ISSN 2078–5097. Вісн. Львів. ун-ту. Сер. прикл. матем. та інф. 2013. Вип. 20. С. 3–9
Visnyk of the Lviv University. Series Appl. Math. and Informatics. Issue 20. P. 3–9

О Б Ч И С Л Ю В А Л Ь Н А   М А Т Е М А Т И К А

УДК 519.6

ТРИКРОКОВИЙ ІТЕРАЦІЙНИЙ МЕТОД МІНІМІЗАЦІЇ ФУНКЦІЙ З КУБІЧНИМ ПОРЯДКОМ ЗБІЖНОСТІ

М. Бартіш, Н. Огородник

Львівський національний університет імені Івана Франка,
вул. Університетська, 1, Львів, 79000, e-mail: ogorodnyk.nataly@gmail.com

Використовуючи ідею трикрокових алгоритмів [1], побудовано новий варіант трикрокового методу розв'язування задач безумовної мінімізації з кубічним порядком збіжності на базі методу з [3]. Теоретично обгрунтовано збіжність методу. Виконано числові експерименти та зроблено висновки про ефективність і можливість застосування запропонованого алгоритму.

Ключові слова: метод Ньютона, трикроковий метод, безумовна мінімізація.

1. ВСТУП

У різних галузях людського життя часто виникає потреба у розв'язуванні задач оптимізації. Такі задачі часто трапляються в науці, техніці, промисловості та багатьох інших галузях. Для розв'язування актуальних проблем побудовано широкі класи алгоритмів. Різні методи виявляють свою ефективність на різних класах задач. Методи мають свої переваги та недоліки: вибір початкового наближення, швидкість збіжності, трудомісткість окремої ітерації тощо. Сьогодні не існує універсального алгоритму, тому і надалі актуальною залишається проблема побудови ефективних методів [2, 4].

Ми досліджуватимемо трикрокові ітераційні методи розв'язування задач безумовної мінімізації, а також метод третього порядку з статті Т. Коган [3]. Використовуючи ідею побудови трикрокових алгоритмів із [1], запропоновано трикроковий метод на підставі методу з [3]. Проведено теоретичні дослідження методу, визначено третій порядок збіжності. При практичному застосуванні алгоритм виявив свою ефективність, у сенсі кількості обчислень перед методом з [3] і методом Ньютона [2] на різних типах задач.

2. ФОРМУЛЮВАННЯ ЗАДАЧІ

Розглянемо задачу безумовної мінімізації функції багатьох змінних

$$f(x) \to \min, \; x \in R^n. \tag{1}$$

Одним з методів розв'язування задачі (1) є метод Ньютона [2, 4]

$$x_{k+1} = x_k - \left(f''(x_k)\right)^{-1} f'(x_k), \; k = 0,1,\dots. \tag{2}$$

У [3] на підставі методу Ньютона (2) запропоновано метод третього порядку для розв'язування функціональних рівнянь. Цей метод можна використовувати для розв'язування задачі (1) у такому вигляді:

$$u_k = x_k - \left(f''(x_k)\right)^{-1} f'(x_k), \tag{3}$$

$$x_{k+1} = x_k - \left(f''\left(\frac{x_k + u_k}{2}\right)\right)^{-1} f'(x_k), \; k = 0,1,\dots. \tag{4}$$

Як і у випадку функціональних рівнянь, метод (3)–(4) розв'язування задачі (1) має третій порядок збіжності.

Використаємо ідею побудови трикрокових методів розв'язування задачі (1) з [1], а саме: на кожній ітерації обчислюємо два проміжні наближення, наступне наближення алгоритму шукаємо як точку мінімуму на прямій, що з'єднує отримані наближення. Згідно з цим підходом як проміжні наближення u_k і v_k використано метод (3)–(4).

Отримаємо такий трикроковий метод:

$$u_k = x_k - \left(f''(x_k)\right)^{-1} f'(x_k), \tag{5}$$

$$v_k = x_k - \left(f''\left(\frac{x_k + u_k}{2}\right)\right)^{-1} f'(x_k), \tag{6}$$

$$x_{k+1} = \arg\min_{\gamma} f\left(u_k + \gamma(v_k - u_k)\right),\ k = 0,1,\dots . \tag{7}$$

3. ОБҐРУНТУВАННЯ ЗБІЖНОСТІ

Теорема. Нехай:

1) $f : R^n \to R$ – тричі неперервно-диференційована і для $x \in D$, $h \in R^n$ правильні оцінки

$$m\|h\| \le \left(f''(x)h, h\right) \le M\|h\|,$$

$$\|f'''(x)\| \le K,$$

де $0 < m \le M$; $m, M, K = const$; $D = \{x \in R^n : f(x) \le f(x_0)\}$;

2) $\forall x, y \in D$, $f'''(x)$ задовольняє умову

$$\|f'''(x) - f'''(y)\| \le N\|x - y\|,$$

де $0 < N < \infty$;

3) початкове наближення x_0 вибирають таким, що виконується умова

$$q = C\left(f(x_0) - f(x_*)\right) < 1,$$

де $C = \frac{2}{m}\left(\frac{K^2}{4m^2} + \frac{3N}{2m}\right)\sqrt{\frac{M}{m}}$.

Тоді послідовність $\{x_k\}$, породжена методом (5)–(7), коректно визначена та збігається до розв'язку x_* задачі (1) і справджується оцінка

$$f(x_k) - f(x_*) \le \left(\prod_{i=0}^{k-1} \mu_{k-i}^{3^i}\right) q^{3^k - 1}\left(f(x_0) - f(x_*)\right), \tag{8}$$

де $\mu_k \in (0,1]$.

Доведення. З умови 1) теореми випливає сильна опуклість функції $f(x)$, отже, розв'язок задачі (1) x_* існує і єдиний. У цьому разі виконується умова для $\forall x \in D$

$$\frac{m}{2}\|x - x_*\|^2 \le f(x) - f(x_*) \le \frac{M}{2}\|x - x_*\|^2. \tag{9}$$

Нехай відоме деяке наближення $x_k \in D$ до розв'язку задачі. Аналогічно до результатів з [2] й умови 1) теореми можемо записати

$$\left\|\left(f''(x)\right)^{-1}\right\| \le \frac{1}{m} \tag{10}$$

і

$$f'(x_k) - f'(x_*) = \left(\int_0^1 f''(x_* + \tau(x_k - x_*))d\tau\right)(x_k - x_*). \tag{11}$$

Використовуючи оцінку (10), формули (5) і (11), отримаємо оцінку

$$\begin{aligned}\|u_k - x_*\| &= \left\|x_k - x_* - \left(f''(x_k)\right)^{-1} f'(x_k)\right\| = \\ &= \left\|\left(f''(x_k)\right)^{-1}\right\| \left\|f''(x_k)(x_k - x_*) - \left(f'(x_k) - f'(x_*)\right)\right\| \le \\ &\le \frac{1}{m}\left\|\int_0^1 \left(f''(x_k) - f''(x_* + \tau(x_k - x_*))\right)d\tau(x_k - x_*)\right\| \le \\ &\le \frac{K}{2m}\|x_k - x_*\|^2.\end{aligned} \tag{12}$$

Використовуючи (6), одержуємо

$$\begin{aligned}\|v_k - x_*\| &= \left\|x_k - x_* - \left(f''\left(\frac{x_k + u_k}{2}\right)\right)^{-1} f'(x_k)\right\| = \\ &= \left\|\left(f''\left(\frac{x_k + u_k}{2}\right)\right)^{-1}\right\| \left\|f''\left(\frac{x_k + u_k}{2}\right)(x_k - x_*) - \left(f'(x_k) - f'(x_*)\right)\right\| \le \\ &\le \frac{1}{m}\left\|f''(x_*)(x_k - x_*) + f'''(x_*)\left(\frac{x_k + u_k}{2} - x_*\right)(x_k - x_*) + \right. \\ &+ \left(f'''\left(x_* + \tau\left(\frac{x_k + u_k}{2} - x_*\right)\right) - f'''(x_*)\right)\left(\frac{x_k + u_k}{2} - x_*\right)(x_k - x_*) - \\ &- f''(x_*)(x_k - x_*) - \frac{1}{2} f'''(x_*)(x_k - x_*)^2 - \\ &\left. - \frac{1}{2}\left(f'''(x_* + \tau(x_k - x_*)) - f'''(x_*)\right)(x_k - x_*)^2\right\| \le \\ &\le \frac{1}{2m}\|f'''(x_*)\| \|u_k - x_*\| \|x_k - x_*\| + \frac{N}{2m}\left\|x_* + \tau\left(\frac{x_k + u_k}{2} - x_*\right) - x_*\right\| \times \\ &\times \|x_k + u_k - 2x_*\| \|x_k - x_*\| + \frac{N}{2m}\|x_* + \tau(x_k - x_*) - x_*\| \|x_k - x_*\|^2 \le \\ &\le \frac{K}{2m}\frac{K}{2m}\|x_k - x_*\|^3 + \frac{N}{4m}\|x_k + u_k - 2x_*\|^2 \|x_k - x_*\| + \frac{N}{2m}\|x_k - x_*\|^3 \le \\ &\le \left(\frac{K^2}{4m^2} + \frac{3N}{2m}\right)\|x_k - x_*\|^3.\end{aligned} \tag{13}$$

Використавши (9) і (13), отримаємо

$$f(v_k)-f(x_*)\le \frac{M}{2}\|v_k-x_*\|^2\le \frac{M}{2}\left(\frac{K^2}{4m^2}+\frac{3N}{2m}\right)^2\|x_k-x_*\|^6\le$$

$$\le \frac{M}{2}\left(\frac{K^2}{4m^2}+\frac{3N}{2m}\right)^2\left(\frac{2}{m}\right)^3(f(x_k)-f(x_*))^3=C^2(f(x_k)-f(x_*))^3, \quad (14)$$

З формули (7) і (14) випливає

$$f(x_{k+1})-f(x_*)\le \mu_k(f(v_k)-f(x_*))\le \mu_k C^2(f(x_k)-f(x_*))^3,$$

де $\mu_k\in(0,1]$.

За допомогою методу математичної індукції, враховуючи умову 3 теореми, доведемо, що оцінка теореми (8) виконується для довільного k. Якщо $k=1$, то одержимо

$$f(x_1)-f(x_*)\le C^2\mu_1(f(x_0)-f(x_*))^3=\mu_1 q^2(f(x_0)-f(x_*))=$$

$$=\left(\prod_{i=0}^{1-1}\mu_{1-i}^{3^i}\right)q^{3^1-1}(f(x_0)-f(x_*)).$$

Припустимо, що ця оцінка виконується для деякого $k>1$

$$(f(x_k)-f(x_*))\le\left(\prod_{i=0}^{k-1}\mu_{k-i}^{3^i}\right)q^{3^k-1}(f(x_0)-f(x_*)).$$

Доведемо, що оцінка правильна для $k+1$

$$(f(x_{k+1})-f(x_*))\le \mu_{k+1}C^2(f(x_k)-f(x_*))^3\le \mu_{k+1}C^2\left(\left(\prod_{i=0}^{k-1}\mu_{k-i}^{3^i}\right)q^{3^k-1}(f(x_0)-f(x_*))\right)^3\le$$

$$\le \mu_{k+1}\left(\prod_{i=0}^{k-1}\mu_{k-i}^{3^i}\right)^3C^2\left(q^{3^k-1}\right)^3(f(x_0)-f(x_*))^3=\left(\prod_{i=0}^{k}\mu_{k-i+1}^{3^i}\right)q^{3^{k+1}-1}(f(x_0)-f(x_*)).$$

Отже, послідовність наближень $\{x_k\}$, побудованих за формулами (5)–(7), збігається до точки розв'язку x_*. Теорему доведено.

Зауважимо, що вибір початкового наближення x_0, яке б задовольняло умову 3 теореми, є досить складною проблемою, тому на практиці доцільно використовувати алгоритм вигляду

$$u_k=x_k-\alpha_k(f''(x_k))^{-1}f'(x_k), \quad (15)$$

$$v_k=x_k-\beta_k\left(f''\left(\frac{x_k+u_k}{2}\right)\right)^{-1}f'(x_k), \quad (16)$$

$$x_{k+1}=\arg\min_{\gamma} f(u_k+\gamma(v_k-u_k)), \quad (17)$$

де параметри $\alpha_k\in(0,1]$, $\beta_k\in(0,1]$ повинні забезпечувати монотонне спадання функції. В алгоритмі (15)–(17), у разі виконання певних умов, збіжність третього порядку буде досягатися локально.

4. ЧИСЛОВІ ЕКСПЕРИМЕНТИ

Ми розглянули приклади і провели порівняння алгоритму (15)–(17) з методом (3)–(4), в якому подібно до (15)–(17) також було введено крокові множники, а також узагальненим методом Ньютона, на підставі (2). Хоча порядок збіжності алгоритмів (3)–(4) та (15)–(17) однаковий, числові експерименти виявили ефективність запропонованої модифікації в сенсі кількості обчислень, що підтверджує теоретично отримані результати про те, що знаменник збіжності трикрокового методу (5)–(7) є меншим, ніж знаменник збіжності методу (3)–(4). (Зазвичай $\mu_k < 1$).

Обчислення проводили до виконання умови $\|x_{k+1} - x_k\| \le \varepsilon$, $\varepsilon = 10^{-8}$. У таблиці подано кількість ітерацій – h і K – кількість обчислень еквівалентних обчисленню функції $f(x)$, які були затрачені для отримання наближеного розв'язку наведених задач, іншими обчисленнями ми нехтували, оскільки вони суттєво не впливали на ефективність роботи методів. Тестові задачі взяли з [5].

Задача 1. Тричленна показникова функція

$$f(x) = \sum_{i=1}^{n/2} \left[\exp(x_{2i-1} + 3x_{2i} - 0.1) + \exp(x_{2i-1} - 3x_{2i} - 0.1) + \exp(-x_{2i-1} - 0.1)\right].$$

Розв'язок задачі: $x^* \approx (0,-0.34,\ldots,0,-0.34)^T$, значення $f(x^*)$ змінюється залежно від значення n.

1) $x_0 = (1,1,\ldots,1)^T$; 2) $x_0 = (2,2,\ldots,2)^T$.

Задача 2. Штрафна функція 2

$$f(x) = \sum_{i=1}^{n} (x_i - 1)^2 + 10^{-3} \left(\sum_{i=1}^{n} x_i^2 - 0.25\right)^2.$$

Значення x^* та $f(x^*)$ залежать від значення n:

1) $x_0 = (10,10,\ldots,10)^T$; 2) $x_0 = (-10,-10,\ldots,-10)^T$.

Порівняльні характеристики роботи методів (2), (3)–(4), (15)–(17)

Ф-ція	Номер x_0	$n = 100$					
		(2)		(3)–(4)		(15)–(17)	
		h	K	h	K	h	K
1	1	9	46460	6	61307	4	41079
1	2	13	67070	9	91911	5	51336
2	1	10	51613	6	61307	2	20603
2	2	10	51613	6	61307	2	20603
3	1	18	92819	12	122514	3	30774
3	2	32	164938	23	234730	16	163778
4	1	15	77368	10	102111	4	40991
4	2	13	67064	9	91913	4	41019

Задача 3. Штрафна функція 1

$$f(x) = 10^{-5} \sum_{i=1}^{n} (x_i - 1)^2 + \left(\sum_{i=1}^{n} x_i^2 - 0.25\right)^2.$$

Значення x^* та $f(x^*)$ залежать від значення n. Матриця Гессе вироджена в точці розв'язку.

1) $x_0 = (10,10,\ldots,10)^T$; 2) $x_0 = (1,0,1,0,\ldots,1,0)^T$.

Задача 5. Штрафна функція

$$f(x) = \sum_{i=1}^{n-1} (x_i - 1)^2 + \left(\sum_{i=1}^{n} x_i^2 - 0.25 \right)^2 .$$

Значення x^* та $f(x^*)$ залежать від значення n:

1) $x_0 = (10,10,\ldots,10,10)^T$; 2) $x_0 = (5,5,\ldots,5)^T$.

5. ВИСНОВКИ

Ми запропонували схему побудови нового класу методів розв'язування задач безумовної мінімізації функції багатьох змінних. Провели теоретичні та числові дослідження цього алгоритму. Запропонований метод (5)–(7) побудований за принципом трикрокових методів з [1], проте замість двох базових методів використано метод запропонований Коган [3]. У цьому методі на кожному кроці обчислюються два наближення до точки розв'язку, які використовують як проміжні в трикроковому методі, наступне наближення трикрокового методу шукають як точку мінімуму на прямій, що їх з'єднує. Доведено, що метод (5)–(7) має збіжність третього порядку.

На підставі проведених обчислень і порівнянні отриманих результатів, з'ясували, що трикроковий метод (5)–(7) за кількістю обчислень та ітерацій переважає метод (3)–(4) з [3]. Зі збільшенням розмірності функції ефективність запропонованого алгоритму, в сенсі кількості обчислень, зростає.

Подані результати числових обчислень узгоджуються з доведеними теоретичними твердженнями.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. *Бартіш М.* Трикрокові методи розв'язування задач безумовної мінімізації / М. Бартіш, О. Ковальчук, Н. Огородник // Вісн. Львів. ун-ту. Сер. прикл. матем. та інформ. – 2007. – Вип. 13. – С. 3–10.
2. *Васильев Ф.П.* Методы оптимизации / Ф.П. Васильев. – М.: Факториал Пресс, 2002. – 824 с.
3. *Коган Т.И.* Об одном итерационном процессе для функциональных уравнений / Т. И. Коган // Сиб. мат. журн. – 1967. – Т. 8, № 4. – С. 958–960.
4. *Пшеничный Б.Н.* Численные методы в экстремальных задачах / Б.Н. Пшеничный, Ю.М. Данилин. – М.: Наука, 1975. – 320 с.
5. *Andrei N.* An unconstrained optimization test functions collection / Neculai Andrei // Advanced modelling and optimization. – 2008. – Vol. 10, № 1. – P. 147–161.

Стаття: надійшла до редколегії 17.10.2012
доопрацьована 28.11.2012
прийнята до друку 05.12.2012

ТРЁХШАГОВЫЙ ИТЕРАЦИОННЫЙ МЕТОД МИНИМИЗАЦИИ ФУНКЦИЙ С КУБИЧЕСКИМ ПОРЯДКОМ СХОДИМОСТИ

М. Бартиш, Н. Огородник

Львовский национальный университет имени Ивана Франко,
ул. Университетская, 1, Львов, 79000, e-mail: ogorodnyk.nataly@gmail.com

Рассмотрено трёхшаговые итерационные методы решения задач безусловной минимизации функции многих переменных (1), а также метод третьего порядка из статьи

Т. Коган [3]. Используя идею построения трёхшаговых алгоритмов с [1], предложено трёхшаговый метод на основании метода из [3]. Проведены теоретические исследования метода, установлен третий порядок сходимости. При практическом применении алгоритм показал свою эффективность в смысле количества вычислений перед методом из [3] и методом Ньютона на различных типах задач.

Ключевые слова: метод Ньютона, трехшаговый метод, безусловная минимизация.

THREE-STEP ITERATIVE METHOD FOR FUNCTION MINIMIZATION WITH THREE-ORDER CONVERGENCE

M. Bartish, N. Ogorodnyk

Ivan Franko National University of Lviv,
Universytetska Str., 1, Lviv, 79000, e-mail: ogorodnyk.nataly@gmail.com

The article deals with three-step iterative methods for solving many variables unconstrained minimization problems (1), as well as the method of the third order of article T. Kogan [3]. Using the idea of building a three-step algorithms [1], we propose three-step method based on the method of [3]. The theoretical research method, set the third order of convergence. In the practical application of the algorithm showed his efficiency, in terms of computation before method [3] and Newton's method on different types of tasks.

Key words: Newton method, three-step method, unconstrained optimization.