ПРИКЛАДНА МАТЕМАТИКА

УДК 517.9

ПРО ДЕЯКІ МЕТОДИ МІНІМІЗАЦІЇ ФУНКЦІЙ З НАДКВАДРАТИЧНОЮ ЗБІЖНІСТЮ

М. Бартіш, О. Ковальчук, Н. Огородник

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

Запропоновано ідею побудови двокрокових методів для мінімізації функцій багатьох змінних. Побудовано один узагальнений метод і два комбінованих методи, які ґрунтуються на відомих методах мінімізації функцій. Виконано теоретичне дослідження запропонованих алгоритмів. Також доведено, як вибір конкретного алгоритму впливає на швидкість збіжності.

Ключові слова: метод Ньютона, різницевий метод, поділена різниця, метод Стеффенсена.

1. Вступ

Розв'язуючи реальні задачі практики, натрапляємо на задача мінімізації функцій багатьох змінних. Знайти точний розв'язок задач на екстремум вдається дуже рідко. Наявність обчислювальної техніки дала змогу досить ефективно знаходити наближені розв'язки таких задач із наперед заданою точністю. Сьогодні існує низка алгоритмів, яка допомагає вибрати ефективний алгоритм для розв'язування таких задач [1, 3, 4]. Це зазвичай ітераційні методи. Існування значної кількості методів розв'язування задач мінімізації функцій багатьох змінних свідчить про те, що не існує універсального алгоритму і як звичайно побудувати такий алгоритм неможливо. Ми розглянули новий підхід до побудови комбінованих алгоритмів [2], що дає змогу досліджувати нові алгоритми, які ефективніші за базові, в сенсі кількості обчислень. У повідомленні побудовано клас двокрокових алгоритмів для розв'язування задач мінімізації функції багатьох змінних.

2. Формулювання задачі

Розглянемо задачу безумовної мінімізації

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

Загальновідомими методами розв'язування задачі (1) є градієнтний метод, метод Ньютона та їхні модифікації [1, 2]. Ми розглянемо нову ідею побудови методів розв'язування задачі (1), а саме

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

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

$$\begin{aligned} \|\Phi(x)-x^*\| \le K\,\|x-x^*\|^{\tau} \\ \tau \in [1,2], \end{aligned} \tag{3}$$

$f'(x,u)$ - поділена різниця вектор-функції $f'(x)$ [5].

3. Обґрунтування збіжності методу (2)

Теорема 1. *Нехай виконуються умови:*

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

Тоді справджується оцінка

$$\|x_k - x^*\| \le q^{(\tau+1)^k-1}\,\|x_0-x^*\|\,, k = 0, 1, ..., \tag{4}$$

де x^* - розв'язок задачі(1).

Доведення. Нехай отримали x_k. Тоді, враховуючи умову $f'(x^*) = 0$, одержимо

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

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

$$\le B\,\|(f'(x_k,u_k) - f'(x_k,x^*))(x_k-x^*)\| \le BMK\,\|x_k-x^*\|^{1+\tau}\,.$$

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

$$\|x_1-x^*\| \le BMK\,\|x_0-x^*\|^{1+\tau} = q^{\tau}\,\|x_0-x^*\| = q^{(1+\tau)^1-1}\,\|x_0-x^*\|\,.$$

Нехай для деякого $k \ge 1$ виконується нерівність

$$\|x_k-x^*\| \le q^{(\tau+1)^k-1}\,\|x_0-x^*\|\,.$$

Тоді

$$\|x_{k+1}-x^*\| \le BMK\,\|x_k-x^*\|^{1+\tau} \le BMKq^{((1+\tau)^k-1)(1+\tau)}\,\|x_0-x^*\|^{1+\tau} \le$$

$$\le q^{(1+\tau)^{k+1}-1}\,\|x_0-x^*\|\,.$$

Отже, виконується оцінка (4). □

Цікавимиі у цьому випадку є конкретні варіанти оператора $\Phi(x)$.

4. Перший вибір оператора $\Phi(x)$

Розглянемо випадок, коли для обчислення u_k використано градієнтний метод. Тоді отримаємо такий алгоритм:

$$\begin{aligned} u_k = x_k - \alpha_k f'(x_k); \\ x_{k+1} = x_k - f'(x_k, u_k)^{-1} f'(x_k); \\ k = 0, 1, ... \end{aligned} \tag{5}$$

Для (5) має виконуватися таке:

$$f(u_k) - f(x_k) \le -\epsilon\alpha_k \left\| f'(x_k) \right\|^2;$$

$$\epsilon \in (0; 1).$$

У цьому випадку під час виконання відповідних умов має справджується оцінка

$$\left\| u_k - x^* \right\| \le C \left\| x_k - x^* \right\|, \tag{6}$$

де $C < 1$.

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) $\|f'(x, y) - f'(x, z)\| \le M\|y - z\|$ *для всіх* $x, y, z \in D$;
3) *існує* $f'(x, y)^{-1}$, *а також* $\left\| f'(x, y)^{-1} \right\| \le B$ *для всіх* $x, y \in D$;
4) *для всіх* $x, y \in D$ *функція* $f(x)$ *сильно опукла і*
   $\mu \left\| x - y \right\|^2 \le (f'(x) - f'(y), x - y) \le L \left\| x - y \right\|^2$;
5) *початкове наближення* x_0 *задовольняє умову*
   $q = BMC \left\| x_0 - x^* \right\| < 1$, *де* C*-деяка константа, така що* $C \in (1 - \frac{\mu}{L}, 1)$.

Тоді справджується оцінка

$$\left\| x_k - x^* \right\| \le q^{2^k - 1} \left\| x_0 - x^* \right\|, k = 0, 1, ..., \tag{7}$$

де x^* *- розв'язок задачі(1).*

Доведення. Аналогічно до попередньої теореми, ми отримаємо

$$\left\| x_{k+1} - x^* \right\| = \left\| x_k - x^* - f'(x_k, u_k)^{-1} f'(x_k)) \right\| \le BM \left\| u_k - x^* \right\| \left\| x_k - x^* \right\|.$$

Використавши оцінку (6), можемо записати

$$\left\| x_{k+1} - x^* \right\| \le BMC \left\| x_k - x^* \right\|^2.$$

Застосуємо метод математичної індукції.

Для $k = 0$ матимемо

$$\left\| x_1 - x^* \right\| \le BMC \left\| x_0 - x^* \right\|^2 \le q \left\| x_0 - x^* \right\| = q^{2^1 - 1} \left\| x_0 - x^* \right\|.$$

Нехай для $k \ge 1$ виконується оцінка

$$\left\| x_k - x^* \right\| \le q^{2^k - 1} \left\| x_0 - x^* \right\|.$$

Тоді

$$\|x_{k+1} - x^*\| \le BMC \|x_k - x^*\|^2 \le BMCq^{2^{k+1}-2} \|x_0 - x^*\|^2 \le q^{2^{k+1}-1} \|x_0 - x^*\| .$$

Отже, виконується оцінка (7). □

Послідовність $\{x_k\}, k = 0, 1, ...$ квадратично збігається до розв'язку задачі. Можна зауважити, що алгоритм (5) за своєю структурою нагадує метод Стеффенсена. В наступному алгоритмі вибір другої точки для побудови матриці поділених різниць визначаємо алгоритмом, що автоматично враховує наближення послідовності $\{x_k\}$ до розв'язку x^*. Кількість обчислень на кожній ітерації практично еквівалентні, а знаменник збіжності для запропонованого алгоритму менший ніж відповідний у методі Стеффенсена.

5. Другий вибір оператора $\Phi(x)$

Розглянемо ще один варіант методу (2), а саме

$$\begin{aligned} u_k = x_k - f'(x_k, x_k - \alpha f'(x_k))^{-1} f'(x_k); \\ x_{k+1} = x_k - f'(x_k, u_k)^{-1} f'(x_k); \\ k = 0, 1, ... \end{aligned} \tag{8}$$

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

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

5) початкове наближення x_0 *задовольняє умову* $q = BM_1\sqrt{1 + \alpha M_2}\, \|x_0 - x^*\| < 1$.

Тоді справджується оцінка

$$\|x_k - x^*\| \le q^{3^k-1} \|x_0 - x^*\|, k = 0, 1, ..., \tag{9}$$

де x^* *- розв'язок задачі(1).*

Доведення. У цьому випадку справджується

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

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

Отже, тепер можна записати

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

$$\cdot \|(f(x_k, u_k)) - f(x_k, x^*))(x_k - x^*)\| \le BM_1 \|u_k - x^*\| \|x_k - x^*\| \le$$

$$\le (BM_1)^2(1 + \alpha M_2) \|x_k - x^*\|^3 .$$

Для доведення теореми також використаємо метод математичної індукції.
Для $k = 0$ маємо

$$\|x_1 - x^*\| \le (BM_1)^2(1 + \alpha M_2) \|x_0 - x^*\|^3 \le q^2 \|x_0 - x^*\| = q^{3^1-1} \|x_0 - x^*\| .$$

Нехай для $k \geq 1$ виконується

$$\|x_k - x^*\| \leq q^{3^k - 1} \|x_0 - x^*\|.$$

Тоді

$$\|x_{k+1} - x^*\| \leq (BM_1)^2(1 + \alpha M_2) \|x_k - x^*\|^3 \leq (BM_1)^2(1 + \alpha M_2) q^{3^{k+1} - 3} \|x_0 - x^*\|^3 \leq$$

$$\leq q^{3^{k+1} - 1} \|x_0 - x^*\|.$$

Теорема доведена, оцінка (9) виконується. □

Зауважимо, що при $n = 1$ індекс ефективності запропонованого алгоритму дорівнює $\sqrt[3]{3}$, що є більшим ніж для класичного методу Ньютона та різницевого методу, для яких індекс ефективності дорівнює $\sqrt{2}$. Якщо у (8) замінити сталий вибір множника α на вибір цього множника на кожному кроці з умови одновимірної мінімізації, ось такого вигляду $\alpha_k = argmin_\alpha f(x_k - \alpha f'(x_k))$, це дасть змогу зменшити знаменник збіжності методу Стеффенсена, а також для запропонованого нами алгоритму.

6. Висновок

Отже наведена методика побудови методів для розв'язування задачі безумовної мінімізації функцій багатьох змінних із надквадратичною швидкістю збіжності. Доведено теореми про швидкість збіжності у загальному випадку та при виборі конкретного оператора $\Phi(x)$.

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 // Вісник Львівського університету. Серія прикладна математика та інформатика. – 2007. – Вип. 13. – С. 3–10.
3. Васильев Ф.П. Методы оптимизации / Ф.П. Васильев. – Москва: Факториал, 2002.
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. – Том 16, № 1. – С. 13–26.

Стаття: надійшла до редколегії 22.09.2021
доопрацьована 20.10.2021
прийнята до друку 24.11.2021

ON SOME METHODS OF MINIMIZING A FUNCTION 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

Large number of descent methods are developed to solve the unconstrained minimization problem. The most popular and investigated are descent methods. Among the classical methods we distinguish gradient type methods. Quasi Newton methods. difference methods as one dimensional methods and methods which using second derivate matrix and building on Newton base. To solve a task minimization function there are a number of methods, these are methods of descent, the classical of which are gradient, Newtonian type methods, difference methods, methods of Quasi-Newtonian type, various modifications of the named methods, etc. The number of methods and their modifications is constantly increasing. Unfortunately, there is no universal method for solving such problems. Depending on the characteristics of the function, the requirements for the final result, the available computing power, etc., to solve a specific problem of minimization, choose one or another method, which in our opinion will be effective in terms of the requirements for the solution. We investigate and offer approach for building new class of methods for solving unconstrained minimization problem. It bases on the fact than on each iteration we can use information about calculated function's value of function using one dimensional method(especially gradient) and some other descent method (some of Newton type method).

Key words: Newton's method, divided difference, Steffensen's method.