

Дослідження чотирикрокового методу мінімізації функцій багатьох змінних

Виконала: студентка денної форми навчання
спеціальності 8.04030101 Прикладна математика
Цимбал Марина Ігорівна

Керівник: к. ф.-м. н., доц. Яровий А. Т.
Рецензент: к. ф.-м. н., доц. Страхов Є. М.

Передімною було поставлено завдання:

1. Побудувати алгоритм для чотирикрокового методу мінімізації функцій;
2. Перевірити його доцільність на тестових функціях;
3. Порівняти 3-х і 4-х крокові методи мінімізації функцій.

Що я зробила:

Я побудувала алгоритм для чотирикрокового методу мінімізації функцій та обґрунтувала його належність до методів спряжених напрямків. Далі я перевірила алгоритм на тестових функціях і порівняла його з трикроковим методом мінімізації функцій.

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.

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 градієнтними методами, бо він у більшій мірі враховує геометричні властивості цільової функції. Опираючись на цю інформацію, можна піти далі і розглянути трикрокові, чотирікрокові, п'ятикрокові і т.д. методи.

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

Алгоритм чотирикрокового методу

Розглянемо задачу мінімізації функції багатьох змінних без обмежень

$$\phi(x) \rightarrow \min, x \in \mathbb{R}^n, \quad (1)$$

де

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

Позначимо її градієнт через $\phi'(x)$. А для її розв'язання розглянемо чотирикроковий метод з таким алгоритмом:

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

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

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

Параметр β_k будемо визначати з умови:

$$\beta_k : \min_{\beta \geq 0} \phi(x^k + \beta s^k) \quad (3)$$

Теоретичні положення методу

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

$$\phi(x) = \frac{1}{2}(Ax, x) + (b, x) + c \quad (4)$$

Побудуємо систему взаємно спряжених напрямків за правилом (2).

$$\begin{aligned} 0 = (s^k, As^{k-1}) &= -(\phi'(x^k), As^{k-1}) + \gamma_1^{k-1}(s^{k-1}, As^{k-1}) + \\ &+ \gamma_2^{k-2} \underbrace{(s^{k-2}, As^{k-1})}_0 + \gamma_3^{k-3} \underbrace{(s^{k-3}, As^{k-1})}_0 \Rightarrow \\ \gamma_1^{k-1} &= \frac{(\phi'(x^k), As^{k-1})}{(s^{k-1}, As^{k-1})} \end{aligned} \quad (5)$$

Аналогічно отримаємо:

$$\gamma_2^{k-2} = \frac{(\phi'(x^k), As^{k-2})}{(s^{k-2}, As^{k-2})} \quad (6)$$

$$\gamma_3^{k-3} = \frac{(\phi'(x^k), As^{k-3})}{(s^{k-3}, As^{k-3})} \quad (7)$$

Теорема 1.

Для диференційовної функції $\phi(x)$ послідовність $\{x^k\}$, що побудована за (2), (5), (6), (7), така, що $(\phi'(x^{k+1}), s^k) = 0$, $k = 0, 1, \dots$

Теорема 2.

Вектори $\phi'(x^k)$ і $\phi'(x^{k+1})$ ортогональні, $k = 0, 1, \dots$

Теорема 3.

Нехай $x^0 \in \mathbb{R}^n$, точки x^1, x^2, \dots, x^{n-1} і вектори s^0, s^1, \dots, s^{n-1} отримані за формулами (2), (5), (6), (7) і $\phi'(x^k) \neq 0$ ($k = \overline{0, n-1}$), тоді вектори s^0, s^1, \dots, s^{n-1} взаємно спряжені, а градієнти $\phi'(x^0), \phi'(x^1), \dots, \phi'(x^{n-1})$ взаємно ортогональні.

Отже розглянутий чотирьохкроковий метод (2), (5), (6), (7) належить до методів спряжених напрямків.

Тепер сформулюємо чотирикроковий метод для мінімізації неквадратичних функцій.

Для цього перетворимо формули (5), (6), (7) так, щоб до них не входила матриця A .

$$\gamma_1^{k-1} = \frac{(\phi'(x^k), As^{k-1})}{(s^{k-1}, As^{k-1})} = \frac{(\phi'(x^k), \phi'(x^k) - \phi'(x^{k-1}))}{\|\phi'(x^{k-1})\|^2} \quad (8)$$

$$\gamma_2^{k-2} = \frac{(\phi'(x^k), As^{k-2})}{(s^{k-2}, As^{k-2})} = \frac{(\phi'(x^k), \phi'(x^{k-1}) - \phi'(x^{k-2}))}{\|\phi'(x^{k-2})\|^2} \quad (9)$$

$$\gamma_3^{k-3} = \frac{(\phi'(x^k), As^{k-3})}{(s^{k-3}, As^{k-3})} = \frac{(\phi'(x^k), \phi'(x^{k-2}) - \phi'(x^{k-3}))}{\|\phi'(x^{k-3})\|^2} \quad (10)$$

Отже, для неквадратичних функцій, чотирикроковий метод має вигляд (2), (8), (9), (10)

Обчислювальний експеримент

Зробимо аналіз роботи описаного чотирикрокового методу порівнянно із трикроковим методом. Для порівняння використаємо такі показники, як точність отриманого 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:

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

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

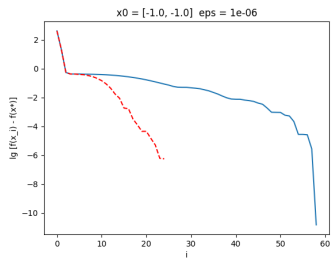
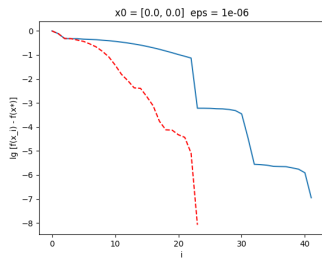
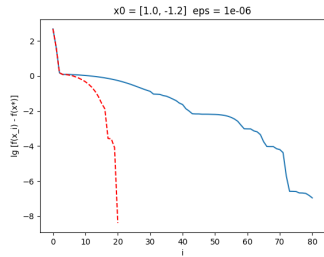
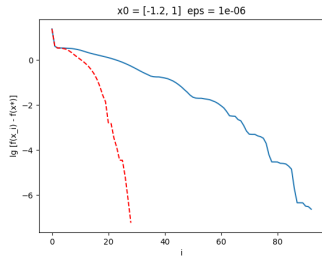
$$\|\phi'(x^k)\| \leq \sqrt[3]{\varepsilon}(1 + |\phi(x^k)|)$$

Пунктирною лінією позначений процес мінімізації трикроковим методом, суцільною - чотирикроковим.

Задача 1. $\phi(x) = (-x_1 + 1)^2 + 100(-x_1^2 + x_2)^2$

Точний розв'язок задачі: $x^* = [1, 1]$ $f^* = 0$

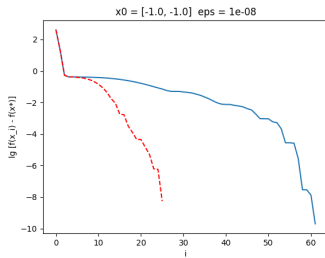
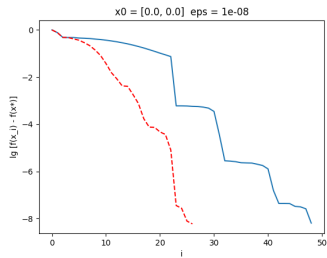
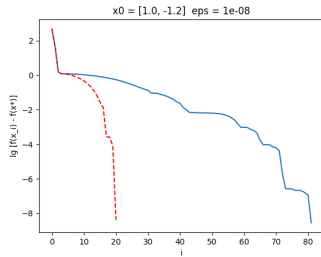
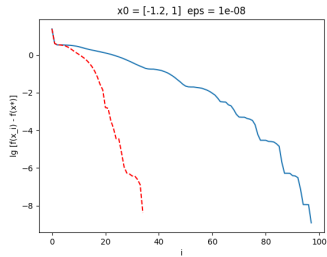
ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
10^{-6}	[-1.2, 1]	24.2	4 кроковий	[1.00027 1.00053]	0	93
			3 кроковий	[1.00069 1.00137]	0.0	29
	[1, -1.2]	484	4 кроковий	[0.99995 0.99991]	0	81
			3 кроковий	[0.99994 0.99988]	0	21
	[0, 0]	1	4 кроковий	[1.00015 1.00029]	0	42
			3 кроковий	[1.00014 1.00029]	0	24
	[-1, -1]	404	4 кроковий	[0.99998 0.99995]	0	59
			3 кроковий	[0.99992 0.99985]	0	25



Задача 1. $\phi(x) = (-x_1 + 1)^2 + 100(-x_1^2 + x_2)^2$

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^{-8}	[-1.2, 1]	24.2	4 кроковий	[0.99999 0.99999]	0	98
			3 кроковий	[0.99997 0.99994]	0	35
	[1, -1.2]	484	4 кроковий	[1. 0.99999]	0	32
			3 кроковий	[1. 1.]	0	25
	[0, 0]	1	4 кроковий	[0.99993 0.99985]	0	24
			3 кроковий	[1. 1.]	0	19
	[-1, -1]	404	4 кроковий	[0.99998 0.99996]	0	20
			3 кроковий	[1. 1.]	0	16

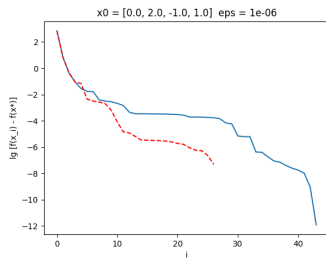
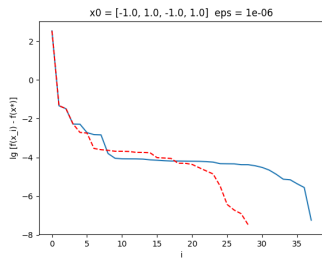
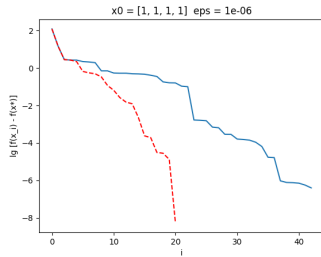
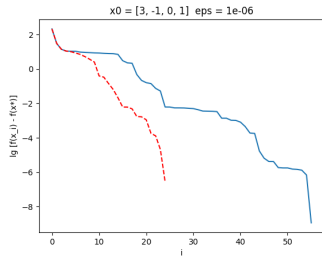


Задача 2. $\phi(x) =$

$$(x_1 + 10x_2)^2 + 10(x_1 - x_4)^4 + (x_2 - 2x_3)^4 + 5(x_3 - x_4)^2$$

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	[0.05 -0.005 0.03 0.03]	3e-05	36
			3 кроковий	[0.0008 -0.00008 -0.008 -0.003]	0.0	36
	[1, 1, 1, 1]	122	4 кроковий	[-0.069 0.0069 -0.04 -0.0411]	7e-05	43
			3 кроковий	[-0.048 0.004 -0.025 -0.025]	1e-05	21
	[-1, 1, -1, 1]	342	4 кроковий	[-0.086 0.008 -0.048 -0.048]	0.00014	38
			3 кроковий	[0 -0 -0.02 -0.01]	0.0	29
	[0, 2, -1, 1]	686	4 кроковий	[0.002 -0.0002 0.0015 0.0015]	0.0	44
			3 кроковий	[-0.03 0.004 -0.0096 -0.0096]	1e-05	27

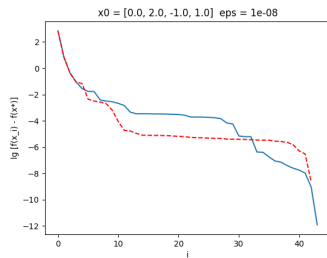
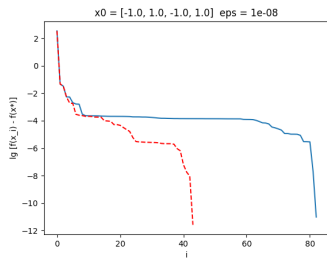
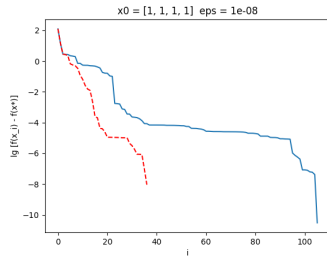
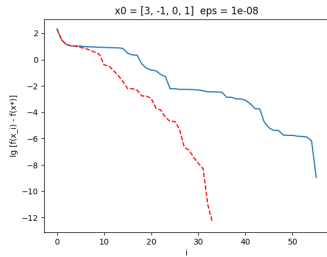


Задача 2. $\phi(x) =$

$$(x_1 + 10x_2)^2 + 10(x_1 - x_4)^4 + (x_2 - 2x_3)^4 + 5(x_3 - x_4)^2$$

Точний розв'язок задачі: $x^* = [0, 0, 0, 0]$ $f^* = 0$

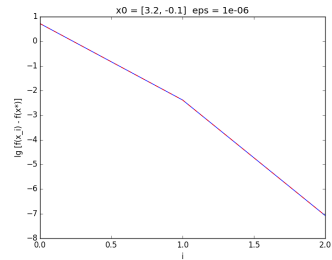
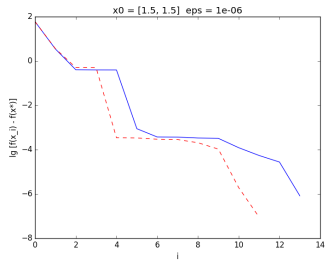
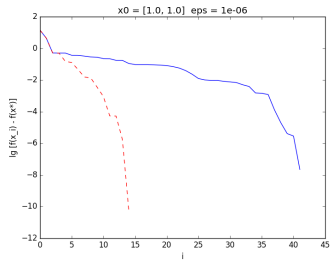
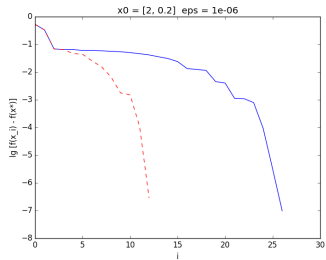
ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
10^{-8}	[3, -1, 0, 1]	215	4 кроковий	[-0.02558 0.00256 -0.01165 -0.01175]	0	56
			3 кроковий	[0.00019 -0.00002 -0.00241 -0.00241]	0	34
	[1, 1, 1, 1]	122	4 кроковий	[0.00081 -0.00008 -0.00795 -0.00794]	0	106
			3 кроковий	[-0.02441 0.00244 -0.01503 -0.01505]	0	37
	[-1, 1, -1, 1]	342	4 кроковий	[0.00905 -0.0009 -0.00682 -0.00667]	0	83
			3 кроковий	[-0.00833 0.00083 -0.00384 -0.00385]	0	44
	[0, 2, -1, 1]	686	4 кроковий	[0.00071 -0.00007 0.00963 0.00962]	0	40
			3 кроковий	[0.00041 -0.00004 -0.00281 -0.00281]	0	62



$$\text{Задача 3. } \phi(x) = (-x_1(-x_2 + 1) + 1.5)^2 + (-x_1(-x_2^2 + 1) + 2.25)^2 + (-x_1(-x_2^3 + 1) + 2.625)^2$$

Точний розв'язок задачі: $x^* = [3, 0.5]$ $f^* = 0$

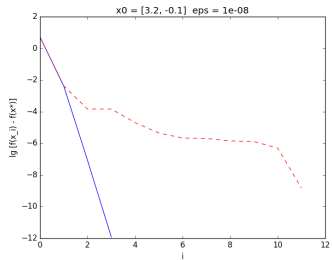
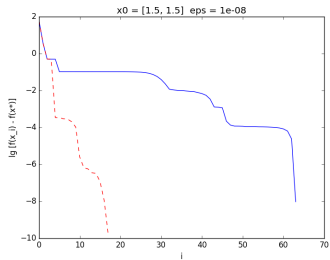
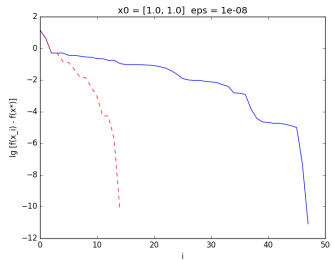
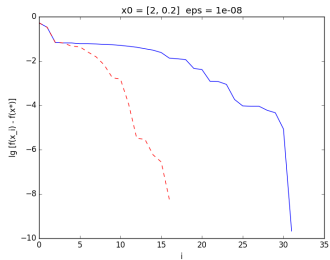
ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
10^{-6}	[2, 0.2]	0.5	4 кроковий	[3.01226 0.50475]	9e-05	27
			3 кроковий	[2.99704 0.49901]	0.0	13
	[1, 1]	14.2	4 кроковий	[2.99354 0.49768]	2e-05	42
			3 кроковий	[3.00004 0.5]	0.0	15
	[1.5, 1.5]	60.3	4 кроковий	[4.53484 0.72223]	0.10485	14
			3 кроковий	[2.99818 0.4995]	0.0	12
	[3.2, -0.1]	5.2	4 кроковий	[2.98676 0.49898]	0.00015	3
			3 кроковий	[2.98676 0.49898]	0.00015	3



$$\text{Задача 3. } \phi(x) = (-x_1(-x_2 + 1) + 1.5)^2 + (-x_1(-x_2^2 + 1) + 2.25)^2 + (-x_1(-x_2^3 + 1) + 2.625)^2$$

Точний розв'язок задачі: $x^* = [3, 0.5]$ $f^* = 0$

ε	x_0	$f(x_0)$	Метод мінімізації	x^* - отриманий розв'язок	$f(x^*)$	Кількість ітерацій
10^{-8}	[2, 0.2]	0.5	4 кроковий	[2.99785 0.49975]	0.0	32
			3 кроковий	[2.99979 0.49992]	0.0	17
	[1, 1]	14.2	4 кроковий	[2.99996 0.5]	0.0	48
			3 кроковий	[3.00004 0.5]	0.0	15
	[1.5, 1.5]	60.3	4 кроковий	[3.00028 0.50053]	0.0	64
			3 кроковий	[3.00014 0.50003]	0.0	18
	[3.2, -0.1]	5.2	4 кроковий	[2.98676 0.49898]	0.00015	4
			3 кроковий	[2.99921 0.49981]	0.0	12



Результати експериментів вказують на неефективність 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.

Дякую за увагу!