ISSN 2078–5097. Вісн. Львів. ун-ту. Сер. прикл. матем. та інф. 2014. Вип. 21. С. 31–37
Visnyk of the Lviv University. Series Appl. Math. and Informatics. Issue 21. P. 31–37

УДК 519.6

ТРИКРОКОВИЙ ІТЕРАЦІЙНО-РІЗНИЦЕВИЙ 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 [2], побудовано новий варіант трикрокового ітераційно-різницевого методу розв'язування задач безумовної мінімізації з кубічним порядком збіжності на базі методу з [1]. Теоретично обгрунтовано збіжність методу. Виконано числові експерименти, зроблено висновки про ефективність і можливість застосування запропонованого алгоритму.

Ключові слова: ітераційно-різницеві методи, 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. Здебільшого математичні моделі зводяться до розв'язування типових задач оптимізації, а зростання потужності комп'ютерів сприяє розширенню сфер успішного застосування методів оптимізації у розв'язанні все складніших задач більших розмірностей. 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 [3, 4].

Ми досліджуватимемо трикроковий ітераційно-різницевий метод розв'язування задач безумовної мінімізації побудований на базі методу Стеффенсена, а також різницевого аналогу методу третього порядку з [5]. Проведено теоретичні дослідження методу, визначено третій порядок збіжності. При практичному застосуванні алгоритм довів свою ефективність, у сенсі кількості обчислень перед методом з [1] і методом Стеффенсена [4] на різних типах задач.

2. ФОРМУЛЮВАННЯ ЗАДАЧІ

Розглянемо задачу безумовної мінімізації функції багатьох змінних

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

Низка методів розв'язування задачі (1) використовує обернену матрицю других похідних, наприклад, метод Ньютона [4, 6] чи метод із [5].

У випадку, коли матрицю других похідних складно або неможливо обчислити, для розв'язування задачі (1) широко використовують ітераційно-різницеві методи, в яких матриця других похідних апроксимується деякою розділеною різницею першого порядку для градієнта $f'(x)$ зі спеціально вибраними вузлами [4]. Це робить ітерацію методу, для певних типів задач, менш трудомісткою, в сенсі кількості обчислень, порівняно з методами, які використовують матрицю других похідних.

Для методу з [5] у [1] запропоновано ітераційно-різницевий аналог, який можна використовувати для розв'язування задачі (1) у такому вигляді:

$$u_k = x_k - \left(f'\left(x_k, \widetilde{x}_k\right)\right)^{-1} f'\left(x_k\right), \tag{2}$$

$$x_{k+1} = x_k - \left(f'\left(x_k, u_k\right)\right)^{-1} f'\left(x_k\right),\ k = 0,1,\dots, \tag{3}$$

де $\widetilde{x}_k = x_k - \lambda_k f'\left(x_k\right)$, $\lambda_k \le 1$. Тобто (2) є методом Стеффенсена [4].

Використовуючи метод (2)-(3), ідею побудови трикрокових методів з [2], побудуємо такий трикроковий метод:

$$u_k = x_k - \left(f'\left(x_k, \widetilde{x}_k\right)\right)^{-1} f'\left(x_k\right), \tag{4}$$

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

$$x_{k+1} = \arg\min_{\gamma} f\left(u_k + \gamma\left(v_k - u_k\right)\right),\ k = 0,1,\dots . \tag{6}$$

Під час дослідження методу (4) – (6) будемо використовувати таке означення розділених різниць першого порядку:

$$f'\left(x, y\right)\left(x - y\right) = f'\left(x\right) - f'\left(y\right). \tag{7}$$

3. ОБҐРУНТУВАННЯ ЗБІЖНОСТІ

Теорема. Нехай

1) $f(x) \in C(D)$ − сильно опукла з константою опуклості $m > 0$,

$$D = \left\{x \in R^n,\ f\left(x\right) \le f\left(x_0\right)\right\};$$

2) $f'\left(x\right)$ має розділені різниці першого порядку в D, причому для $x, y, z, t \in D$ виконується

$$\left\|\left(f'\left(x, y\right) - f'\left(t, z\right)\right)\right\| \le K\left(\left\|x - t\right\| + \left\|y - z\right\|\right),$$

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

де $0 < K < \infty$, $0 < m \le M$;

3) $\forall x, y \in D$ існує $\left[f'\left(x, y\right)\right]^{-1}$, причому $\left\|\left[f'\left(x, y\right)\right]^{-1}\right\| \le B$, де $B < \infty$ − деяка стала;

4) початкове наближення x_0 вибирають таким, що виконується умова

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

де

$$C = \frac{\left(BK\right)^2}{m}\sqrt{\frac{M}{m}}\ .$$

Тоді послідовність $\left\{x_k\right\}$, породжена методом (4)-(6), коректно визначена та збігається до розв'язку x_* задачі (1) і справджується оцінка

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

де $\mu_k \in \left(0,1\right]$.

Доведення. Існування та єдиність розв'язку (1) випливають з умови сильної опуклості функції $f\left(x\right)$ [6].

Припустимо, що наближення x_k до розв'язку задачі (1) знайдено. Тоді, враховуючи (4) і (7), отримаємо

$$u_k - x_* = x_k - x_* - \left(f'(x_k, \tilde{x}_k)\right)^{-1}\left(f'(x_k) - f'(x_*)\right) = \left(f'(x_k, \tilde{x}_k)\right)^{-1} \times$$
$$\times \left(f'(x_k, \tilde{x}_k)(x_k - x_*) - f'(x_k, x_*)(x_k - x_*)\right) =$$
$$= \left(f'(x_k, \tilde{x}_k)\right)^{-1}\left(f'(x_k, \tilde{x}_k) - f'(x_k, x_*)\right)(x_k - x_*).$$

Тут $\tilde{x}_k = x_k + \zeta(x_* - x_k)$, $\zeta \in (0,1)$.

Врахувавши, що $\tilde{x}_k$ обчислюється згідно з градієнтним методом, отримаємо
$$\|\tilde{x}_k - x_*\| = \|x_k - \lambda_k f'(x_k) - x_*\| \le \|x_k - x_*\|. \tag{9}$$
З урахуванням (9) та умов 2) і 3) матимемо
$$\|u_k - x_*\| \le BK\|x_k - x_*\| \, \|\tilde{x}_k - x_*\| \le BK\|x_k - x_*\|^2. \tag{10}$$
Використовуючи (5) і (10), одержимо
$$\|v_k - x_*\| = \left\|x_k - x_* - \left(f'(x_k, u_k)\right)^{-1}\left(f'(x_k) - f'(x_*)\right)\right\| \le$$
$$\le \left\|\left(f'(x_k, u_k)\right)^{-1}\right\| \left\|\left(f'(x_k, u_k) - f'(x_k, x_*)\right)\right\| \|x_k - x_*\| \le$$
$$\le BK\|u_k - x_*\| \, \|x_k - x_*\| \le (BK)^2\|x_k - x_*\|^3.$$
Використаємо умову 2) і розвинення функції $f(x)$ у ряд Тейлора, тоді
$$f(v_k) - f(x_*) = \left(f'(\tilde{v}_k), v_k - x_*\right) = \left(f'(\tilde{v}_k, x_*)(v_k - x_*), v_k - x_*\right) \le M\|v_k - x_*\|^2,$$
де $\tilde{v}_k = v_k + \zeta(x_* - v_k)$, $\zeta \in (0,1)$.

Отже,
$$f(v_k) - f(x_*) \le M(BK)^4\|x_k - x_*\|^6. \tag{11}$$
На підставі властивості опуклих функцій [6] можемо записати
$$f(x_k) - f(x_*) = \left(f'(\tilde{x}_k), x_k - x_*\right) \ge m\|x_k - x_*\|^2,$$
де $\tilde{x}_k = x_k + \zeta(x_* - x_k)$, $\zeta \in (0,1)$. Отож,
$$\|x_k - x_*\|^2 \le \frac{1}{m}\left(f(x_k) - f(x_*)\right). \tag{12}$$
Використавши (11) і (12), отримаємо
$$f(v_k) - f(x_*) \le M(BK)^4\|x_k - x_*\|^6 \le$$
$$\le M(BK)^4\left(\frac{1}{m}\right)^3\left(f(x_k) - f(x_*)\right)^3 = C^2\left(f(x_k) - f(x_*)\right)^3, \tag{13}$$
З формули (6) і (13) випливає
$$f(x_{k+1}) - f(x_*) \le \mu_k\left(f(v_k) - f(x_*)\right) \le \mu_k C^2\left(f(x_k) - f(x_*)\right)^3,$$
де $\mu_k \in (0,1]$.

За допомогою методу математичної індукції, враховуючи умову 4 теореми, доведемо, що оцінка теореми (8) виконується для довільного k. Якщо $k = 1$, то одержимо
$$f(x_1) - f(x_*) \le C^2\mu_1\left(f(x_0) - f(x_*)\right)^3 = \mu_1 q^2\left(f(x_0) - f(x_*)\right) =$$
$$= \left(\prod_{i=0}^{1-1}\mu_{1-i}^{3^i}\right)q^{3^1-1}\left(f(x_0) - f(x_*)\right).$$

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

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

Доведемо, що оцінка правильна для $k+1$

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

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

Отже, послідовність наближень $\{x_k\}$, побудованих за формулами (4)-(6) збігається до точки розв'язку x_*. Теорему доведено.

Зауважимо, що вибір початкового наближення x_0, яке б задовольняло умову 4 теореми, є досить складною проблемою, тому на практиці доцільно використовувати алгоритм вигляду

$$u_k=x_k-\alpha_k\left(f'(x_k,\widetilde{x}_k)\right)^{-1}f'(x_k),\tag{14}$$

$$v_k=x_k-\beta_k\left(f'(x_k,u_k)\right)^{-1}f'(x_k),\tag{15}$$

$$x_{k+1}=\arg\min_{\gamma}f\left(u_k+\gamma(v_k-u_k)\right),\tag{16}$$

де параметри $\alpha_k\in(0,1]$, $\beta_k\in(0,1]$ повинні забезпечувати монотонне спадання функції. В алгоритмі (14)-(16), у разі виконання певних умов, збіжність третього порядку буде досягатися локально.

4. ЧИСЛОВІ ЕКСПЕРИМЕНТИ

Ми розглянули приклади та провели порівняння алгоритму (14)-(16) з методом (2)-(3), в якому подібно до (14)-(16) також було введено крокові множники, а також узагальненим методом Стеффенсена

$$x_{k+1}=x_k-\alpha_k\left(f'(x_k,x_k-\lambda_kf'(x_k))\right)^{-1}f'(x_k).\tag{17}$$

У [4] для обчислення розділених різниць першого порядку запропоновано формулу

$$f'(x,y)_{i,j}=\frac{f_i'(x_1,\dots,x_j,y_{j+1},\dots,y_n)-f_i'(x_1,\dots,x_{j-1},y_j,\dots,y_n)}{x_j-y_j}.$$

Параметр приймемо сталим на кожному кроці і таким, що $\lambda_k=10^{-4}$.

Хоча порядок збіжності алгоритмів (2)-(3) і (4)-(6) однаковий, числові експерименти виявили ефективність запропонованої модифікації в сенсі кількості обчислень, що підтверджує теоретично отримані результати про те, що знаменник збіжності трикрокового методу (4)-(6) є меншим, ніж знаменник збіжності методу (2)-(3). (Зазвичай $\mu_k<1$).

Обчислення проводили до виконання умови

$$\left\|x_{k+1}-x_k\right\|\le\varepsilon,\ \varepsilon=10^{-8}.$$

У табл. 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 [7].

Порівняльні характеристики роботи методів (17), (2)-(3), (14)-(16)

Ф-ція	№ x_0	$n=100$					
		(17)		(2)-(3)		(14)-(16)	
		h	K	h	K	h	K
1	1	9	91910	6	121907	4	81417
1	2	11	112313	6	121910	4	81401
2	1	10	102111	6	121913	3	61072
2	2	10	102112	6	121913	3	61074
3	1	16	163317	11	223423	3	61067
3	2	32	326538	16	324999	10	203518
4	1	16	163328	11	223426	5	101716
4	2	8	81712	6	121917	4	81412

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:

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

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

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

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

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

Задача 4. 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^* і $f(x^*)$ залежать від значення n:

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

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 (4)-(6) побудований за принципом трикрокових методів з [2], проте замість двох базових методів використано різницевий аналог методу з [1]. 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 (4)-(6) має збіжність третього порядку.

На підставі проведених обчислень і порівнянні отриманих результатів, з'ясували, що трикроковий метод (4)-(6) за кількістю обчислень та ітерацій переважає метод (2)-(3) з [1]. 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. *Бартиш М. Я.* О некоторых итерационных методах решения функциональных уравнений / М. Я. Бартиш // Сиб. мат. журн. – 1969. – Т. 10, № 3. – С. 488-493.
2. *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.
3. *Бейко І. В.* Задачі, методи та алгоритми оптимізації / І. В. Бейко, П. М. Зінько, О. Г. Наконечний. – К.: ВПЦ Київський університет, 2012. – 800 с.
4. *Васильев Ф. П.* Методы оптимизации / Ф. П. Васильев. – М.: Факториал Пресс, 2002. – 824 с.
5. *Коган Т. И.* Об одном итерационном процессе для функциональных уравнений / Т. И. Коган // Сиб. мат. журн. – 1967. – Т. 8, № 4. – С. 958-960.
6. *Пшеничный Б. Н.* Численные методы в экстремальных задачах / Б. Н. Пшеничный, Ю. М. Данилин. – М.: Наука, 1975. – 320 с.
7. *Andrei N.* An unconstrained optimization test functions collection / Neculai Andrei // Advanced modelling and optimization. – 2008. – Vol. 10, No 1. – P. 147-161.

Стаття: надійшла до редколегії 22.01.2014
доопрацьована 12.02.2014
прийнята до друку 05.03.2014

ТРЕХШАГОВЫЙ ИТЕРАЦИОННО-РАЗНОСТНЫЙ 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

Рассматриваются методы третьего порядка сходимости. На основании метода с [5] в [1] построен разностный метод и доказан его третий порядок сходимости. Используя идею построения трехшаговых алгоритмов с [2], предложено трёхшаговый метод на основании метода из [1]. Проведены теоретические исследования метода, установлен третий порядок сходимости. При практическом применении алгоритм показал свою эффективность, в смысле количества вычислений перед методом из [1] и методом Стеффенсена, на различных типах задач.

Ключевые слова: метод Ньютона, метод Стеффенсена, итерационно-разностные методы, трехшаговый метод, безусловная минимизация.

THREE-STEP ITERATIVE DIFFERENCE 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 M. Bartish [1]. Using the idea of building a three-step algorithms [2], we propose three-step method based on the method of [1]. 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 [1] and Steffensen method on different types of tasks.

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