = \left(\sum_{i=1}^n |f'_i|^q \right)^{\frac{1}{q'}}, \text{ де } \frac{1}{q} + \frac{1}{q'} = 1. \quad (2.10)$$

Таким чином, спряженість норм $\|p\|_q$ і $\|f'\|_{q'}$ векторів напрямку і градієнта полягає в тому, що октаедричній нормі $\|p\|_1$ відповідає кубічна норма вектора градієнта $\|f'\|_\infty$, сферичній нормі $\|p\|_2$ відповідає також сферична норма $\|f'\|_2$, а кубічній нормі $\|p\|_\infty$ - октаедрична норма $\|f'\|_1$.

Відповідні формули для напрямків p_k можна отримати в результаті розв'язання задачі на умовний екстремум:

$$\langle f'_k, p_k \rangle = \text{const}, \quad \|p_k\|_q = \left(\sum_{i=1}^n |p_{ki}|^q \right)^{\frac{1}{q}} = 1. \quad (2.11)$$

Застосування метода Лагранжа до задачі (2.11) дає наступний результат:

$$p_{ki} = -f'_{ki} \frac{|f'_{ki}|^{q-2}}{\left(\|f'_k\|_q \right)^{q-1}}, \quad i=1, 2, \dots, n. \quad (2.12)$$

Якщо розглядається октаедрична норма $\|f'\|_1$, з формули (2.12) маємо:

$$p_{ki} = -\frac{f'_{ki}}{|f'_{ki}|} = -\text{sign}(f'_{ki}), \quad i=1, 2, \dots, n. \quad (2.13)$$

Якщо застосовується сферична норма $\|f'\|_2$, то для напрямку пошуку маємо:

$$\rho_{ki} = -\frac{f'_{ki}}{\|f'_k\|_2} = -\frac{f'_{ki}}{\sqrt{\sum_{i=1}^n (f'_{ki})^2}}, \quad i=1, 2, \dots, n. \quad (2.14)$$

При кубічній нормі $\|f'\|_\infty$ напрямок пошуку задається формулою:

$$\rho_{ki} = \begin{cases} 0, & \text{якщо } |f'_{ki}| < M \\ \text{sign}(f'_{ki}), & \text{якщо } |f'_{ki}| = M \end{cases}, \quad i=1, 2, \dots, n. \quad (2.15)$$

де $M = \max |f'_{ki}|$.

Алгоритм градієнтного методу з різними метриками

Вхідними даними є цільова функція двох змінних $f(x)$, $x=(x_1, x_2)$, початкова точка x_0 , точність σ , номер ітерації $k=0$.

Вихідними даними є точка мінімуму x і значення мінімуму цільової функції $f(x)$.

1) Визначити поточний вектор градієнта цільової функції $f'(x_k)$.

2) Перевірити умову закінчення пошуку в точці $x_k=(x_{k1}, x_{k2})$:

$$\|f'_k\| \leq \sigma. \quad (2.16)$$

Якщо умова (2.16) виконується, то покласти $x=x_k$ і обчислити $f(x)$, інакше перейти до 3).

3) Визначити напрямок пошуку $\rho_k=(\rho_{k1}, \rho_{k2})$ в точці x_k для вибраного випадку норми (октаедричної, сферичної або кубічної) за однією з формул (2.13)-(2.15) для кількості змінних $n=2$.

4) Визначити довжину кроку α_k в напрямку ρ_k або на основі розв'язання задачі одновимірної мінімізації $f(x_k + \alpha \rho_k)$, або за схемою дроблення кроку.

Якщо цільова функція є квадратичною, то величина кроку обчислюється за формулою:

$$\alpha_k = -\frac{f'_{k, \rho_k}}{\text{H}_{\rho_k, \rho_k}} \cdot \epsilon.$$

5) Визначити $x_{k+1} = x_k + \alpha_k \rho_k$.

6) Повторити операції 1)-5), поклавши $k=k+1$ в відповідних формулах.

2.2. Лабораторна робота "Мінімізація функції двох змінних градієнтним методом з різними метриками"

Мета роботи – застосувати алгоритм градієнтного методу з різними метриками для пошуку точки мінімуму цільової функції $f(x)$. Визначити

найкращий варіант норми для мінімізації заданої функції за кількістю ітерацій (кількістю обчислень цільової функції).

Проміжні результати пошуку необхідно представити у вигляді таблиць 2.1.-2.3.

Кінцевими результатами є знайдена точка мінімуму x^* і мінімальне значення цільової функції $f(x^*)$.

Таблиця 2.1 – Результати пошуку мінімуму градієнтним методом з октаедричною нормою

№ ітерації k	x_k	ρ_k	α_k	$f(x_k)$	$\ f'(x_k)\ $
0					
...					
n					

Таблиця 2.2 – Результати пошуку мінімуму градієнтним методом зі сферичною нормою

№ ітерації k	x_k	ρ_k	α_k	$f(x_k)$	$\ f'(x_k)\ $
0					
...					
n					

Таблиця 2.3 – Результати пошуку мінімуму градієнтним методом з кубічною нормою

№ ітерації k	x_k	ρ_k	α_k	$f(x_k)$	$\ f'(x_k)\ $
0					
...					
n					

Контрольні питання

1. Як визначається величина кроку в градієнтному методі найшвидшого спуску?
2. Як визначається величина кроку в градієнтному методі з подрібненням кроку?
3. Наведіть умову закінчення процесу пошуку в градієнтному методі.

4. Як обчислюється вектор напрямку пошуку в градієнтному методі з октаедричною нормою? Наведіть формулу.
5. Наведіть формулу для обчислення напрямку пошуку в градієнтному методі із сферичною Нормою.
6. Запишіть формулу для обчислення напрямку пошуку в градієнтному методі з кубічною нормою.
7. За яким правилом змінюється величина кроку в градієнтному методі з дробленням кроку?
8. Як співвідносяться між собою градієнти цільової функції в сусідніх точках пошуку?
9. Чим відрізняється градієнтний метод найскорішого спуску від градієнтного методу з дробленням кроку?
10. Дайте визначення q' -норми вектора градієнта f'_k .

2.3. Метод спряжених градієнтів Флетчера-Рівса

Метод Флетчера-Рівса відноситься до *методів спряжених напрямків* і використовує градієнт цільової функції, тобто є методом першого порядку.

❖ *Визначення.*

Розглянемо вектор $p_0, p_1, \dots, p_{n-1} \in \mathbb{R}^n$. Візьмемо позитивно визначену симетричну матрицю H розміру $n \times n$. Якщо виконуються умови:

$$\left\{ \begin{array}{l} (Hp_j, p_k) = 0, j \neq k \\ (Hp_k, p_k) \neq 0 \end{array} \right\}, \quad (2.17)$$

то вектори p_0, p_1, \dots, p_{n-1} називаються *H-спряженими*, або спряженими по відношенню до матриці H . Зауважимо, що якщо $H=E$, то умова (2.17) перетворюється на звичайну умову ортогональності векторів. Тобто спряженість векторів - це узагальнення властивості ортогональності векторів.

Властивості спряжених векторів.

1. Спряжені вектори є *лінійно незалежними* векторами.
2. Спряжені вектори $p_0, p_1, \dots, p_{n-1} \in \mathbb{R}^n$ утворюють *базис* в \mathbb{R}^n . (Кількість спряжених векторів повинна дорівнювати розміру простору).

❖ *Визначення.*

Методи пошуку мінімуму, які засновані на формуванні спряжених напрямків пошуку. називаються *методами спряжених напрямків*.

Твердження.

Мінімізація квадратичної функції $f(x)$ n змінних з позитивно визначеною матрицею Гессе H може бути реалізована за n або *меншу кількість кроків* одновимірної мінімізації, якщо ці кроки робити послідовно в спряжених напрямках p_0, p_1, \dots, p_{n-1} .

Це означає, що для квадратичної функції двох змінних метод спряжених напрямків гарантовано приводить в точку мінімуму за дві ітерації. В деяких випадках, які залежать від вдалого завдання початкової точки x_0 , це може відбутися за одну ітерацію.

Формування спряжених напрямків.

Нехай заданий деякий початковий напрямок пошуку p_0 в точці x_0 , де $p_0, x_0 \in R^n$, і в результаті мінімізації функції $f(x)$ в цьому напрямку отримана точка x_1 :

$$f(x_0 + \alpha p_0).$$

Для формування напрямку пошуку p_1 в точці x_1 , який буде спряженим до p_0 , треба розв'язати рівняння

$$H p_0, p_1 \geq 0. \quad (2.18)$$

Якщо цільова функція $f(x)$ квадратична, то H - це матриця Гессе цієї функції, інакше в якості H може бути використана будь-яка симетрична позитивно визначена матриця розміру $n \times n$.

Розглянемо випадок функції двох змінних. Тоді $p_0 = (p_{01}, p_{02})$, $H = \begin{pmatrix} h_{11} & h_{12} \\ h_{12} & h_{22} \end{pmatrix}$

і рівняння (2.18) набуває вигляду:

$$\begin{pmatrix} h_{11} & h_{12} \\ h_{12} & h_{22} \end{pmatrix} \begin{pmatrix} p_{01} \\ p_{02} \end{pmatrix} \begin{pmatrix} p_{11} \\ p_{12} \end{pmatrix} \geq 0,$$

де p_{11} , p_{12} - невідомі компоненти вектора p_1 . Після перетворень отримаємо лінійне однорідне рівняння з двома невідомими

$$(h_{11} p_{01} + h_{12} p_{02}) p_{11} + (h_{12} p_{01} + h_{22} p_{02}) p_{12} = 0. \quad (2.18)$$

Достатньо отримати будь-який розв'язок цього рівняння. Для цього треба, наприклад, довільно задати $p_{11} \neq 0$ і потім визначити p_{12} :

$$p_{12} = - \frac{(h_{11} p_{01} + h_{12} p_{02}) p_{11}}{h_{12} p_{01} + h_{22} p_{02}}.$$

➤ **Зауваження.**

Довільність завдання $p_{11} \neq 0$ впливає тільки на довжину вектора p_1 , але не впливає на напрямок вектора p_1 , отже це не впливає на кінцевий результат мінімізації в напрямку p_1 , бо положення наступної точки $x_2 = x_1 + \alpha_1 p_1$ коректується довжиною кроку α_1 .

В методі спряжених градієнтів Флетчера-Рівса початковий напрямок пошуку в точці x_0 відповідає напрямку антиградієнта в цій точці: $p_0 = -f'_0$, а в будь-якій іншій поточній точці x_k релаксаційної послідовності задається як лінійна комбінація антиградієнта в поточній точці $-f'_k$ і напрямку пошуку на попередньому кроці:

$$p_k = -f'_k + \beta_{k-1} p_{k-1}.$$

Для метода Флетчера-Рівса мають місце наступні твердження.

Твердження 1. В методі Флетчера-Рівса два сусідні градієнта f'_k і f'_{k+1} ортогональні:

$$f'_{k+1}, f'_k \geq 0.$$

Твердження 2. В методі Флетчера-Рівса мінімізації квадратичної цільової функції напрямки пошуку p_0, p_1, \dots, p_{n-1} Н-спряжені, а градієнти $f'_0, f'_1, \dots, f'_{n-1}$ взаємно ортогональні.

Алгоритм методу Флетчера-Рівса для випадку квадратичної цільової функції.

Вхідними даними є квадратична цільова функція двох змінних $f(x)$, початкова точка x_0 та точність σ , номер ітерації $k=0$.

Вихідними даними є точка мінімуму x і значення мінімуму цільової функції $f(x)$.

1) Визначити поточний вектор градієнта цільової функції $f'(x_k)$.

2) Перевірити умову закінчення пошуку в точці $x_k = (x_{k1}, x_{k2})$:

$$\|f'_k\| \leq \sigma. \quad (2.19)$$

Якщо умова (2.19) виконується, то покласти $x=x_k$ і обчислити $f(x)$, інакше перейти до 3).

3) Визначити напрямок пошуку в поточній точці x_k :

$$p_0 = -f'_0, \text{ якщо } k=0,$$

$$p_k = -f'_k + \beta_{k-1} p_{k-1}, \text{ якщо } k=1, 2, \dots, \beta_{k-1} = f'_k, \text{ якщо } \text{Нр}_{k-1} > \frac{\text{Нр}_{k-1}, p_{k-1}}{\epsilon}. \quad (2.20)$$

4) Визначити довжину кроку α_k за формулою:

$$\alpha_k = -\frac{f'_k, p_k}{\text{Нр}_{k, p_k} > \epsilon}.$$

5) Визначити $x_{k+1} = x_k + \alpha_k p_k$.

6) Повторити операції 1)-5), збільшуючи в відповідних формулах номер ітерації k на 1.

Алгоритм методу Флетчера-Рівса для випадку неквадратичної цільової функції

Вхідними даними є цільова функція двох змінних $f(x)$, початкова точка x_0 та точність σ , номер ітерації $k=0$.

Вихідними даними є точка мінімуму x і значення мінімуму цільової функції $f(x)$.

1) Визначити поточний вектор градієнта цільової функції $f'(x_k)$.

2) Перевірити умову закінчення пошуку в точці x_k :

$$\|f'_k\| \leq \sigma. \quad (2.20)$$

Якщо умова (2.20) виконується, то покласти $x=x_k$ і обчислити $f(x)$, інакше перейти до 3).

3) Визначити напрямок пошуку в поточній точці x_k :

$$p_k = -f'_k, \text{ якщо } k \in K, K = (0, n, 2n, 3n, \dots),$$

$$p_k = -f'_k + \beta_{k-1} p_{k-1}, \text{ якщо } k = 1, 2, \dots, \beta_{k-1} = f'_k, f'_k > \frac{f'_{k-1}, f'_{k-1} > \epsilon}{f'_{k-1}, f'_{k-1} > \epsilon}. \quad (2.20)$$

4) Визначити довжину кроку α_k в напрямку p_k або на основі розв'язання задачі одновимірної мінімізації $f(x_k + \alpha p_k)$, або за схемою подрібнення кроку.

5) Визначити $x_{k+1} = x_k + \alpha_k p_k$.

6) Повторити операції 1)-5), збільшуючи в відповідних формулах номер ітерації k на 1.

➤ **Зауваження.**

В модифікації Полака-Рібьєра метода спряжених градієнтів коефіцієнт β_{k-1} в формулах (2.20) обчислюється за формулою:

$$\beta_{k-1} = f'_k, f'_k - f'_{k-1} > \frac{f'_{k-1}, f'_{k-1} > \epsilon}{f'_{k-1}, f'_{k-1} > \epsilon}.$$

2.4. Лабораторна робота "Мінімізація функції двох змінних методом спряжених градієнтів Флетчера-Рівса "

Мета роботи: застосувати алгоритм методу Флетчера-Рівса для пошуку точки мінімуму цільової функції $f(x)$.

Проміжні результати пошуку необхідно представити у вигляді таблиці 2.4.

Кінцевими результатами є знайдена точка мінімуму x і мінімальне значення цільової функції $f(x)$.

Таблиця 2.4 – Результати пошуку мінімуму методом Флетчера-Рівса

№ ітерації k	x_k	p_k	α_k	$f(x_k)$	$\ f'(x_k)\ $
0					
...					
n					

Контрольні питання

1. Дайте визначення спряженим векторам.
2. Доведіть, що спряжені вектори є лінійно незалежними.

3. Як співвідносяться властивості спряженості векторів і ортогональності векторів?
4. Коли спряжені вектори утворюють базис в просторі R^n ?
5. Сформулюйте твердження про мінімізацію квадратичної функції n змінних методом спряжених напрямків.
6. Як формуються спряжені напрямки за допомогою симетричної позитивно визначеної матриці?
7. Сформулюйте напрямок p_1 , який є Н-спряженим по відношенню до напрямку s . $p_0 = (1, -1)$, $H = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}$.
8. Як формується напрямок пошуку в методі Флетчера-Рівса у випадку квадратичної цільової функції?
9. Як формується напрямок пошуку в методі Флетчера-Рівса у випадку неквадратичної цільової функції?
10. Як формується напрямок пошуку в модифікації Полака-Ріб'єра?
11. Сформулюйте твердження про властивості градієнтів і напрямків пошуку в методі Флетчера-Рівса.

2.5. Метод Пауелла

Властивість паралельного підпростору.

Розглянемо квадратичну функцію $f(x)$ з позитивно визначеною матрицею Гессе H . Задамо напрямок пошуку d_1 і проведемо мінімізацію $f(x)$ в напрямку d_1 з двох точок x_0 і y_0 . Отримаємо в результаті мінімізації точки x_1 і y_1 (рис. 2.2). Можна показати, що напрямок d_1 і напрямок $(x_1 - y_1)$ будуть H -спряженими. Для квадратичної цільової функції двох змінних це означає, що точка мінімуму x^* буде належати напрямку $(x_1 - y_1)$. Ця ситуація в теорії оптимізації називається *властивістю паралельного підпростору*.

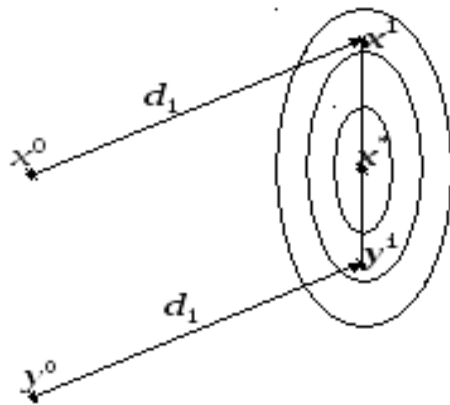


Рисунок 2.2 – Формування спряжених напрямків

Таким чином, формування H -спряжених напрямків засноване на подвійній мінімізації квадратичної функції в одному і тому ж напрямку пошуку. Недоліком представленої схеми формування H -спряжених напрямків є наявність двох точок x_0 і y_0 , які не являються точками релаксаційної послідовності. Цього недоліку позбавлений метод Пауелла, який також формує систему H -спряжених напрямків мінімізації на основі властивості паралельного підпростору.

Метод Пауелла - це метод, який найбільш ефективно застосовується для мінімізації функцій, близьких до квадратичних. Метод Пауелла є розвитком методу циклічного покоординатного спуску, в якому послідовний одновимірний пошук здійснюється в напрямку ортів координатних осей $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, \dots, 0)$, ..., $e_n = (0, 0, \dots, 1)$. Але на відміну від методу покоординатного спуску в методі Пауелла формується система спряжених напрямків.

Кожна ітерація методу Пауелла починається з визначення n лінійно незалежних напрямків p_1, p_2, \dots, p_n . Спочатку ці напрямки співпадають з оортами координатних осей e_1, e_2, \dots, e_n .

З початкової точки x_0 здійснюється одновимірний пошук в напрямку p_n :

$$x_1 = x_0 + \alpha_0 p_n, \quad f(x_0 + \alpha p_n). \quad (2.21)$$

Потім з точки x_1 здійснюється послідовний покоординатний пошук вздовж напрямків p_1, p_2, \dots, p_n :

$$x_{k+2} = x_{k+1} + \alpha_{k+1} p_{k+1}, \quad f(x_{k+1} + \alpha p_{k+1}), \quad k = \overline{0, n-1}. \quad (2.22)$$

Останньою точкою після виконання процедури покоординатного пошуку (2.21) є точка x_{n+1} , яка отримана в результаті мінімізації функції $f(x)$ з точки x_n в напрямку p_n . Точка x_1 також отримана в результаті мінімізації вздовж напрямку p_n . Тоді у випадку квадратичної функції $f(x)$ напрямком $d = x_{n+1} - x_1$ буде спряженим до напрямку p_1 на основі властивості паралельного підпростору. Далі із системи напрямків p_1, p_2, \dots, p_n виключається напрямок p_1 , а решта напрямків перепозначається і до них додається напрямок d :

$$p_i = p_{i+1}, \quad i = \overline{1, n-1}, \quad p_n = d. \quad (2.23)$$

Отже, для квадратичної функції $f(x)$ після першої ітерації останні напрямки p_{n-1} і p_n будуть спряженими.

Наступна ітерація починається з точки $x_0 = x_{n+1}$ і виконується за формулами (2.21), (2.22). Ітерації продовжуються, поки норма вектора d залишається меншою за задану точність σ .

Таким чином, для квадратичної функції після k ітерацій методу Пауелла останні $k+1$ напрямки пошуку p_{n-k}, \dots, p_n , які сформовані перед $(k+1)$ -ю ітерацією, будуть взаємно спряженими. Після виконання $n-1$ ітерацій всі напрямки пошуку будуть взаємно спряженими, тому метод Пауелла забезпечує мінімізацію квадратичної функції n змінних не більше, ніж за n ітерацій. Для мінімізації квадратичної функції 2-х змінних достатньо виконання двох ітерацій. При цьому точка мінімуму буде знайдена після виконання чотирьох одновимірних пошуків.

Алгоритм методу Пауелла.

Вхідними даними є цільова функція двох змінних $f(x)$, початкова точка x_0 та точність σ .

Вихідними даними є точка мінімуму x^* і значення мінімуму цільової функції $f(x^*)$.

1) Задати початкові напрямки пошуку $p_1 = (1, 0)$, $p_2 = (0, 1)$.

Виконати одновимірний пошук:

$$x_1 = x_0 + \alpha_0 p_2, \quad f(x_0 + \alpha p_2).$$

2) З точки x_1 виконати послідовний одновимірний пошук в напрямках p_1 , p_2 :

$$x_2 = x_1 + \alpha_1 p_1, \quad f(x_1 + \alpha p_1),$$

$$x_3 = x_2 + \alpha_2 p_2, \quad f(x_2 + \alpha p_2).$$

3) Визначити напрямок

$$d = x_3 - x_1.$$

4) Перевірити умову закінчення:

$$\|d\| \leq \sigma. \quad (2.24)$$

Якщо умова (2.24) виконується, покласти $x = x_3$, $f(x) = f(x_3)$.

Інакше перейти до 5).

5) Перепозначити: $p_1 = p_2$, $p_2 = d$, $x_0 = x_3$. Перейти до 2).

2.6. Лабораторна робота "Мінімізація функції двох змінних методом Пауелла"

Мета роботи – застосувати алгоритм методу Пауелла для пошуку точки мінімуму цільової функції $f(x)$.

Проміжні результати пошуку представити у вигляді таблиці 2.5.

Кінцевими результатами є знайдена точка мінімуму x і мінімальне значення цільової функції $f(x)$.

Таблиця 2.5 – Результати пошуку мінімуму методом Пауелла

№ ітерації	k	x_k	p_k	α_k	d	$\ d\ $
1	0					
	1					
	2					
	3					
2	0					
...

Контрольні питання

1. Сформулюйте властивість паралельного підпростору.
2. В чому полягає основна ідея методу Пауелла?
3. До якої групи методів відноситься метод Пауелла? Класифікуйте.
4. Опишіть ітераційну частину алгоритма методу Пауелла.

5. Запишіть умову закінчення процесу пошуку в алгоритмі методу Пауелла.
6. Для мінімізації яких функцій метод Пауелла є найбільш ефективним?
7. Як задається початкова система напрямків пошуку в методі Пауелла?
8. Скільки ітерацій потрібно виконати, щоб мінімізувати квадратичну функцію двох змінних методом Пауелла?
9. Скільки одновимірних пошуків треба здійснити, щоб мінімізувати квадратичну функцію двох змінних методом Пауелла?
10. Як формується система спряжених напрямків в методі Пауелла?

2.7. Метод Ньютона-Рафсона

Метод Ньютона-Рафсона є модифікацією методу Ньютона, який заснований на квадратичній апроксимації цільової функції $f(x)$ в припущенні, що $f(x)$ є двічі диференційована, а її матриця Гессе $f''(x)$ є позитивно визначеною. Розкладемо $f(x)$ в околі деякої заданої точки x_0 в ряд Тейлора, утримуючи члени другого порядку малості по $\|\Delta x\|^2$:

$$f(x_0 + \Delta x) = f(x_0) + < . \quad (2.25)$$

Результатом апроксимації є квадратична функція $g(\Delta x)$:

$$g(\Delta x) = f(x) + < .$$

Визначимо значення аргументу Δx , яке забезпечує мінімум цієї функції, враховуючи умови оптимальності першого порядку:

$$\frac{d}{d\Delta x} g(\Delta x) = f'(x_0) + f''(x_0) \Delta x = 0. \quad (2.26)$$

Отримаємо з (2.26), що точка мінімуму x знаходиться від точки x_0 на відстані

$$\Delta x = - (f''(x_0))^{-1} f'(x_0),$$

тобто

$$x = x_0 - (f''(x_0))^{-1} f'(x_0). \quad (2.27)$$

Оскільки відносно точки x_0 не робилось ніяких припущень, (2.27) означає, що точка мінімуму для цільової квадратичної функції знаходиться всього за одну ітерацію.

Якщо $f(x)$ не квадратична функція, (2.27) набуває вигляду:

$$x_{k+1} = x_k - (f''(x_k))^{-1} f'(x_k), \quad (2.28)$$

де $(f''(x_k))^{-1}$ - матриця, зворотня до матриці Гессе функції $f(x)$ в точці x_k .

Це ітераційна формула *методу Ньютона*. Очевидно, що довжина кроку в напрямку $-(f''(x_k))^{-1} f'(x_k)$ на кожній ітерації методу Ньютона дорівнює *одиниці*. Метод Ньютона є методом другого порядку, бо використовує значення матриці Гессе на кожній ітерації. Якщо цільова функція $f(x)$ є квадратичною функцією двох змінних, метод Ньютона потребує всього *одну ітерацію* для визначення точки мінімуму x . Швидкість збіжності методу Ньютона є *квадратичною*.

➤ **Зауваження.**

Збіжність методу Ньютона залежить від розташування початкової точки x_0 відносно положення точки мінімуму x . Якщо $\|x - x_0\| > 1$, метод Ньютона не гарантує збіжності.

Метод *Ньютона-Рафсона* є покращенням методу Ньютона (2.28), основна відмінність якого є застосування методу одновимірної мінімізації для

знаходження оптимального кроку на поточній ітерації. Метод Ньютона-Рафсона - це метод Ньютона із змінним кроком. Робоча формула методу Ньютона-Рафсона має вигляд:

$$x_{k+1} = x_k - \alpha_k (f''(x_k))^{-1} f'(x_k), \quad (2.29)$$

де

$$f(x_k - \alpha (f''(x_k))^{-1} f'(x_k)). \quad (2.30)$$

Якщо для розв'язання задачі (2.30) вибрано метод з подрібненням кроку, то відповідна умова (1.7) для методу Ньютона-Рафсона має вигляд:

$$f(x) - f(x_k) \leq -\varepsilon \alpha (f'_k, (f''_k)^{-1} f'_k), \quad (2.31)$$

де $f(x) = x_k - \alpha (f''_k)^{-1} f'_k$.

Алгоритм методу Ньютона-Рафсона.

Вхідними даними є цільова функція двох змінних $f(x)$, $x = (x_1, x_2)$, початкова точка x_0 , точність σ , початкове значення кроку на кожній ітерації $\alpha^0 = 1$, мале число $\varepsilon \in (0, 1)$, номер ітерації $k=0$.

Вихідними даними є точка мінімуму x і значення мінімуму цільової функції $f(x)$.

1) Обчислити значення градієнта $f'_k = f'(x_k)$. Перевірити умову закінчення пошуку:

$$\|f'_k\| \leq \sigma. \quad (2.32)$$

Якщо умова (2.32) виконується, процес пошуку закінчено, і $x = x_k$. Інакше перейти до 2).

2) Обчислити в поточній точці x_k значення функції $f(x_k)$ і зворотню матрицю Гессе $(f''_k)^{-1} = (f''(x_k))^{-1}$, покласти $\alpha = \alpha^0$.

3) Визначити напрямок пошуку в точці x_k :

$$p_k = - (f''_k)^{-1} f'_k.$$

4) Покласти $x = x_k - \alpha (f''_k)^{-1} f'_k$, обчислити $f(x)$.

5) Перевірити виконання умови:

$$f(x) - f(x_k) \leq -\varepsilon \alpha (f'_k, (f''_k)^{-1} f'_k). \quad (2.33)$$

Якщо умова (2.33) виконується, покласти $x_{k+1} = x$, збільшити номер ітерації k , перейти до 1). Інакше зменшити величину кроку вдвічі $\alpha := \alpha/2$, перейти до 4).

➤ **Зауваження.**

Величину кроку α на кожній ітерації слід зменшувати до тих пір, поки $\alpha > \sigma$. Якщо $\alpha \leq \sigma$ і умова $\|f'_k\| \leq \sigma$ не виконана, це означає, що задача не може бути розв'язана з заданою точністю.

2.8. Лабораторна робота "Мінімізація функції двох змінних методом Ньютона-Рафсона"

Мета роботи – застосувати алгоритм методу Ньютона-Рафсона для пошуку точки мінімуму цільової функції $f(x)$.

Проміжні результати пошуку необхідно представити у вигляді таблиці 2.6.

Кінцевими результатами є знайдена точка мінімуму x^* і мінімальне значення цільової функції $f(x)$.

Таблиця 2.6 – Результати пошуку мінімуму методом Ньютона-Рафсона

№ ітерації k	x_k	p_k	α_k	$f(x_k)$	$\ f'(x_k)\ $
0					
...					
n					

Контрольні питання

1. До якої групи методів належить метод Ньютона і модифікація Рафсона? Класифікуйте.

2. Яка умова збіжності методу Ньютона?

3. Доведіть, що матриця $\begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix}$ є позитивно-визначеною.

4. Знайдіть матрицю Гессе $f''(x)$ для функції

$$f(x) = 5x_1^2 + 2x_2^2 - 2x_1x_2 - x_1 + 6x_2 - 3.$$

5. Знайдіть зворотню матрицю Гессе $(f''(x))^{-1}$ для функції

$$f(x) = 8x_1^2 + x_2^2 + x_1x_2 - x_1 + x_2.$$

6. Чому дорівнює величина кроку на кожній ітерації в методі Ньютона?

7. Як визначається величина кроку на кожній ітерації в методі Ньютона-Рафсона?

8. Як виглядає умова закінчення процесу пошуку в алгоритмі методу Ньютона-Рафсона?

9. Скільки кроків методу Ньютона-Рафсона треба виконати, щоб мінімізувати квадратичну функцію двох змінних з позитивно визначеною матрицею Гессе?

2.9. Метод Хука-Дживса

Метод Хука-Дживса - це прямий метод, який є розвитком методу покоординатного спуску. Оскільки процедура пошуку в методі покоординатного спуску базується на циклічній зміні координатних напрямків пошуку, то це може привести до неефективності пошуку і навіть до відсутності збіжності до точки локального мінімуму x^* , якщо лінії рівня цільової функції мають "яружний" характер. Метод Хука-Дживса дозволяє суттєво прискорити збіжність пошуку за рахунок того, що пошук періодично проводиться в напрямку так званого "зразка".

Метод Хука-Дживса включає в себе дві процедури: "досліджуючий пошук" і "пошук за зразком".

"Досліджуючий пошук" здійснюється навколо початкової (базисної) точки x_{k-1} і нагадує схему покоординатного спуску. Задається величина кроку Δ_i , яка може бути різною для різних координатних напрямків і змінюватись в процесі пошуку. Значення кожної i -ї координати послідовно змінюють спочатку на величину Δ_i в позитивному напрямку, решту координат залишають незмінними. Якщо значення функції зменшилося, то це "успіх", цю координату фіксують і переходять до іншої координати. Якщо значення функції не зменшилося, то це "невдача", тоді початкове значення цієї координати зменшують на Δ_i і також перевіряють, як змінилося значення функції. Якщо при зміні координати на величину $\pm \Delta_i$ функція не зменшилася, цю координату залишають незмінною і переходять до іншої координати. Після перебору всіх координат маємо нову точку x_k . "Досліджуючий пошук" вважається успішним, якщо

$$f(x_k) < f(x_{k-1}). \quad (2.34)$$

Якщо отримали, що $x_k = x_{k-1}$, то треба зменшити величину Δ_i і повторити "досліджуючий пошук". Зменшення величини кроку Δ_i відбувається, поки $\Delta_i > \sigma$. Якщо всі змінені $\Delta_i \leq \sigma$, це означає, що $x = x_{k-1}$.

При $x_k \neq x_{k-1}$ переходять до "пошуку за зразком".

"Пошук за зразком" здійснюється з точки x_k в напрямку $(x_k - x_{k-1})$:

$$x_p = x_k + (x_k - x_{k-1}) = 2x_k - x_{k-1}. \quad (2.35)$$

Якщо $f(x_p) < f(x_k)$, точки перепозначаються: $x_{k-1} = x_k$, $x_k = x_p$ і процедура (2.35) повторюється. Далі виконується "досліджуючий пошук" навколо точки x_p . В результаті отримуємо точку x_{k+1} . Цей пошук вважається "успішним", якщо

$$f(x_{k+1}) < f(x_k). \quad (2.36)$$

Далі базисна точка перепозначається:

$$x_{k-1} = x_{k+1}, \quad (2.37)$$

і процедура пошуку Хука-Дживса повторюється.

Алгоритм методу Хука-Дживса.

Вхідними даними є цільова функція двох змінних $f(x)$, $x=(x_1, x_2)$, початкова точка x_{k-1} , точність σ , початкове значення прирощень по координатам Δ_i ($i=1,2$), коефіцієнт зменшення кроку $r>1$, номер ітерації $k=1$.

Вихідними даними є точка мінімуму x і значення мінімуму цільової функції $f(x)$.

1) Провести "досліджуючий пошук" навколо точки x_{k-1} , знайти точку x_k .

2) Перевірити, чи був "досліджуючий пошук" успішним:

$$f(x_k) < f(x_{k-1}) ?$$

Якщо так, то перейти до 4). Інакше перейти до 3).

3) Перевірка кінця пошуку:

$$\|x_k - x_{k-1}\| \leq \sigma.$$

Якщо умова закінчення виконана, то $x=x_k$. Інакше треба зменшити прирощення Δ_i/r , ($i=1,2$) і перейти до 1).

4) Провести "пошук за зразком":

$$x_p = 2x_k - x_{k-1}.$$

5) Перевірити виконання умови:

$$f(x_p) < f(x_k). \quad (2.38)$$

Якщо (2.38) виконується, то перепозначити точки: $x_{k-1}=x_k$, $x_k=x_p$, і продовжити "пошук за зразком" 4). Інакше перейти до 5).

5) Виконати "досліджуючий пошук" навколо точки x_p , отримати в результаті точку x_{k+1} .

б) Перевірити успішність "пошуку за зразком":

$$f(x_{k+1}) < f(x_k).$$

Якщо пошук успішний, то $x_{k-1}=x_k$, $x_k=x_{k+1}$, перейти до 1). Інакше $x_{k-1}=x_k$, перейти до 3).

2.10. Лабораторна робота "Мінімізація функції двох змінних методом Хука-Дживса"

Мета роботи – застосувати алгоритм методу Хука-Дживса для пошуку точки мінімуму цільової функції $f(x)$.

Проміжні результати пошуку представити у вигляді таблиці 2.7.

Кінцевими результатами є знайдена точка мінімуму x^* і мінімальне значення цільової функції $f(x^*)$.

Таблиця 2.7 – Результати пошуку мінімуму методом Хука-Дживса

№ ітерації k	x_{k-1}	x_k	x_p	x_{k+1}	$f(x_{k-1})$	$f(x_k)$	$f(x_{k+1})$
1							
2							
...							

Контрольні питання

1. Дайте класифікацію методу Хука-Дживса.
2. Як виконується операція "досліджуючий пошук" навколо точки x_{k-1} .
3. Коли "досліджуючий пошук" навколо точки x_{k-1} є успішним?
4. В чому полягає операція "пошук за зразком"?
5. Коли "пошук за зразком" вважається успішним?
6. Опишіть операцію "досліджуючий пошук" навколо точки x_p .
7. Наведіть умову закінчення пошуку в методі Хука-Дживса.

2.11. Метод деформованого багатогранника

Метод деформованого багатогранника є прямим методом пошуку і відноситься до групи методів мінімізації по симплексу. Симплекс (від лат. simplex - простий) - це найпростіший опуклий багатогранник в просторі E^n з $(n+1)$ вершинами (в E^2 симплекс - це трикутник з точками-вершинами, в E^3 - тетраедр). Кожна вершина багатогранника називається *вершиною симплекса*, відрізок, що з'єднує дві вершини - *ребром симплекса*. В *регулярному* (правильному) симплексі довжина ребер однакова (рис. 2.3). В *нерегулярному* симплексі довжина ребер може змінюватися в залежності від характеру розташування ліній рівня цільової функції. *Важливою властивістю* симплекса є те, що новий симплекс можна побудувати на будь-якій грані початкового симплекса шляхом переносу обраної вершини вздовж прямої, проведеної через *центр тяжіння* X_c решти вершин початкового симплекса (рис. 2.4).

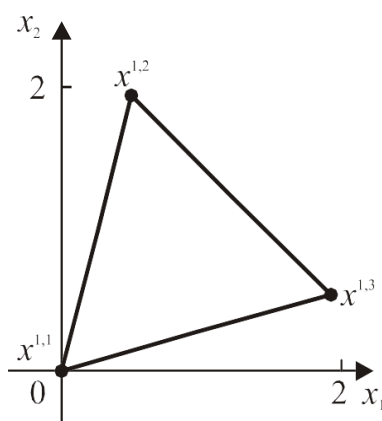


Рисунок 2.3 - Регулярний симплекс в E^2

В методі *деформованого багатогранника* Нелдера-Міда пошук мінімуму ведеться по нерегулярному симплексу.

Пошук в E^n починається із завдання координат вершин початкового симплексу x_1, x_2, \dots, x_{n+1} і обчислення значень цільової функції в точках-вершинах: $f_i = f(x_i)$, $(i=1, 2, \dots, n+1)$.

В методі Нелдера-Міда виконуються наступні операції:

1. {OMAX} - визначення найгіршої точки-вершини x_h симплексу $f(x_i)$, $(i=\overline{1, n+1})$. (2.39)

2. {OMIN}- визначення найкращої точки-вершини x_l симплексу $f(x_i)$, $(i=\overline{1, n+1})$. (2.40)

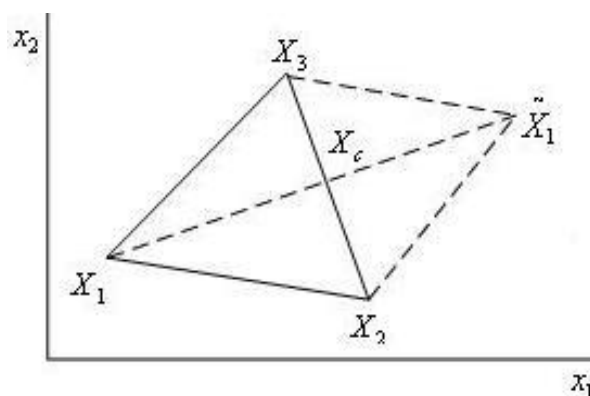


Рисунок 2.4 - Побудування нового симплексу (X_2, X_3, \tilde{X}_1) на грані (X_2, X_3) початкового симплексу (X_1, X_2, X_3) в E^2

3. {OCT} - операція визначення *центра тяжіння*:

$$x_{n+2} = \frac{\sum_{i=1}^{n+1} x_i - x_h}{n}. \quad (2.41)$$

4. {OB} - операція *відображення* (рис.2.4):

$$x_{n+3} = x_{n+2} + \alpha(x_{n+2} - x_h). \quad (2.41)$$

Рекомендується брати $\alpha=1$.

Операції {OMAX}, {OMIN}, {OCT} і {OB} виконуються завжди на кожній ітерації. Інші операції виконуються в залежності від виконання умов.

5. {OP} - операція *розтягнення* (рис.2.5):

$$x_{n+4} = x_{n+2} + \gamma(x_{n+3} - x_{n+2}), \quad (\gamma > 1). \quad (2.42)$$

Операція виконується, якщо виконується умова $f(x_{n+3}) < f_l$. Рекомендується брати $\gamma=2$.

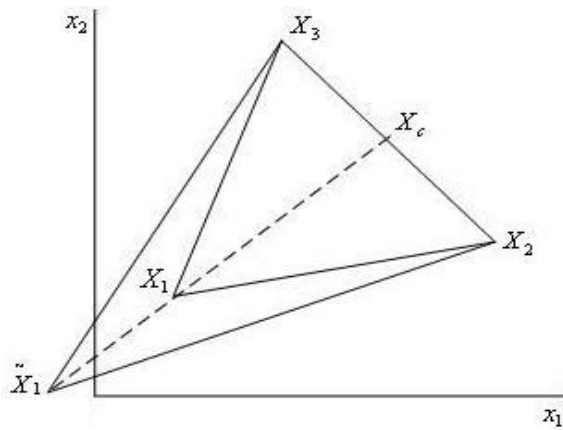


Рисунок 2.5 - Операція розтягнення в E^2

6. {ОСТ}- операція *стискання* (рис. 2.6):

$$x_{n+5} = x_{n+2} + \beta(x_h - x_{n+2}), \quad (0 < \beta < 1). \quad (2.43)$$

В результаті цієї операції вектор $(x_h - x_{n+2})$ стискається. Виконується за умови $f(x_{n+3}) \geq f_i$ для всіх $i \neq h$. Рекомендується брати $\beta = 0,5$.

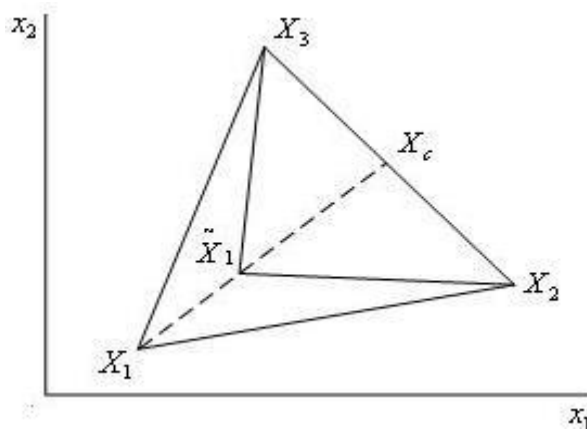


Рисунок 2.6 - Операція стискання в E^2

7. {ОРЕД}- операція *редукції* (рис. 2.7). Виконується, коли $f(x_{n+3}) \geq f_h$. В результаті операції редукції всі вектори $(x_i - x_l)$ зменшуються в ρ разів, рахуючи від точки x_l :

$$x_i = x_l + \rho(x_i - x_l), \quad i = \overline{1, n+1}, \quad (i \neq l). \quad (2.44)$$

Рекомендується брати $\rho = 0,5$.

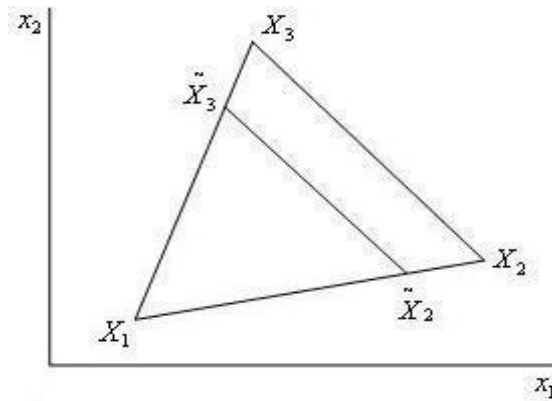


Рисунок 2.7 - Операція редукції симплекса в E^2

8. Умова завершення процесу пошуку. Процес пошуку завершується, якщо виконується одна з умов:

$$|f_h - f_l| \leq \sigma, \quad (2.45)$$

або

$$\frac{\sum_{i=1}^{n+1} \|x_i - x_{n+2}\|}{n+1} \leq \sigma, \quad (2.46)$$

або

$$\left(\frac{1}{n+1} \sum_{i=1}^{n+1} (f(x_i) - f(x_{n+2}))^2 \right)^{\frac{1}{2}} \leq \sigma. \quad (2.47)$$

Алгоритм методу Нелдера-Міда.

Вхідними даними є цільова функція двох змінних $f(x)$, $x=(x_1, x_2)$, початковий симплекс x_1, x_2, x_3 , точність σ , номер ітерації $k=1$.

Вихідними даними є точка мінімуму x і значення мінімуму цільової функції $f(x)$.

1) Обчислити значення цільової функції в вершинах симплекса (x_1, x_2, x_3) :

$$f_1=f(x_1), f_2=f(x_2), f_3=f(x_3).$$

2) Виконати операції $\{OMAX\}$, $\{OMIN\}$. Упорядкувати точки-вершини в порядку зростання значень функції в них:

$$f(x_1) < f(x_2) < f(x_3), x_i = \arg(f(x_i)), (i=\overline{1,3}),$$

$$f_h=f(x_3), f_l=f(x_1), x_l=x_1, x_h=x_3.$$

3) Виконати операцію $\{OCT\}$, визначити точку центра тяжіння:

$$x_4 = \frac{\sum_{i=1}^3 x_i - x_h}{2}.$$

4) Перевірити виконання умови закінчення пошуку:

$$\left(\frac{1}{3} \sum_{i=1}^3 (f(x_i) - f(x_4))^2\right)^{\frac{1}{2}} \leq \sigma.$$

Якщо умова виконана, пошук закінчено. Покласти $x=x_i$, $f(x)=f_i$. Інакше перейти до 5).

5) Виконати операцію віддзеркалення $\{OB\}$, знайти:

$$x_5 = x_4 + (x_4 - x_h), \quad f_5 = f(x_5).$$

6) Якщо $f(x_5) < f_i$, то виконати операцію розтягування $\{OP\}$:

$$x_6 = x_{n+2} + 2(x_5 - x_4), \quad f_6 = f(x_6).$$

Точка x_6 включається в число точок-вершин нового симплексу замість точки x_h :

$$x_3 = x_6, \quad f_3 = f_6.$$

Перейти до 3).

7) Якщо $f(x_5) \geq f_i$ для всіх $i \neq h$, то виконати операцію стискання $\{OST\}$:

$$x_7 = x_4 + 0,5(x_h - x_4), \quad f_7 = f(x_7).$$

Точка x_7 включається в число точок-вершин нового симплексу замість точки x_h :

$$x_3 = x_7, \quad f_3 = f_7.$$

Перейти до 3).

8) Якщо $f(x_5) \geq f_h$, то виконати операцію редукції $\{ORED\}$:

$$x_i = x_i + 0,5(x_i - x_i), \quad i = \overline{1,3}, \quad (i \neq 1).$$

Обчислити

$$f_i = f(x_i), \quad (i = \overline{1,3}).$$

Перейти до 2).

2.12. Лабораторна робота "Мінімізація функції двох змінних методом деформованого багатогранника

Мета роботи – застосувати алгоритм методу Нелдера-Міда для пошуку точки мінімуму цільової функції $f(x)$.

Проміжні результати пошуку представити у вигляді таблиці 2.8.

Кінцевими результатами є знайдена точка мінімуму x і мінімальне значення цільової функції $f(x)$.

Таблиця 2.8 – Результати пошуку мінімуму методом Нелдера-Міда

№ ітерації k	x_1	x_2	x_3	x_4	$f(x_1)$	$f(x_2)$	$f(x_3)$
1							
2							
...							

Контрольні питання

1. Дайте класифікацію методу деформованого багатогранника Нелдера-Міда.
2. Що таке регулярний симплекс в E^n ?
3. Які початкові умови в алгоритмі методу Нелдера- Міда?
4. В чому полягає операція {OMAX} ?
5. В чому полягає операція {OMIN}?
6. Опишіть операцію визначення *центра тяжіння* {OЦТ} .
7. Запишіть формулу операції *віддзеркалення* {OB} .
8. За яких умов виконується операція *розтягування* {OP} і в чому вона полягає?
9. Коли виконується операція *стискання* {OСТ} і в чому вона полягає?
10. Що таке операція *редукції* {OPЕД} і коли вона виконується?
11. Запишіть умову закінчення пошуку в методі Нелдера-Міда.

3. МЕТОДИ УМОВНОЇ МІНІМІЗАЦІЇ ФУНКЦІЙ БАГАТЬОХ ЗМІННИХ

3.1. Метод зовнішніх штрафних функцій

Розглянемо задачу умовної оптимізації

$$f(x) = f(x), \quad x \in X, \quad (3.1)$$

де $X \subset E^n$ - замкнена опукла множина, яка задана системою рівностей і нерівностей.

Зіставимо функції $f(x)$ наступну допоміжну функцію:

$$F(x) = f(x) + \delta(x|X). \quad (3.2)$$

Функція $\delta(x|X)$ називається *штрафом* або *індикаторною функцією*:

$$\delta(x|X) = \begin{cases} 0, & \text{якщо } x \in X \\ +\infty, & \text{якщо } x \notin X \end{cases}. \quad (3.3)$$

Побудуємо послідовність $\delta_k(x|X)$ таку, що для всіх $x \in E^n$

$$\delta_k(x|X) = \delta(x|X).$$

Тоді задача, еквівалентна до початкової задачі (3.1), запишеться у вигляді:

$$\delta_k(x|X), \quad \text{де } x \in E^n. \quad (3.4)$$

В цих умовах задача (3.1) на умовний екстремум може бути зведена до послідовності задач безумовної оптимізації:

$$f(x). \quad (3.5)$$

В залежності від способу формування штрафу розрізняють методи *внутрішніх штрафних функцій* і методи *зовнішніх штрафних функцій*.

Умовою використання методу *зовнішніх штрафних функцій* є присутність в системі обмежень як рівностей, так і нерівностей:

$$X = \{x: \phi_i(x) \geq 0, i = \overline{1, m}, \phi_i(x) = 0, i = \overline{m+1, l}\}. \quad (3.6)$$

Для обмежень-нерівностей вводиться функція зрізання:

$$\phi_i^-(x) = \min(0, \phi_i(x)), \quad i = \overline{1, m}.$$

Штрафна функція при цьому може формуватися у вигляді:

$$\delta(x|X) = \sum_{i=1}^m (\phi_i^-(x))^2 + \sum_{i=m+1}^l (\phi_i(x))^2. \quad (3.7)$$

Тоді допоміжна функція буде мати вигляд:

$$F(x, \lambda_k) = f(x) + \lambda_k \left(\sum_{i=1}^m (\phi_i^-(x))^2 + \sum_{i=m+1}^l (\phi_i(x))^2 \right),$$

де послідовність чисел $\{\lambda_k\}$ є необмежено зростаючою:

$$\lambda_k \rightarrow +\infty.$$

Зауважимо, що на відміну від методів внутрішніх штрафних функцій початкову точку x_0 можна задавати довільно, а *штраф* повинен бути додатною функцією.

Інший варіант завдання *штрафної функції* має вигляд:

$$\delta(x|X) = \sum_{i=1}^m |\phi_i^-(x)| + \sum_{i=m+1}^l |\phi_i(x)|. \quad (3.8)$$

Допоміжна функція при цьому набуває вигляду:

$$F(x, \lambda_k) = f(x) + \lambda_k \left(\sum_{i=1}^m |\phi_i^-(x)| + \sum_{i=m+1}^l |\phi_i(x)| \right).$$

Алгоритм методу зовнішніх штрафних функцій.

Вхідними даними є цільова функція двох змінних $f(x)$, $x=(x_1, x_2)$, обмеження $\phi_1(x) \geq 0$, $\phi_2(x) = 0$, початкова точка x_0 , початковий елемент послідовності $\{\lambda_k\}$: $\lambda_0 = 1$, точність σ , номер ітерації $k=0$.

Вихідними даними є точка мінімуму x і значення мінімуму цільової функції $f(x)$.

1) Обчислити $\phi_1(x_k)$. Якщо $\phi_1(x_k) < 0$, покласти $\phi_1^-(x) = \phi_1(x)$. Інакше покласти $\phi_1^-(x) = 0$.

2) З початкової точки x_k розв'язати задачу *безумовної оптимізації* з заданою точністю σ :

$$x_{k+1} = \operatorname{argmin} F(x, \lambda_k),$$

де $F(x, \lambda_k) = f(x) + \lambda_k ((\phi_1^-(x))^2 + \phi_2(x)^2)$.

3) Перевірити виконання обмежень:

$$0 < \phi_1(x_{k+1}) \leq \sigma, \quad |\phi_2(x_{k+1})| \leq \sigma. \quad (3.9)$$

Якщо умови (3.9) виконуються, пошук завершено: $x = x_{k+1}$, $f(x) = f(x_{k+1})$. Інакше покласти $\lambda_k := \lambda_k \cdot 10$, $k := k+1$. Перейти до 1).

➤ **Зауваження.**

Оскільки збіжність методу зовнішніх штрафних функцій забезпечується при $\lambda_k \rightarrow +\infty$, виникає питання, чи можна отримати розв'язок початкової задачі на умовний екстремум (3.1) в результаті однократного розв'язання задачі безумовної мінімізації допоміжної функції $F(x, \lambda_k)$, задавши зразу велике значення параметра штрафу λ_k , наприклад, поклавши $\lambda_0 = 10^{20}$? Однак така заміна послідовного розв'язання допоміжних задач безумовної оптимізації не вважається можливою, оскільки із зростанням λ_k функція $F(x, \lambda_k)$ набуває яскраво вираженої яружної структури. Тому швидкість збіжності будь-якого методу безумовної мінімізації до вирішення допоміжної задачі різко падає.

3.2. Лабораторна робота "Мінімізація функції двох змінних методом зовнішніх штрафних функцій

Мета роботи – застосувати алгоритм методу зовнішніх штрафних функцій для пошуку точки мінімуму цільової функції $f(x)$.

Проміжні результати пошуку представити у вигляді таблиці 3.1.

Кінцевими результатами є знайдена точка мінімуму x^* і мінімальне значення цільової функції $f(x^*)$.

Таблиця 3.1 – Результати пошуку мінімуму методом зовнішніх штрафних функцій

№ ітерації k	x_k	x_{k+1}	$f(x_k)$	$f(x_{k+1})$	$\phi_1(x_{k+1})$	$ \phi_2(x_{k+1}) $
0						
1						
...						

Контрольні питання

1. Дайте визначення *штрафної функції*.
2. Дайте визначення *штрафу* або *індикаторної функції*.
3. Сформулюйте умови застосування методу *зовнішніх штрафних функцій*.
4. Дайте визначення *функції зрізання*.
5. Як формується послідовність чисел $\{\lambda_k\}$ в методі *зовнішніх штрафних функцій*?
6. Наведіть приклади *індикаторних функцій* в методі *зовнішніх штрафних функцій*?
7. Які умови закінчення пошуку в методі *зовнішніх штрафних функцій*?
8. Які вимоги до завдання початкової точки x_0 в методі *зовнішніх штрафних функцій*?
9. Чому не можна зразу задати дуже велике λ_0 , щоб за одну ітерацію визначити точку мінімуму?

3.3 Метод проекції градієнта

Метод проекції градієнта є узагальненням градієнтного методу безумовної оптимізації на випадок задач оптимізації при обмеженнях.

Розглянемо задачу

$$f(x) = f(x), \quad x \in X, \quad (3.10)$$

де $X \subset E^n$ - замкнена опукла множина, функція $f(x)$ є диференційованою в X .

❖ *Визначення.*

Проекцією $P_X(a)$ точки a на множину X називається точка $x \in X$, відстань якої до точки a є мінімальною:

$$\|x - a\|^2 \rightarrow \min, \quad x \in X. \quad (3.11)$$

Твердження:

Нехай X - замкнена опукла множина в E^n . Тоді:

- 1) проекція $P_X(a)$ будь-якої точки $a \in E^n$ на множину X існує і єдина;
- 2) для того, щоб точка $x \in X$ була проекцією точки $a \in E^n$ необхідно і достатньо, щоб для усіх $y \in X$ виконувалась нерівність

$$\langle x - a, x - y \rangle \geq 0.$$

З цього твердження витікає, що $\|x - a\| \leq \|x - y\|$;

- 3) для будь-яких точок $a_1, a_2 \in E^n$ і їх проекцій $P_X(a_1)$, $P_X(a_2)$ на множину X справедлива нерівність

$$\|P_X(a_1) - P_X(a_2)\| \leq \|a_1 - a_2\|,$$

тобто оператор проектування має властивість нерозтягнення відстаней.

В методі проекції градієнту релаксаційна послідовність формується за формулою

$$x_{k+1} = P_X(x_k - \alpha_k f'_k), \quad (3.12)$$

де величина кроку α_k знаходиться, як і раніше, на основі розв'язання задачі одновимірної мінімізації.

Отже, в описаному методі на кожній ітерації треба виконувати операцію проектування точки $a = (x_k - \alpha_k f'_k)$ на множину X , тобто розв'язувати задачу (3.11). Для деяких випадків множини X можна вказати явний вигляд проекції:

- 1) якщо $X = \{x \in E^n : \|x - d\| \leq R\}$ - куля, то

$$P_X(a) = d + \frac{a - d}{\|a - d\|} R;$$

- 2) якщо $X = \{x \in E^n : b_i \leq x_i \leq c_i, i = \overline{1, n}\}$ - n -вимірний паралелепіпед, то

$$(P_X(a))_i = \begin{cases} b_i, & a_i < b_i \\ a_i, & b_i \leq a_i \leq c_i \\ c_i, & a_i > c_i \end{cases};$$

3) якщо $X = \{x \in E^n: x_i \geq 0, i = \overline{1, n}\}$, то додатний ортант, то

$$P_X(a) = (\max(0, a_1), \max(0, a_2), \dots, \max(0, a_n));$$

4) якщо $X = \{x \in E^n: \langle p, x \rangle \geq \beta, (p \neq 0)\}$ - півпростір, то

$$P_X(a) = a + \max(0, \beta - \langle p, a \rangle) \frac{p}{\|p\|^2};$$

5) якщо $X = \{x \in E^n: Ax = b\}$ - гіперплощина, при чому рядки матриці A лінійно незалежні, то

$$P_X(a) = a - A^T(AA^T)^{-1}(Aa - b),$$

де A^T - транспонована матриця.

Метод проекції градієнта Розена.

Напрямок найшвидшого спуску цільової функції в поточній точці визначається напрямком її антиградієнта. Але при наявності обмежень пошук вздовж напрямку найшвидшого спуску може привести до виходу за межі допустимої області. В методі проекції градієнта Розена процедура пошуку мінімуму базується на побудованні такої послідовності точок пошуку, яка є релаксаційною і в той же час не призводить до появи недопустимих точок.

Нехай обмеження в задачі (3.10) є лінійними нерівностями і рівностями:
 $X = \{x: \phi_i(x) \geq 0, i = \overline{1, m}, \phi_i(x) = 0, i = \overline{m+1, l}\}$. Перепишемо їх у вигляді:

$$X = \{x: \langle a_i, x \rangle \geq b_i, i \in I_1, \langle a_i, x \rangle = b_i, i \in I_2\}, \quad (3.13)$$

де $a_i, i = \overline{1, l}$ - вектори з простору E^n , $I_1 = \{i: i = \overline{1, m}\}$, $I_2 = \{i: i = \overline{m+1, l}\}$.

Нехай x_k - деяка допустима точка і $I_0(x_k) = \{i: \langle a_i, x_k \rangle = b_i, i \in I_1\} \cup I_2$ - множина індексів обмежень, які є *активними* (тобто виконуються як рівності) в точці x_k . Знайдемо напрямок p_k ($\|p_k\| = 1$), який дозволяє мінімізувати $f'_k, p_k > \epsilon$ і не вийти за межі допустимої області. Отримаємо, що напрямок p_k повинен задовольняти умовам:

$$\langle a_i, p_k \rangle \geq 0, \quad \forall i \in I_0(x_k). \quad (3.14)$$

Позначимо через A_0 підматрицю матриці обмежень A , яка складається з $i \in I_0(x_k)$ рядків матриці A . Припустимо, що ранг матриці A_0 дорівнює q , що відповідає кількості елементів множини $I_0(x_k)$.

Твердження.

Якщо матриця A_0 має розмірність $q \times n$ і ранг $q \leq n$, то оптимальним розв'язком задачі

$$\begin{aligned} \min \langle f'_k, p_k \rangle &> \epsilon \\ A_0 p_k &= 0, \quad \|p_k\| = 1 \end{aligned}$$

є вектор $p_k = \frac{y}{\|y\|}$, де y - проекція вектора антиградієнта $-f'_k$ на гіперплощину $S_0 = \{y: A_0 y = 0\}$ і визначається за формулою

$$y = -P_0 f'_k = -(E - A_0^T (A_0 A_0^T)^{-1} A_0) f'_k.$$

Матриця $P_0 = E - A_0^T (A_0 A_0^T)^{-1} A_0$ називається *матрицею проектування*.

Алгоритм методу проекції градієнта Розена.

Вхідними даними є цільова функція двох змінних $f(x)$, $x = (x_1, x_2)$, лінійні обмеження $\phi_1(x) \geq 0$, $\phi_2(x) = 0$, початкова точка x_0 , точність σ , номер ітерації $k=0$.

Вихідними даними є точка мінімуму x^* і значення мінімуму цільової функції $f(x)$.

1) Визначити множину $I_0(x_k)$ індексів активних обмежень та множину $J_0(x_k)$ індексів неактивних обмежень.

2) Нехай A_0 - матриця, рядки якої відповідають обмеженням $i \in I_0(x_k)$, а A_1 - матриця, рядки якої відповідають обмеженням $i \in J_0(x_k)$. Якщо A_0 порожня, то покласти $P_0 = E$. Інакше обчислити матрицю проектування

$$P_0 = E - A_0^T (A_0 A_0^T)^{-1} A_0$$

та вектор напрямку пошуку $p_k = -P_0 f'_k$.

Якщо $p_k = 0$, перейти до 4).

3) Якщо $p_k \neq 0$, то знайти

$$\left(\frac{e_i}{d_i}, d_i > 0 \right), d > 0 \infty, d \leq 0, \quad (3.15)$$

де $e = (e_1, e_2) = b_1 - A_1 x_k$, $d = (d_1, d_2) = b_1 - A_1 p_k$. Потім знайти таке x_{k+1} , що

$$f(x_k + \alpha p_k), \alpha = \alpha_{\max}.$$

Покласти $x_{k+1} = x_k + \alpha_k p_k$, $k = k+1$ та перейти до кроку 1).

4) Якщо $p_k = 0$ і матриця A_0 порожня, то зупинитися. Нехай

$$u = -(A_0 A_0^T)^{-1} A_0 f'(x_k).$$

Якщо $u \geq 0$, то зупинитися: x_k задовольняє умовам Куна-Такера. У протилежному випадку нехай u_i - найбільша за модулем з від'ємних компонент вектора u . Покласти $I_0(x_k) = I_0(x_k) - \{i\}$ та перейти до 1).

➤ **Зауваження.**

При реалізації методу проекції градієнта Розена на ПК можуть виникати додаткові труднощі. Наприклад, навіть, якщо $p_k = -P_0 f'_k = 0$, то при наявності помилок округлення вектор p_k , який обчислений на комп'ютері, може відрізнитися від нульового. Тоді потрібно використовувати інше правило

переходу до викреслювання рядку з A_0 . Одне з правил таке: завжди потрібно обчислювати $f'(x_k)p(x_k)$, $p(x_k) = -P_0 f'(x_k)$, де \bar{A}_0 може бути отримана з A_0 шляхом викреслювання рядка, який відповідає найбільшому за модулем від'ємному u_i , коли є хоча б одне від'ємне u_i . Потім треба вибрати напрямок, який відповідає більшому з цих двох чисел.

3.4. Лабораторна робота "Мінімізація функції двох змінних методом проекції градієнта Розена

Мета роботи – застосувати алгоритм методу проекції градієнта Розена для пошуку точки мінімуму цільової функції $f(x)$.

Проміжні результати пошуку представити у вигляді таблиці 3.2.

Кінцевими результатами є знайдена точка мінімуму x^* і мінімальне значення цільової функції $f(x^*)$.

Таблиця 3.2 – Результати пошуку мінімуму методом проекції градієнта Розена

№ ітерації k	x_k	x_{k+1}	$f(x_k)$	$f(x_{k+1})$	$\phi_1(x_{k+1})$	$ \phi_2(x_{k+1}) $
0						
1						
...						

Контрольні питання

1. Сформулюйте задачу мінімізації в умовах обмежень.
2. Яке обмеження називається активним в точці x_k ?
3. Чи може оптимальне значення цільової функції $f(x)$ в задачі умовної мінімізації бути менше оптимального значення цієї ж функції в задачі безумовної мінімізації? Поясніть.
4. Дайте визначення *проекції* $P_X(a)$ точки a на множину X .
5. Сформулюйте необхідні і достатні умови того, щоб точка $x \in X$ була проекцією точки $a \in E^n$ на множину X .
6. Запишіть умову того, що оператор проектування має властивість нерозтягнення відстаней.
7. Запишіть формулу для *матриці проектування* P_0 у випадку лінійних обмежень.
8. Як формується в методі проекції градієнту релаксаційна послідовність?
9. Дайте класифікацію методу Розена.

3.5. Метод можливих (допустимих) напрямків Зойтендейка

Для кожної допустимої точки $x_k \in X$ може існувати безліч різноманітних допустимих напрямків пошуку. Сутність методів допустимих напрямків полягає в наступному: пошук починається в допустимій точці x_k простору розв'язків і реалізується (при лінеаризованих обмеженнях) вздовж траєкторії, яка забезпечує покращення значення цільової функції і разом з цим ніколи не виходить за межі допустимої області. В цьому сенсі розглянутий метод проекції градієнта Розена є одною з модифікацій методів допустимих напрямків. Однак, слід звернути увагу на наступне: проекція градієнта $f'(x_k)$ на підмножину обмежень X , що містить x_k , задає напрямок пошуку, який відповідає (відносно евклідової метрики) напрямку найскорішого спуску для цільової функції, тоді як в методі Зойтендейка використовується інша метрика, яка визначається співвідношенням

$$\|\Delta x\| = \max\{|x_1|, |x_2|, \dots, |x_n|\}.$$

Отже, це кубічна метрика вектора Δx .

Застосування кубічної метрики призводить до того, що в якості допустимого напрямку вибирається саме такий, який забезпечує найбільшу мінімізуючу поправку до значення цільової функції без порушення будь-якого з обмежень.

Метод допустимих напрямків Зойтендейка.

Метод допустимих напрямків Зойтендейка дозволяє оперувати як з лінійними, так і з нелінійними обмеженнями; однак за його допомогою не вдається розв'язувати задачі з обмеженнями у вигляді рівностей. Нехай обмеження мають вигляд нерівностей: $X = \{x \mid g_i(x) \geq 0, i = \overline{1, m}\}$.

В результаті лінеаризації присутніх в задачі умовної мінімізації функції шляхом розкладання їх в ряд Тейлора в околі точки x_k прийдемо до наступної задачі:

$$\min_{x \in E^n},$$

при обмеженнях

$$g_i(x_k) + <, i = \overline{1, m}. \quad (3.16)$$

Процедура Зойтендейка встановлює найбільш сприятливі з можливих напрямків оптимізаційного пошуку в точці x_k (в сенсі отримання найбільшої мінімізуючої поправки до значення цільової функції $f(x)$ при переміщенні із точки x_k в точку x_{k+1} без порушення будь-якого з обмежень) шляхом розв'язання деякої підзадачі, пов'язаної з початковою задачею.

Переміщення з точки x_k в точку x_{k+1} , як звичайно, задається співвідношенням $x_{k+1} = x_k + \alpha_k p_k$. Підстановка в задачі (3.16), (3.17) замість вектора x його поточного значення x_{k+1} призводить до задачі

$$\min,$$

при обмеженнях

$$g_i(x_k) + \alpha_k < g_i'(x_k), p(x_k) > 0, \quad i = \overline{1, m}. \quad (3.17)$$

Оскільки $f(x_k)$ і $g_i(x_k) \geq 0$ є постійними, то необхідні і достатні умови реалізації переміщення з точки x_k в точку x_{k+1} на основі розв'язання задачі (3.17) з урахуванням того, що точка x_{k+1} залишається допустимою для задачі (3.16)), мають вигляд:

$$f'(x_k), p(x_k) > < 0, \quad (3.18)$$

$$g_i'(x_k), p(x_k) > \geq 0 \quad i = \overline{1, m}. \quad (3.19)$$

Будь-який вектор $p(x_k)$, який задовольняє умовам (3.18), (3.19), буде в той же час і вектором допустимого напрямку пошуку.

В методі Зойтендейка допустимий напрямок пошуку, що забезпечує отримання найбільшої мінімізуючої поправки до значення цільової функції $f(x)$ при переміщенні з точки x_k в точку x_{k+1} , знаходиться на основі розв'язання наступної задачі лінійного програмування:

$$\min < f'(x_k), p(x_k) >, \quad p(x_k) \in E^n \quad (3.20)$$

при обмеженнях

$$g_i'(x_k), p(x_k) > \geq 0 \quad i = \overline{1, m}. \quad (3.21)$$

Алгоритм методу допустимих напрямків Зойтендейка.

Вхідними даними є цільова функція двох змінних $f(x)$, $x = (x_1, x_2)$, лінійне обмеження-нерівність $g(x) \geq 0$, початкова точка x_0 , точність σ , номер ітерації $k=0$.

Вихідними даними є точка мінімуму x^* і значення мінімуму цільової функції $f(x^*)$.

1) Обчислити градієнти $f'(x_k)$ і $g_i'(x_k)$.

2) Розв'язати задачу лінійного програмування

$$\min \langle f'(x_k), p(x_k) \rangle, p(x_k) \in E^n \quad \text{при обмеженнях} \quad g'(x_k), p(x_k) \geq 0.$$

Визначити допустимий напрямок $p(x_k)$.

3) Якщо $\langle f'(x_k), p(x_k) \rangle \geq 0$, то пошук закінчено і $x = x_k$.

Інакше, якщо $\langle f'(x_k), p(x_k) \rangle < 0$, визначити максимально можливу довжину кроку α в напрямку $p(x_k)$ без виходу за межі допустимої області на основі розв'язання задачі одновимірної оптимізації:

$$\alpha = \max \{ \alpha \mid (x_k + \alpha p_k) \in X \}, \quad \text{де } X = \{x \mid g(x) \geq 0\}.$$

4) Розв'язати задачу одновимірної мінімізації з точністю σ :

$$f(x_k + \alpha p_k),$$

визначити довжину кроку α_k .

5) Знайти $x_{k+1} = x_k + \alpha_k p_k$. Покласти $k = k + 1$, перейти до 1).

3.6. Лабораторна робота "Мінімізація функції двох змінних методом допустимих напрямків Зойтендейка"

Мета роботи – застосувати алгоритм методу допустимих напрямків Зойтендейка для пошуку точки мінімуму цільової функції $f(x)$.

Проміжні результати пошуку представити у вигляді таблиці 3.3.

Кінцевими результатами є знайдена точка мінімуму x і мінімальне значення цільової функції $f(x)$.

Таблиця 3.3 – Результати пошуку мінімуму методом допустимих напрямків Зойтендейка

№ ітерації k	x_k	$f(x_k)$	α	α_k	p_k	$f'_k, p_k > \epsilon$	$g(x_k)$
0							
1							
...							

Контрольні питання

1. Дайте визначення допустимим напрямкам в задачі умовної мінімізації.
2. Яка метрика використовується в методі Зойтендейка?
3. Який вигляд мають обмеження в задачі мінімізації методом Зойтендейка?
4. Як записується лінеаризована задача в методі Зойтендейка?

5. Сформулюйте задачу лінійного програмування, що розв'язується в методі Зойтендейка.

6. Запишіть умову закінчення пошуку в алгоритмі Зойтендейка.

СПИСОК ЛІТЕРАТУРИ

1. Барвінський А.Ф. Математичне програмування/А.Ф.Барвінський, І.Я. Олексів та ін. – Львів: НУ "ЛП", 2004.– 468 с.

2. Кісельова О.М. Чисельні методи оптимізації: навч. посіб./ О.М.Кісельова, А.Є.Шевельова.-Д.:Вид-во ДНУ, 2008.– 208 с.

2. Himmelblau David M. Applied Nonlinear Programming. – The University of Texas, Austin, Texas. McGraw-Hill Book Company, 1972.

3. Уханська О.М. Тексти лекцій з курсу "Методи оптимізації". – Львів: НУ "ЛП", 2003.– 107 с.

4. Вітлінський В.В., Наконечний С.І., Терещенко Т.О. Математичне програмування. – К: КНЕУ, 2001.– 248 с.

5. Жалдак М.І. Основи теорії і методів оптимізації: навч. посіб./ М.І.Жалдак, Ю.В.Триус.-Черкаси:Брама-Україна,2005.–306 с.

6. Бейко И.В. Методы и алгоритмы решения задач оптимизации / И.В. Бейко, Б.Н. Бублик, П.Н. Зинько. – К.: Вища школа, 1983. – 512 с.

7. Попов Ю.Д. Методи оптимізації: навч. електронний посіб./Ю.Д.Попов, В.І.Тюптя, В.І.Шевченко.- К.:2003.–215 с.

8. Сікора Я.Б. Методи оптимізації: навч. посіб./Житомир: Вид-во ЖДУ ім І.Франка, 2012.– 82 с.

9. Мовчан А.П. Методи статичної оптимізації: навч. посіб./А.П.Мовчан, О.В.Степанець.-К.: НТТУ "КП", 2012.– 138 с.

10. Плаксі́й Ю.А. Методи оптимізації функцій. Частина 1. Методи мінімізації функцій однієї змінної: навч.-метод. посіб./ Ю.А.Плаксі́й, О.А.Татарінова. – Харків, НТУ «ХП».-2016.– 79 с.

11. Плаксі́й Ю.А., Успенський В.Б. Методичні вказівки до лабораторно-практичного заняття «Вивчення методів мінімізації функцій багатьох змінних» з курсу «Методи оптимізації», Харків, ХДПУ, 1999.–16 с.

12. Reklaitis G.V., Ravindran A., Ragsdell K.M. Engineering Optimization. Methods and Applications. – Aerospace and Mechanical Engineering University of Arizona. A Wiley-Interscience Publication. John Wiley & Sons. – New York, 1983.
13. Fletcher R. Practical methods of optimization / R. Fletcher. – 2-d edition. – UK, Chichester : John Wiley & Sons Ltd, 2000. – 436 p.
14. Forsythe G.E., Malcolm M.A. , Moler C.B. Computer Methods for Mathematical Computations. Englewood Cliffs, New Jersey 07632. Prentice Hall, Inc., 1977, XI. – 259 p.

Додаток А
ЦІЛЬОВІ ФУНКЦІЇ І ОБМЕЖЕННЯ

Таблиця А1. Цільові функції і початкова точка

№ варіанта	Цільова функція $f(x,y)$	Початкова точка x_0
1	$12x^2+y^2-4xy-2y+5$	(-1, -2)
2	$x^2+10y^2-3xy-4x+2y-1$	(2, -1)
3	$3x^2+6y^2+2xy-x-3y$	(0, -1)
4	$6x^2+y^2+2xy-2x-y$	(0, -1)
5	$6x^2+y^2-x+2y-3$	(1, -2)
6	$8x^2+y^2-3xy+2x-y+4$	(-1, 0)
7	$5x^2+2y^2+3xy-x-y+2$	(-1, -1)
8	$6x^2+y^2+2xy-x-y$	(1, 0)
9	$8x^2+y^2+xy-x+y$	(-1, -1)
10	$6x^2+y^2-2xy-x-y$	(1, 1)
11	$6x^2+y^2+xy-2x-y$	(-1, 0)
12	$5x^2+2y^2-2xy-x+6y-3$	(-1, -1)
13	$8x^2+y^2-4xy+3x-y+4$	(1, 0)
14	$9x^2+2y^2+xy-x-y+2$	(-1, 1)
15	$14x^2+y^2-2xy-3x+y+2$	(0, -1)
16	$11x^2+y^2-2xy-2x-y+1$	(1, 1)
17	$8x^2+y^2+xy-x-2y$	(1, -1)
18	$6x^2+y^2+2xy-3x+2y-2$	(1, 0)
19	$10x^2+4y^2-2xy+2x-y-1$	(-1, 1)
20	$10x^2+y^2+xy-12x-y$	(-1, 2)
21	$4x^2+y^2-xy-6y-4x+3$	(0, -1)
22	$2x^2+y^2+xy-x-y$	(1, -1)
23	$3x^2+y^2+xy-x+y$	(-1, -1)
24	$4x^2+y^2-2xy-x-y$	(1, 0)
25	$6x^2+y^2+xy-2x-y$	(0, -1)
26	$5x^2+y^2-x+2y-3$	(1, -2)
27	$8x^2+y^2-4xy+2x-y+4$	(-1, 0)
28	$4x^2+2y^2+3xy-x-y+2$	(-1, -1)
29	$5x^2+y^2+xy-x-y$	(1, 0)
30	$3x^2+y^2+xy-x+y+1$	(-1, -1)
31	$7x^2+y^2-2xy-x-y+2-3$	(1, 1)
32	$7x^2+y^2+xy-2x-y-1$	(-1, 0)
33	$5x^2+y^2-2xy-x+6y-3$	(-1, -1)
34	$8x^2+y^2-5xy+3x-y+4$	(1, 0)
35	$4x^2+y^2-3xy-x+y$	(1, 1)
36	$5x^2+2y^2-3xy-4x+2y-3$	(0, -1)
37	$3x^2+2y^2+3xy-x+y$	(-1, -1)
38	$2x^2+2y^2+3xy-x-2y$	(-1, 0)

Таблиця А2. Обмеження на змінні

№ варіанта	$\phi_1(x) \geq 0$	$\phi_2(x) = 0$
1	$1 - 2x \geq 0$	$x - 2y = 0$
2	$2 + 3x \geq 0$	$x + 2y = 0$
3	$2 - x \geq 0$	$1,5x - y = 0$
4	$2 + x \geq 0$	$-1 + x + y = 0$
5	$x - 1 \geq 0$	$1,2 - x - y = 0$
6	$x - 2 \geq 0$	$1,3 + x - y = 0$
7	$1 + y \geq 0$	$1,2 - x - 2y = 0$
8	$1 - y \geq 0$	$-1,5 + x + y = 0$
9	$y - 1 \geq 0$	$2 - x - y = 0$
10	$y - 2 \geq 0$	$1,4 - x + y = 0$
11	$1,5 + x \geq 0$	$1 - 2x + y = 0$
12	$1,5 - x \geq 0$	$-1 + x - 2y = 0$
13	$2,2 + x \geq 0$	$2 - x + 2y = 0$
14	$2,1 - x \geq 0$	$2,4 + x + y = 0$
15	$x - 1,3 \geq 0$	$-1 + x - 3y = 0$
16	$x - 2,2 \geq 0$	$-x - 2y + 2 = 0$
17	$2,1 + y \geq 0$	$2x - y + 2 = 0$
18	$3,1 - y \geq 0$	$x - 2y + 4 = 0$
19	$y - 1,5 \geq 0$	$x + y - 1,8 = 0$
20	$y - 2,1 \geq 0$	$x - y - 3 = 0$
21	$1 + 2x \geq 0$	$x + y - 4 = 0$
22	$1 + x \geq 0$	$1 + x + 2y = 0$
23	$1 - x \geq 0$	$1 - 2x + y = 0$
24	$2 + x \geq 0$	$1,5 + x - y = 0$
25	$2 - x \geq 0$	$-1 + x + y = 0$
26	$x - 1 \geq 0$	$1,2 - x - 2y = 0$
27	$x - 2 \geq 0$	$1,3 + 2x - y = 0$
28	$1 + x \geq 0$	$-1,2 + x - y = 0$
29	$1 - y \geq 0$	$1,5 + x + y = 0$
30	$y - 1 \geq 0$	$2,4 - x - 2y = 0$
31	$y - 2 \geq 0$	$1,4 - x + y = 0$
32	$1,5 + x \geq 0$	$1 - 2x + y = 0$
33	$1,5 - x \geq 0$	$-1 + x - 2y = 0$
34	$2,2 + x \geq 0$	$2 - x + 2y = 0$
35	$2 - x \geq 0$	$1 - 2x + y = 0$
36	$1 - 2x \geq 0$	$-1 + x + y = 0$
37	$1,2 + x \geq 0$	$1 + x - y = 0$
38	$1 + x + y \geq 0$	$2 - x + 2y = 0$

Навчальне видання

МЕТОДИЧНІ ВКАЗІВКИ
до виконання лабораторних робіт
«МЕТОДИ МІНІМІЗАЦІЇ ФУНКЦІЙ БАГАТЬОХ ЗМІННИХ»
з курсу «Методи оптимізації»
для студентів денної та заочної форми навчання
спеціальності 122 «Комп'ютерні науки»

Укладач:
ПЛАКСІЙ Юрій Андрійович

Відповідальний за випуск доц. Татарінова О. А.
Роботу до видання рекомендував проф. Бреславський Д. В.

В авторській редакції

План 2023р., поз. 309

Підп. до друку 2023 р.

Гарнітура Times New Roman

Видавничий центр НТУ «ХП».
Свідоцтво про державну реєстрацію ДК № 5478 від 21.08.2017 р.
61002, Харків, вул. Кирпичова, 2
