**ПРИКЛАДНА МАТЕМАТИКА**

*УДК* 519.6

# ТРИКРОКОВІ МЕТОДИ МІНІМІЗАЦІЇ ФУНКЦІЙ З НАДКВАДРАТИЧНОЮ ЗБІЖНІСТЮ

**М. Бартіш, О. Ковальчук, Н. Огородник**

*Львівський національний університет імені Івана Франка,*
*вул. Університеська, 1, Львів, 79000,*
*e-mail:* *mykhaylo.bartish@lnu.edu.ua*, *olha.kovalchuk@lnu.edu.ua*,
*nataliya.ohorodnyk@lnu.edu.ua*

Досліджено новий варіант побудови трикрокових методів мінімізації функції багатьох змінних. Розглянуто загальну схему побудови таких методів, а також два конкретних алгоритми. Доведено збіжність запропонованих алгоритмів. У використанні запропонованих алгоритмів детальніше використовується отримана інформація про значення функції її похідних і поділених різниць. З'ясовано, що швидкість збіжності запропонованих алгоритмів вища або дорівнює швидкості збіжності базових алгоритмів.

*Ключові слова*: градієнтний метод, поділені різниці, ітераційно-різницеві методи, трикрокові методи.

## 1. Вступ

На практиці часто виникають задачі, математичні моделі яких призводять до задач оптимізації. В літературі подано низку алгоритмів [3, 4] для розв'язування таких задач. На жаль, не має універсального методу їх розв'язування. Ми пропонуємо новий підхід для побудови комбінованих алгоритмів, які ефективні в сенсі кількості обчислень для розв'язування таких задач порівняно з базовими. Запропоновано клас трикрокових алгоритмів розв'язування задач мінімізації функції багатьох змінних. Запропоновані алгоритми більш ефективно використовують інформацію, отриману на кожній ітерації, про значення функції, її похідних, поділених різниць на кожній ітерації.

## 2. Формулювання задачі

Розглянемо задачу безумовної мінімізації

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

Загальновідомими методами розв'язування задачі (1) є градієнтний метод, метод Ньютона та їхні модифікації тощо [3–5]. Використовуючи інформацію про уже відомі методи, ми побудуємо новий клас методів, який дає змогу вибрати ефективний алгоритм в сенсі кількості обчислень для розв'язування конкретної задачі (1), а саме

$$\begin{aligned} u_k &= \Phi(x_k); \\ v_k &= x_k - f'(x_k, u_k)^{-1} f'(x_k); \\ x_{k+1} &= arg\min_{\gamma} f(u_k + \gamma(v_k - u_k)); \\ k &= 0, 1, ..., \end{aligned} \tag{2}$$

де $\Phi(x)$ оператор, який задовольняє умову

$$\|\Phi(x) - x^*\| \leq K \|x - x^*\|^\tau, \quad \tau \in [1, 2], \tag{3}$$

$f'(x, y)$ – поділена різниця першого порядку вектор-функції $f'(x)$ [5]; $x^*$ – розв'язок задачі (1).

## 3. Обґрунтування збіжності методу (2)

**Теорема 1.** *Нехай виконуються умови:*

1) $f(x) \in C^1(D)$, *де* $D = \{x \in R^n, f(x) \leq f(x_0)\}$, $x_0$ – *початкове наближення;*
2) $\|f'(x, y) - f'(x, z)\| \leq M_1\|y - z\|$ *для всіх* $x, y, z \in D$;
3) $m \|x - y\|^2 \leq (f'(x) - f'(y), x - y) \leq M \|x - y\|^2$, *де* $0 < m \leq M$ *для всіх* $x, y \in D$;
4) *існує* $f'(x, y)^{-1}$ *, а також* $\left\|f'(x, y)^{-1}\right\| \leq B$ *для всіх* $x, y \in D$;
5) *для всіх* $x \in D$ *оператор* $\Phi(x)$ *задовольняє умову* $\|\Phi(x) - x^*\| \leq K \|x - x^*\|^\tau$, $\tau \in [1, 2]$, $K$ – *деяка константа, що* $0 < K < \infty$;
6) *початкове наближення* $x_0$ *задовольняє умову*

$$q^\tau = L(f(x_0) - f(x^*)^\tau < 1,$$

*де* $L = \frac{MB^2M_1^2K^22^\tau}{m^{1+\tau}}$.

*Тоді справджується оцінка*

$$f(x_k) - f(x^*) \leq \prod_{i=1}^{k} \gamma_i^{(1+\tau)^{k-i}} q^{(1+\tau)^k - 1}(f(x_0) - f(x^*)), k = 0, 1, ...,$$

*де* $\gamma_i \in (0, 1]$ *і* $x^*$ – *розв'язок задачі(1).*

*Доведення.* Із умови сильної опуклості функцій випливає існування точки $x^*$, тоді $f'(x^*) = 0$, $f(x^*) \leq f(x)$ для всіх $x \in D$.

Нехай для $k \geq 1$ отримали $x_k$, тоді одержимо

$$\|v_k - x^*\| = \left\|x_k - x^* - f'(x_k, u_k)^{-1}f'(x_k))\right\| = \left\|f'(x_k, u_k)^{-1}\right\| \|(f'(x_k, u_k)-$$

$$-f'(x_k, x^*))(x_k - x^*)\| \leq BM_1 \|(u_k - x^*)(x_k - x^*)\| \leq BM_1K \|x_k - x^*\|^{1+\tau}.$$

Використавши умову (3), отримаємо

$$f(x) - f(x^*) = \int_0^1 \frac{1}{\tau}(f'(x^* + \tau(x - x^*)) - f'(x^*), \tau(x - x^*))d\tau \geq$$

$$\geq m \int_0^1 \tau \|x - x^*\|^2 d\tau = \frac{m}{2} \|x - x^*\|^2$$

і

$$\|x - x^*\|^2 \leq \frac{2}{m}(f(x) - f(x^*)).$$

Одночасно

$$f(x) - f(x^*) = \int_0^1 \frac{1}{\tau}(f'(x^* + \tau(x - x^*)) - f'(x^*), \tau(x - x^*))d\tau \leq \frac{M}{2} \|x - x^*\|^2.$$

Тепер можемо записати

$$f(v_k) - f(x^*) \le \frac{M}{2}\|v_k - x^*\|^2 \le \frac{M}{2}(BM_1K\|x_k - x^*\|^{1+\tau})^2 \le$$

$$\le \frac{M(BM_1K)^2}{2}\|x_k - x^*\|^{2(1+\tau)} \le \frac{M(BM_1K)^2}{2}(\frac{2}{m}(f(x_k) - f(x^*))^{1+\tau} \le$$

$$\le \frac{M(BM_1K)^2 2^\tau}{m^{1+\tau}}(f(x_k) - f(x^*))^{1+\tau} = L(f(x_k) - f(x^*))^{1+\tau}.$$

Отримаємо

$$(f(x_{k+1}) - f(x^*)) \le \gamma_{k+1}L(f(x_k) - f(x^*))^{1+\tau}. \tag{4}$$

Оскільки $f(x_{k+1}) \le min\{f(u_k), f(v_k)\}$, то $\gamma_{k+1} \le 1$.
Застосуємо метод математичної індукції.
Для $k = 0$ матимемо

$$(f(x_1) - f(x^*)) \le \gamma_1 L(f(x_0) - f(x^*))^{1+\tau} \le \gamma_1 q^\tau (f(x_0) - f(x^*)).$$

Нехай для деякого $k \ge 1$ виконується умова

$$f(x_k) - f(x^*) \le \prod_{i=1}^{k} \gamma_i^{(1+\tau)^{k-i}} q^{(1+\tau)^k - 1}(f(x_0) - f(x^*)). \tag{5}$$

Тоді згідно з (4) і (5)

$$(f(x_{k+1}) - f(x^*)) \le \gamma_{k+1}L(f(x_k) - f(x^*))^{1+\tau} \le \gamma_{k+1}L(\prod_{i=1}^{k} \gamma_i^{(1+\tau)^{k-i}} q^{(1+\tau)^k - 1}\cdot$$

$$\cdot(f(x_0) - f(x^*)))^{1+\tau} \le \prod_{i=1}^{k+1} \gamma_i^{(1+\tau)^{k-i+1}} q^{(1+\tau)^{k+1} - 1}(f(x_0) - f(x^*)).$$

Отже, виконується оцінка збіжності і теорема доведена. □

Цікаві, у цьому випадку, конкретні варіанти оператора $\Phi(x)$.

## 4. Перший вибір оператора $\Phi(x)$

Розглянемо випадок, коли для обчислення $u_k$ використано градієнтний метод

$$\Phi(x_k) = x_k - \alpha_k f'(x_k); \tau = 1.$$

Тоді отримаємо наступний алгоритм:

$$\begin{aligned} u_k &= x_k - \alpha_k f'(x_k); \\ v_k &= x_k - f'(x_k, u_k)^{-1} f'(x_k); \\ x_{k+1} &= arg\min_{\gamma} f(u_k + \gamma(v_k - u_k)); \\ k &= 0, 1, .... \end{aligned} \tag{6}$$

**Теорема 2.** *Нехай виконуються умови:*
*1)* $f(x) \in C^1(D)$, *де* $D = \{x \in R^n, f(x) \le f(x_0)\}$;

*2)* $\|f'(x,y) - f'(x,z)\| \le M_1\|y-z\|$ *для всіх* $x,y,z \in D$;

*3)* $m\left\|x-y\right\|^2 \le (f'(x)-f'(y), x-y) \le M\left\|x-y\right\|^2$, де $0 < m \le M$ *для всіх* $x,y \in D$;

*4) існує* $f'(x,y)^{-1}$ , *а також* $\left\|f'(x,y)^{-1}\right\| \le B$ *для всіх* $x,y \in D$;

*5) початкове наближення* $x_0$ *задовольняє умову*

$$q = L(f(x_0) - f(x^*) < 1,$$

де $L = \frac{2MB^2M_1^2}{m^2}$.

*Тоді справджується оцінка*

$$f(x_k) - f(x^*) \le \prod_{i=1}^{k} \gamma_i^{2^{k-i}} q^{2^k-1}(f(x_0) - f(x^*)), k = 0, 1, ...,$$

*де* $\gamma_i \in (0,1]$ *і* $x^*$ *– розв'язок задачі (1).*

*Доведення.* Із умови сильної опуклості функцій випливає існування точки $x^*$, тоді $f'(x^*) = 0$, $f(x^*) \le f(x)$ для всіх $x \in D$.

Нехай для $k \ge 1$ отримали $x_k$, тоді одержимо

$$\|v_k - x^*\| = \left\|x_k - x^* - f'(x_k,u_k)^{-1}f'(x_k)\right\| = \left\|f'(x_k,u_k)^{-1}\right\| \|(f'(x_k,u_k)-$$

$$-f'(x_k,x^*))(x_k - x^*)\| \le BM_1\left\|(u_k - x^*)(x_k - x^*)\right\| \le BM_1q_1\left\|x_k - x^*\right\|^2.$$

У нашому випадку $\|u_k - x^*\| \le q_1\left\|x_k - x^*\right\|$ де $q_1 < 1$.

Оскільки

$$\|u_k - x^*\|^2 \le \frac{2}{m}(f(u_k) - f(x^*)),$$

$$(f(v_k) - f(x^*)) \le \frac{M}{2}\left\|v_k - x^*\right\|^2,$$

то тепер можемо записати

$$f(v_k) - f(x^*) \le \frac{M}{2}\left\|v_k - x^*\right\|^2 \le \frac{M}{2}(BM_1q_1\left\|x_k - x^*\right\|^2)^2 =$$

$$= \frac{M(BM_1q_1)^2}{2}\left\|x_k - x^*\right\|^4 \le \frac{M(BM_1q_1)^2}{2}(\frac{2}{m}(f(x_k) - f(x^*))^2 \le$$

$$\le \frac{M(BM_1q_1)^2 2}{m^2}(f(x_k) - f(x^*))^2 < L(f(x_k) - f(x^*))^2.$$

Отже,

$$(f(x_{k+1}) - f(x^*)) \le \gamma_{k+1}L(f(x_k) - f(x^*))^2, \qquad (7)$$

де $\gamma_{k+1} \le 1$.

Застосуємо метод математичної індукції. Для $k = 0$ матимемо

$$(f(x_1) - f(x^*)) \le \gamma_1 L(f(x_0) - f(x^*))^2 \le \gamma_1 q((f(x_0) - f(x^*)).$$

Нехай для деякого $k \ge 1$ виконується умова

$$f(x_k) - f(x^*) \le \prod_{i=1}^{k} \gamma_i^{2^{k-i}} q^{2^k-1}(f(x_0) - f(x^*)). \qquad (8)$$

Тоді згідно з (7) і (8)

$$f(x_{k+1}) - f(x^*) \le \gamma_{k+1} L (f(x_k) - f(x^*))^2 \le \gamma_{k+1} L (\prod_{i=1}^{k} \gamma_i^{2^{k-i}} q^{2^k - 1} \times$$

$$\times (f(x_0) - f(x^*)))^2 \le \prod_{i=1}^{k+1} \gamma_i^{2^{k-i+1}} q^{2^{k+1}-1} (f(x_0) - f(x^*)).$$

Отже, виконується оцінка збіжності і теорема доведена. □

## 5. Другий вибір оператора $\Phi(x)$

Розглянемо ще один варіант методу (2), а саме

$$\begin{aligned} u_k = x_k - f'(x_k, x_k - \alpha f'(x_k))^{-1} f'(x_k); \\ v_k = x_k - f'(x_k, u_k)^{-1} f'(x_k); \\ x_{k+1} = arg \min_{\gamma} f(u_k + \gamma(v_k - u_k)); \\ k = 0, 1, .... \end{aligned} \tag{9}$$

**Теорема 3.** *Нехай виконуються умови:*

*1)* $f(x) \in C^1(D)$, де $D = \{x \in R^n, f(x) \le f(x_0)\}$;

*2)* $\|f'(x,y) - f'(x,z)\| \le M_1 \|y - z\|$ *для всіх* $x, y, z \in D$;

*3)* $m \|x - y\|^2 \le (f'(x) - f'(y), x - y) \le M \|x - y\|^2$, *де* $0 < m \le M$ *для всіх* $x, y \in D$;

*4) існує* $f'(x,y)^{-1}$ , *а також* $\left\|f'(x,y)^{-1}\right\| \le B$ *для всіх* $x, y \in D$;

*5)* $\|f'(x,y)\| \le M_2$ *для всіх* $x, y \in D$;

*6) початкове наближення* $x_0$ *задовольняє умову*

$$q = L(f(x_0) - f(x^*) < 1,$$

*де* $L = \frac{2\sqrt{M} B^2 M_1^2 (1 + \alpha M_2)}{\sqrt{m^3}}$.

*Тоді справджується оцінка*

$$f(x_k) - f(x^*) \le \prod_{i=1}^{k} \gamma_i^{3^{k-i}} q^{3^k - 1} (f(x_0) - f(x^*)), k = 0, 1, ...,$$

*де* $x^*$ *– розв'язок задачі(1).*

*Доведення.* Нехай для деякого $k > 1$ отримаємо $x_k$

$$\|u_k - x^*\| = \left\|x_k - x^* - f'(x_k, x_k - \alpha f'(x_k))^{-1} f'(x_k)\right\| =$$

$$= \left\|f'(x_k, x_k - \alpha f'(x_k))^{-1} \left(f'(x_k, x_k - \alpha f'(x_k)) - f'(x_k, x^*)\right)(x_k - x^*)\right\| \le$$

$$\le BM_1 \|x_k - \alpha f'(x_k) - x^*\| \|x_k - x^*\| \le BM_1(1 + \alpha M_2) \|x_k - x^*\|^2 .$$

Отже, тепер можна записати

$$\|v_k - x^*\| = \left\|x_k - x^* - f'(x_k, u_k)^{-1} f'(x_k))\right\| = \left\|f'(x_k, u_k)^{-1}\right\| \cdot$$

$$\cdot \|(f'(x_k, u_k)) - f'(x_k, x^*))(x_k - x^*)\| \le BM_1 \|u_k - x^*\| \|x_k - x^*\| \le$$
$$\le (BM_1)^2(1 + \alpha M_2) \|x_k - x^*\|^3 .$$

Оскільки
$$\|u_k - x^*\|^2 \le \frac{2}{m}(f(u_k) - f(x^*)),$$
то
$$(f(v_k) - f(x^*)) \le \frac{M}{2} \|v_k - x^*\|^2 .$$

Отримаємо
$$f(v_k) - f(x^*) \le \frac{M}{2} \|v_k - x^*\|^2 \le \frac{M}{2}((BM_1)^2(1 + \alpha M_2) \|x_k - x^*\|^3)^2 =$$
$$= \frac{M(BM_1)^4}{2}(1 + \alpha M_2)^2 \|x_k - x^*\|^6 \le \frac{M(BM_1)^4}{2}(1 + \alpha M_2)^2(\frac{2}{m}(f(x_k) - f(x^*))^3 =$$
$$= \frac{4M(BM_1)^4}{m^3}(1 + \alpha M_2)^2(f(x_k) - f(x^*))^3 = L^2(f(x_k) - f(x^*))^3.$$

Звідси
$$(f(x_{k+1}) - f(x^*)) \le \gamma_{k+1} L^2(f(x_k) - f(x^*))^3. \tag{10}$$

Для доведення теореми також використаємо метод математичної індукції. При $k = 0$
$$(f(x_1) - f(x^*)) \le \gamma_1 L^2(f(x_0) - f(x^*))^3 \le \gamma_1 q^2(f(x_0) - f(x^*))^3.$$

Нехай для деякого $k \ge 1$ виконується умова
$$f(x_k) - f(x^*) \le \prod_{i=1}^{k} \gamma_i^{3^{k-i}} q^{3^k-1}(f(x_0) - f(x^*)). \tag{11}$$

Тоді згідно з (10) і (11)
$$f(x_{k+1}) - f(x^*) \le \gamma_{k+1} L^2(f(x_k) - f(x^*))^3 \le \gamma_{k+1} L^2(\prod_{i=1}^{k} \gamma_i^{3^{k-i}} q^{3^k-1} \times$$
$$\times (f(x_0) - f(x^*)))^3 \le \prod_{i=1}^{k+1} \gamma_i^{3^{k-i+1}} q^{3^{k+1}-1}(f(x_0) - f(x^*)).$$

Отже, виконується оцінка збіжності і теорема доведена.

□

Зауважимо, якщо у (9) параметр $\alpha$ вибрати таким, що змінюється аналогічно як у градієнтному методі, то це дасть змогу зменшити знаменник збіжності методу.

## 6. Висновок

На підставі запропонованого у [1, 2] підходу до побудови методів розв'язування задач безумовної мінімізації функції багатьох змінних запропоновано схему трикрокових алгоритмів із надквадратичною швидкістю збіжності. Доведено збіжність у загальному випадку. Також розглянуто часткові випадки, а саме трикроковий метод з надквадратичною збіжністю на базі градієнтного методу, а також на базі методу Стефенсена. Доведено збіжність запропонованих методів.

## Список використаної літератури

1. Бартіш М.Я. Про один трикроковий метод мінімізації функцій / М. Бартіш, Н. Огородник // Математичні студії. – 2010. – Том 34, № 1. – С. 106–112.
2. Бартіш М.Я. Трикрокові методи розв'язування задач безумовної мінімізації / М. Бартіш, О. Ковальчук, Н. Огородник // Вісн. Львів. національного ун-ту ім. І. Франка. Сер. прикл. матем. та інформ. – 2007. – Вип. 13. – С. 3-10.
3. Бейко І.В. Задачі, методи та алгоритми оптимізації / І. Бейко, П. Зінько, О. Наконечний. – Київ: Видавництво поліграфічного центру Київського університету, 2012.
4. Дэннис Дж. мл. Численные методы безусловной оптимизации и решения нелинейных уравнений. / Дж. Дэннис, Р. Шнабель. – Мир, 1988.
5. Ульм С.Ю. Об обобщенных разделенных разностях / С. Ульм, // Известия АН ЭССР. – 1967. – Т. 16, № 1. – С. 13-26.

*Стаття: надійшла до редколегії* 15.06.2022
*доопрацьована* 20.08.2022
*прийнята до друку* 07.09.2022

**THREE-STEP METHODS OF MINIMIZING FUNCTIONS WITH SUPERQUADRATIC CONVERGENCE**

**M. Bartish, O. Kovalchuk, N. Ogorodnyk**

*Ivan Franko National University of Lviv,*
*Universytetska str., 1, Lviv, 79000, Ukraine*
*e-mail: mykhaylo.bartish@lnu.edu.ua, olha.kovalchuk@lnu.edu.ua,*
*ogorodnyk.nataly@gmail.com*

The article considers the solving to the problem of unconstrained minimization of a function of many variables. Well-known methods for solving such problems are gradient methods and Newton's methods. But all existing methods have their advantages and disadvantages and solve well different classes of problems. In previous works, we proposed a three-step method. On each iteration, it uses the information about the function and the function's derivatives obtained at the previous iteration more efficiently. Using the same approach as in the three-step algorithm, we propose new methods. In the article, we presented a general scheme of construction of new three-step methods and two particular variants. Given the conditions of the method's convergence and proved convergence in the general case.

As a basic method, we use the method of divided differences and gradient method in the first case; the Steffensen method and method of divided differences in the second. The main calculations of function and its derivatives are made in the first method of divided differences. So the three-step method reaches the solution in fewer steps than the basic method if it is used alone.

In both cases, theorems of convergence are proved. Also, we showed that the rate of convergence of the new-proposed algorithms is greater or equal then the rate of convergence of basic methods.

Numerical experiments were done and their results proved the theoretical research.

*Key words*: gradient method, divided differences, iterative-difference methods, and three-step methods.