УДК 519.853.6

Страхов Є. М.[1], к. ф.-м. н.,
Яровий А. Т.[2], к. ф.-м. н., доц.

Трикроковий метод для задачі багатовимірної оптимізації

[1] Одеський національний університет імені І. І. Мечникова, 65082, м. Одеса, вул. Дворянська, 2,
e-mail: swebus86@gmail.com
[2] Одеський національний університет імені І. І. Мечникова, 65082, м. Одеса, вул. Дворянська, 2,
e-mail: ouek@onu.edu.ua

E. M. Strakhov[1], Ph. D.,
A. T. Yarovyy[2], Ph. D.

A Three-Term Method for Unconstrained Optimization

[1] Odessa I. I. Mechnikov National University, 65082, Odessa, Dvoryanska str., 2,
e-mail: swebus86@gmail.com
[2] Odessa I. I. Mechnikov National University, 65082, Odessa, Dvoryanska str., 2,
e-mail: ouek@onu.edu.ua

У статті розглянуто трикроковий метод для розв'язування задач безумовної мінімізації функцій багатьох змінних. Доведено, що даний метод належить до методів спряжених напрямків. Проведено чисельний експеримент, який показує більшу ефективність запропонованого алгоритму порівняно із класичним двокроковим методом спряжених градієнтів.
Ключові слова: трикроковий метод, спряжені напрямки, задача безумовної оптимізації.

We consider an unconstrained minimization problem with a continuously differentiable function. Conjugate gradient methods (e. g. Fletcher – Reeves algorithm, Polak – Ribiere – Polyak algorithm) are very efficient for this branch of optimization. In this paper we suggest a three-term type modification of the classical conjugate gradient algorithm. Note that this modification needs the same number of function and gradient evaluations as the two-term method. We prove that the descent directions evaluated by the algorithm are conjugate for a quadratic function. Further we construct a minimization algorithm for nonquadratic functions. Some numerical results attained by our method are presented. These results demonstrate the advantages of three-term type algorithm. Namely, it requires less iteration to achieve greater accuracy. However, these benefits are less noticeable for large-scale problems. The study of three-term methods can be further generalized to the general case of k-term conjugate directions methods.
Key Words: three-term method, conjugate directions, unconstrained optimization.

Статтю представив д. т. н. Гаращенко Ф. Г.

Вступ

На сучасному етапі розвитку науки та комп'ютерних технологій теорія чисельних методів нелінійного програмування є досить розвиненою. Проте при виборі алгоритму для розв'язування конкретної реальної задачі дослідникам доводиться у значній мірі покладатися на досвід та інтуїцію. Це пов'язано з тим, що досить важко дати чіткі рекомендації щодо застосування того чи іншого методу; немає й «універсальних» методів, які можна було б використати у будь-якій ситуації. Тому теорія методів оптимізації продовжує розвиватися, з'являються нові методи, які мають свої переваги у порівнянні з попередніми.

У загальному вигляді задача нелінійного програмування (або задача оптимізації) полягає у визначенні екстремального значення деякої функції на допустимій множині, яка визначається набором обмежень. У випадку відсутності обмежень маємо задачу безумовної оптимізації. Такі задачі мають як самостійне прикладне значення, так і можуть використовуватися як допоміжні при розв'язуванні задач з обмеженнями та більш складних проблем.

Традиційно чисельні методи безумовної оптимізації поділяють на методи нульового (не потребують обчислення похідних від цільової функції), першого (використовується тільки перша похідна) та другого (відповідно, друга похідна) порядків. Окремим класом серед них є методи спуску. Вони мають ітераційну природу, тобто виходячи з деякої початкової точки будують послідовні наближення до точки мінімуму (максимуму), використовуючи напрямки спуску, тобто зменшення (зростання) функції. Щодо побудови цих напрямків розроблено різні підходи, які і визначають подальшу класифікацію методів. Зокрема, ті методи, які при побудові напрямків спуску використовують лише інформацію поточної ітерації, називають однокроковими. Якщо використовується додатково інформація попереднього кроку, то двокроковими, і т. д. Класичним двокроковим методом є метод спряжених градієнтів.

Розглянемо задачу

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

де $\varphi(x): R^n \to R^1$ є неперервно диференційовною функцією. Позначимо її градієнт через $\varphi'(x)$.

Метод спряжених градієнтів має такий загальний вигляд:

$$x^{k+1} = x^k + \beta_k s^k,\ k = 0,1,\dots,$$
$$s^k = \begin{cases} -\varphi'(x^k), & k = 0, \\ -\varphi'(x^k) + \gamma_{k-1} s^{k-1}, & k = 1, 2, \dots \end{cases} \tag{2}$$

де $x^0, x^1, \dots, x^k, \dots$ — послідовні наближення, $s^0, s^1, \dots, s^k, \dots$ — напрямки спуску, β_k — величина кроку вздовж напрямку спуску, γ_{k-1} — числовий параметр.

Одним із можливих варіантів вибору кроку β_k є розв'язування задачі одновимірної мінімізації:

$$\beta_k : \min_{\beta \ge 0} \varphi\left(x^k + \beta s^k\right). \tag{3}$$

На практиці також використовується вибір β_k з умов Вольфе [5]:

$$\varphi\left(x^k + \beta_k s^k\right) - \varphi\left(x^k\right) \le \delta \beta_k \left[\varphi'\left(x^k\right)\right]^T s^k,$$
$$\left[\varphi'\left(x^k + \beta_k s^k\right)\right]^T s^k \ge \sigma \left[\varphi'\left(x^k\right)\right]^T s^k,$$

де $0 < \delta \le \sigma < 1$ — деякі константи.

Практичні дослідження доводять, що якість розв'язку вихідної задачі (1) та швидкість збіжності алгоритму суттєво залежать від якості розв'язку одновимірної задачі (3).

Різновиди методу (2) визначаються способом обчислення параметру γ_{k-1}, зокрема

- $\gamma_{k-1} = \dfrac{\left\|\varphi'\left(x^k\right)\right\|^2}{\left\|\varphi'\left(x^{k-1}\right)\right\|^2}$ (метод Флетчера — Рівза);
- $\gamma_{k-1} = \dfrac{\left(\varphi'\left(x^k\right), \varphi'\left(x^k\right) - \varphi'\left(x^{k-1}\right)\right)}{\left\|\varphi'\left(x^{k-1}\right)\right\|^2}$ (метод Полака — Ріб'єра — Поляка)

та інші. Тут і далі $\|\cdot\|$ означає евклідову норму вектора, а $(\cdot,\cdot)$ — скалярний добуток векторів.

Метод спряжених градієнтів вважається досить ефективним для задач великої розмірності. Він у більшій мірі порівняно із однокроковими градієнтними методами враховує геометричні властивості цільової функції. Недоліком методу є чутливість до похибок обчислень, особливо при великій кількості змінних.

Зважаючи на переваги двокрокових методів перед однокроковими, можна очікувати, що врахування інформації, отриманої не тільки на попередньому, $(k-1)$-му, кроці, а й на $(k-2)$-му, дозволить побудувати ще більш ефективний алгоритм.

Теоретичні положення методу

Отже, повернемося до задачі нелінійного програмування (1). Для її розв'язання розглянемо трикроковий метод з таким алгоритмом:

$$x^{k+1} = x^k + \beta_k s^k,\ k = 0,1,\dots,$$
$$s^k = \begin{cases} -\varphi'\left(x^k\right), & k = 0, \\ -\varphi'\left(x^k\right) + \xi_{k-1} s^{k-1}, & k = 1, \\ -\varphi'\left(x^k\right) + \xi_{k-1} s^{k-1} + \gamma_{k-2} s^{k-2}, & k = 2,3,\dots \end{cases} \tag{4}$$

де, як і раніше, $x^0, \dots, x^k, \dots$ — послідовні наближення, $s^0, \dots, s^k, \dots$ — напрямки спуску, β_k, ξ_{k-1}, γ_{k-2} — числові параметри. Крок β_k будемо визначати з умови (3).

Означення. [2] Вектори s' і s'' називаються спряженими (відносно матриці A), якщо вони відмінні від нуля і $(As', s'') = 0$. Вектори $s^0, s^1, \dots, s^k$ називаються взаємно спряженими (відносно матриці A), якщо всі вони відмінні від нуля і $\left(As^i, s^j\right) = 0$, $i \ne j$, $0 \le i, j \le k$. Матриця

A вважається симетричною і додатно означеною ($A > 0$).

Розглянемо деякі властивості методу за умови, що функція $\varphi(x)$ є квадратичною

$$\varphi(x) = \frac{1}{2}(Ax, x) + (b, x) + c\,. \qquad (5)$$

Побудуємо систему взаємно спряжених напрямків за правилом (4).

$$0 = \left(s^k,\, As^{k-1}\right) = -\left(\varphi'\left(x^k\right),\, As^{k-1}\right) + \\ +\xi_{k-1}\left(s^{k-1},\, As^{k-1}\right) + \gamma_{k-2}\left(s^{k-2},\, As^{k-1}\right).$$

Звідси

$$\xi_{k-1} = \frac{\left(\varphi'\left(x^k\right),\, As^{k-1}\right)}{\left(s^{k-1},\, As^{k-1}\right)}\,. \qquad (6)$$

Знаменник у (6) не дорівнює нулеві, так як матриця A додатно означена.

Далі,

$$0 = \left(s^k,\, As^{k-2}\right) = -\left(\varphi'\left(x^k\right),\, As^{k-2}\right) + \\ +\xi_{k-1}\left(s^{k-1},\, As^{k-2}\right) + \gamma_{k-2}\left(s^{k-2},\, As^{k-2}\right).$$

Отримаємо

$$\gamma_{k-2} = \frac{\left(\varphi'\left(x^k\right),\, As^{k-2}\right)}{\left(s^{k-2},\, As^{k-2}\right)}\,. \qquad (7)$$

Теорема 1. Для диференційовної функції $\varphi(x)$ послідовність $\left\{x^k\right\}$, що побудована за (4), (6), (7), така, що

$$\left(\varphi'\left(x^{k+1}\right),\, s^k\right) = 0,\; k = 0,1,\ldots \qquad (8)$$

Доведення. Враховуючи (4), отримаємо $Ax^{k+1} = Ax^k + \beta_k As^k$; разом з $\varphi'(x) = Ax + b$ маємо

$$\varphi'\left(x^{k+1}\right) = \varphi'\left(x^k\right) + \beta_k As^k\,. \qquad (9)$$

З (3) отримаємо, що при $\beta_k > 0$

$$\frac{d}{d\beta}\varphi\left(x^k + \beta s^k\right)\Big|_{\beta=\beta_k} = 0\,,$$

а при $\beta_k = 0$

$$\frac{d}{d\beta}\varphi\left(x^k + \beta s^k\right)\Big|_{\beta=0} \geq 0\,.$$

Якщо $\beta_k > 0$, то

$$0 = \frac{d}{d\beta}\varphi\left(x^k + \beta s^k\right)\Big|_{\beta=\beta_k} = \\ = \left(\varphi'\left(x^k + \beta_k s^k\right),\, s^k\right) = \left(\varphi'\left(x^{k+1}\right),\, s^k\right).$$

Отже, $\left(\varphi'\left(x^{k+1}\right),\, s^k\right) = 0$.

Доведення того, що співвідношення (8) справедливе і при $\beta_k = 0$, проведемо за індукцією. Маємо

$$0 \leq \frac{d}{d\beta}\varphi\left(x^0 + \beta s^0\right)\Big|_{\beta=0} = \left(\varphi'\left(x^1\right),\, s^0\right) = -\left\|\varphi'\left(x^0\right)\right\|^2,$$

звідки отримаємо рівність $\left(\varphi'\left(x^1\right),\, s^0\right) = 0$.

Нехай $\left(\varphi'\left(x^k\right),\, s^{k-1}\right) = 0$ і доведемо, що $\left(\varphi'\left(x^{k+1}\right),\, s^k\right) = 0$ при $\beta_k = 0$. Так як $x^{k+1} = x^k$, то маємо, враховуючи (9)

$$0 \leq \frac{d}{d\beta}\varphi\left(x^k + \beta s^k\right)\Big|_{\beta=0} = \\ = \left(\varphi'\left(x^{k+1}\right),\, s^k\right) = \left(\varphi'\left(x^k\right),\, s^k\right) = \\ = \left(\varphi'\left(x^k\right), -\varphi'\left(x^k\right) + \xi_{k-1}s^{k-1} + \gamma_{k-2}s^{k-2}\right) = \\ = -\left\|\varphi'\left(x^k\right)\right\|^2 + \gamma_{k-2}\left(\varphi'\left(x^k\right),\, s^{k-2}\right) = \\ = -\left\|\varphi'\left(x^k\right)\right\|^2 + \gamma_{k-2}\left(\varphi'\left(x^{k-1}\right) + \beta_{k-1}As^{k-1},\, s^{k-2}\right) = \\ = -\left\|\varphi'\left(x^k\right)\right\|^2 + \gamma_{k-2}\left(\varphi'\left(x^{k-1}\right),\, s^{k-2}\right) + \\ + \gamma_{k-2}\beta_{k-1}\left(As^{k-1},\, s^{k-2}\right) = -\left\|\varphi'\left(x^k\right)\right\|^2 \leq 0\,.$$

Отримали $\left(\varphi'\left(x^{k+1}\right), s^k\right) = 0$. Теорему доведено.

Доведемо, що метод (4), (6), (7) відноситься до методів спряжених напрямків.

Теорема 2. Вектори $\varphi'\left(x^k\right)$ і $\varphi'\left(x^{k+1}\right)$ ортогональні, $k = 0,1,\ldots$.

Доведення. Відомо [2], що квадратична функція (5) досягає мінімального значення при

$$\beta_k = -\frac{\left(\varphi'\left(x^k\right),\, s^k\right)}{\left(As^k,\, s^k\right)}\,. \qquad (10)$$

Тоді, враховуючи (9) і (10), отримаємо

$$\left(\varphi'\left(x^{k+1}\right),\, \varphi'\left(x^k\right)\right) = \\ = \left(\varphi'\left(x^k\right) - \frac{\left(\varphi'\left(x^k\right),\, s^k\right)}{\left(As^k,\, s^k\right)}As^k,\, \varphi'\left(x^k\right)\right) =$$

$$= \left(\varphi'\left(x^k \right), \varphi'\left(x^k \right) \right) - \frac{\left(\varphi'\left(x^k \right), s^k \right)}{\left(As^k, s^k \right)} \left(As^k, \varphi'\left(x^k \right) \right).$$

Розглянемо $\left(\varphi'\left(x^k \right), s^k \right)$. Отже,

$$\left(\varphi'\left(x^k \right), s^k \right) = -\left(\varphi'\left(x^k \right), \varphi'\left(x^k \right) \right) + \\ +\xi_{k-1}\left(\varphi'\left(x^k \right), s^{k-1} \right) + \gamma_{k-2}\left(\varphi'\left(x^k \right), s^{k-2} \right) = \\ = -\left(\varphi'\left(x^k \right), \varphi'\left(x^k \right) \right) + \gamma_{k-2}\left(\varphi'\left(x^k \right), s^{k-2} \right).$$

Далі, враховуючи (9), маємо $\left(\varphi'\left(x^k \right), s^{k-2} \right) =$ $= \left(\varphi'\left(x^{k-1} \right), s^{k-2} \right) + \beta_{k-1}\left(As^{k-1}, s^{k-2} \right) = 0$.
Тоді

$$\left(\varphi'\left(x^k \right), s^k \right) = -\left(\varphi'\left(x^k \right), \varphi'\left(x^k \right) \right). \qquad (11)$$

Маємо

$$\left(As^k, s^k \right) = -\left(As^k, \varphi'\left(x^k \right) \right) + \xi_{k-1}\left(As^k, s^{k-1} \right) + \\ + \gamma_{k-2}\left(As^k, s^{k-2} \right) = -\left(As^k, \varphi'\left(x^k \right) \right). \qquad (12)$$

З урахуванням (11) і (12) отримаємо

$$\left(\varphi'\left(x^{k+1} \right), \varphi'\left(x^k \right) \right) = \left(\varphi'\left(x^k \right), \varphi'\left(x^k \right) \right) - \\ - \frac{-\left(\varphi'\left(x^k \right), \varphi'\left(x^k \right) \right)}{-\left(As^k, \varphi'\left(x^k \right) \right)} \left(As^k, \varphi'\left(x^k \right) \right) = 0.$$

Теорему доведено.

Теорема 3. Нехай $x^0 \in R^n$, точки $x^1, \ldots, x^{n-1}$ і вектори $s^0, s^1, \ldots, s^{n-1}$ отримані за формулами (4), (6), (7) і $\varphi'\left(x^k \right) \neq 0$ $(i = 0, \ldots, n-1)$. Тоді вектори $s^0, s^1, \ldots, s^{n-1}$ взаємно спряжені, а градієнти $\varphi'\left(x^0 \right)$, $\varphi'\left(x^1 \right)$, $\ldots$, $\varphi'\left(x^{n-1} \right)$ взаємно ортогональні.

Доведення. Теорему будемо доводити методом математичної індукції. Ортогональність векторів $\varphi'\left(x^0 \right)$ і $\varphi'\left(x^1 \right)$ отримаємо за теоремою 2. Вектор $s^0 \neq 0$ за умовою, s^1 теж не дорівнює нулеві, інакше $s^1 = -\varphi'\left(x^1 \right) - \beta_0 \varphi'\left(x^0 \right) = 0$, а це неможливо, враховуючи ортогональність $\varphi'\left(x^0 \right)$ і $\varphi'\left(x^1 \right)$. Спряженість s^0 і s^1 отримаємо за (6) і (7).

Припустимо, що $k \leq n-1$, вектори $s^0, s^1, \ldots, s^{k-1}$ взаємно спряжені, а градієнти $\varphi'\left(x^0 \right), \ldots, \varphi'\left(x^{k-1} \right)$ взаємно ортогональні. Тоді за теоремою 2 $\left(\varphi'\left(x^k \right), \varphi'\left(x^{k-1} \right) \right) = 0$. При $i \leq k-2$, використовуючи (9), індукцію та (4), маємо

$$\left(\varphi'\left(x^k \right), \varphi'\left(x^i \right) \right) = \\ = \left(\varphi'\left(x^{k-1} \right), \varphi'\left(x^i \right) \right) + \beta_{k-1}\left(As^{k-1}, \varphi'\left(x^i \right) \right) = \\ = \beta_{k-1}\left(As^{k-1}, -s^i + \xi_{i-1}s^{i-1} + \gamma_{i-2}s^{i-2} \right) = 0.$$

Отже, доведена взаємна ортогональність векторів $\varphi'\left(x^0 \right)$, $\varphi'\left(x^1 \right)$, $\ldots$, $\varphi'\left(x^{n-1} \right)$.

Вектор $s^k \neq 0$, інакше з (4) отримали б, що вектори $\varphi'\left(x^0 \right)$, $\ldots$, $\varphi'\left(x^k \right)$ лінійно залежні, що суперечить їх взаємній ортогональності. Доведемо, що $s^0, \ldots, s^k$ взаємно спряжені.

Дійсно, $\left(s^k, As^{k-1} \right) = 0$ за (6), (7). Враховуючи (10), маємо

$$\beta_i = -\frac{\left(\varphi'\left(x^i \right), -\varphi'\left(x^i \right) - \xi_{i-1}\varphi'\left(x^{i-1} \right) - \ldots \right)}{\left(As^i, s^i \right)} = \\ = \frac{\left(\varphi'\left(x^i \right), \varphi'\left(x^i \right) \right)}{\left(As^i, s^i \right)},$$

і тому $\beta_i \neq 0$, $i \leq k$. Тоді з (9) отримаємо

$$As^i = \frac{\varphi'\left(x^{i+1} \right) - \varphi'\left(x^i \right)}{\beta_i}. \qquad (13)$$

При $i \leq k-2$, використовуючи (4), індукцію, (13) та доведену взаємну ортогональність градієнтів, отримаємо

$$\left(s^k, As^i \right) = \left(-\varphi'\left(x^k \right) + \xi_{k-1}s^{k-1} + \gamma_{k-2}s^{k-2}, As^i \right) = \\ = -\left(\varphi'\left(x^k \right), \frac{\varphi'\left(x^{i+1} \right) - \varphi'\left(x^i \right)}{\beta_i} \right) = 0.$$

Теорему доведено.

Отже, розглянутий метод (4), (6), (7) належить до методів спряжених напрямків.

Сформулюємо тепер трикроковий метод для мінімізації неквадратичних функцій. Для цього перетворимо формули (6) і (7) так, щоб до них не входила матриця A.

Перетворимо ξ_{k-1}:

$$\xi_{k-1} = \frac{\left(\varphi'\left(x^k \right), As^{k-1} \right)}{\left(s^{k-1}, As^{k-1} \right)} =$$

$$=\frac{\left(\varphi'\left(x^k\right),\varphi'\left(x^k\right)-\varphi'\left(x^{k-1}\right)\right)}{\left(s^{k-1},\varphi'\left(x^k\right)-\varphi'\left(x^{k-1}\right)\right)}=$$

$$=\frac{\left(\varphi'\left(x^k\right),\varphi'\left(x^k\right)-\varphi'\left(x^{k-1}\right)\right)}{\left(-\varphi'\left(x^{k-1}\right)-\xi_{k-2}\varphi'\left(x^{k-2}\right)+\ldots,\varphi'\left(x^k\right)-\varphi'\left(x^{k-1}\right)\right)}=$$

$$=\frac{\left(\varphi'\left(x^k\right),\varphi'\left(x^k\right)-\varphi'\left(x^{k-1}\right)\right)}{\left\|\varphi'\left(x^{k-1}\right)\right\|^2}.$$

Далі аналогічно отримаємо

$$\gamma_{k-2}=\frac{\left(\varphi'\left(x^k\right),As^{k-2}\right)}{\left(s^{k-2},As^{k-2}\right)}=$$

$$=\frac{\left(\varphi'\left(x^k\right),\varphi'\left(x^{k-1}\right)-\varphi'\left(x^{k-2}\right)\right)}{\left\|\varphi'\left(x^{k-2}\right)\right\|^2}.$$

Отже, для неквадратичних функцій будемо використовувати трикроковий метод (4), де крок β_k визначається з умови (3), а параметри ξ_{k-1}, γ_{k-2} обчислюються наступним чином:

$$\xi_{k-1}=\frac{\left(\varphi'\left(x^k\right),\varphi'\left(x^k\right)-\varphi'\left(x^{k-1}\right)\right)}{\left\|\varphi'\left(x^{k-1}\right)\right\|^2}, \quad (14)$$

$$\gamma_{k-2}=\frac{\left(\varphi'\left(x^k\right),\varphi'\left(x^{k-1}\right)-\varphi'\left(x^{k-2}\right)\right)}{\left\|\varphi'\left(x^{k-2}\right)\right\|^2}. \quad (15)$$

Обчислювальний експеримент

Зробимо аналіз роботи описаного трикрокового методу порівняно із класичним методом спряжених градієнтів. 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 трикроковому методі не змінюється порівняно із двокроковим.

Наведемо результати розрахунків для декількох тестових задач [3, 4]. Обчислення проводилися із точністю $\varepsilon=10^{-6}$. Критерій зупинки процесу обчислень полягав у одночасному виконанні трьох умов:

$$\varphi\left(x^{k-1}\right)-\varphi\left(x^k\right)<\varepsilon\left(1+\left|\varphi\left(x^k\right)\right|\right),$$
$$\left\|x^{k-1}-x^k\right\|<\sqrt{\varepsilon}\left(1+\left\|x^k\right\|\right),$$
$$\left\|\varphi'\left(x^k\right)\right\|\le\sqrt[3]{\varepsilon}\left(1+\left|\varphi\left(x^k\right)\right|\right).$$

Для розрахунків використовувалася відкрита система комп'ютерної математики Sage (www.sagemath.org).

Задача 1. Кількість змінних $n=3$:

$$\varphi(x)=100\left[x_3-\left(\frac{x_1+x_2}{2}\right)^2\right]^2+\left(1-x_1\right)^2+\left(1-x_2\right)^2,$$

початкові точки $x_0^1=(-1.2,\,2,\,0)^T$, $\varphi(x_0^1)=8.40$ та $x_0^2=(-2,\,2,\,4)^T$, $\varphi(x_0^2)=1610$; точний розв'язок задачі $x^*=(1,1,1)^T$, $\varphi(x^*)=0$.

Таблиця 1

Результати обчислень для задачі 1

Початкова точка x_0^1	МСГ	Трикроковий метод
Значення функції	$3{,}27\cdot10^{-5}$	$9{,}78\cdot10^{-8}$
Норма градієнту	0,0086	0,0042
К-ть ітерацій	150	34

Задача 2. Функція Пауелла, кількість змінних $n=4$:

$$\phi(x)=(x_1+10x_2)^2+5(x_3-x_4)^2+$$
$$+(x_2-2x_3)^4+10(x_1-x_4)^4,$$

початкові точки $x_0^1=(3,-1,\,0,\,1)^T$, $\varphi(x_0^1)=215$ та $x_0^2=(1,1,1,1)^T$, $\varphi(x_0^1)=125$; точний розв'язок задачі $x^*=(0,0,0,0)^T$, $\varphi(x^*)=0$.

Таблиця 2

Результати обчислень для задачі 2

Початкова точка x_0^1	МСГ	Трикроковий метод
Значення функції	$1{,}05\cdot10^{-5}$	$2{,}32\cdot10^{-8}$
Норма градієнту	0,0094	0,0009
К-ть ітерацій	58	33

Задача 3. Узагальнена функція Розенброка, кількість змінних n:

$$\phi(x)=\sum_{i=1}^{n/2}\left[\left(1-x_{2i-1}\right)^2+100\left(x_{2i}-x_{2i-1}^2\right)^2\right],$$

початкова точка $x_0=(-1.2,1,-1.2,1,\ldots)^T$, точний розв'язок задачі $x^*=(1,\ldots,1)^T$, $\varphi(x^*)=0$.

Таблиця 3
Результати обчислень для задачі 3 при $n = 8$

Початкова точка x_0	МСГ	Трикроковий метод
Значення функції	$2{,}47 \cdot 10^{-6}$	$2{,}34 \cdot 10^{-6}$
Норма градієнту	0,0098	0,0056
К-ть ітерацій	152	60

Таблиця 4
Результати обчислень для задачі 3 при $n = 20$

Початкова точка x_0	МСГ	Трикроковий метод
Значення функції	$1{,}70 \cdot 10^{-6}$	$2{,}07 \cdot 10^{-6}$
Норма градієнту	0,00984	0,00978
К-ть ітерацій	283	268

Задача 4. Узагальнена функція Біла (Beale), кількість змінних $n = 100$:

$$\varphi(x) = \sum_{i=1}^{n/2} \Big[\left(1.5 - x_{2i-1}\left(1 - x_{2i}\right)\right)^2 +$$
$$+\left(2.25 - x_{2i-1}\left(1 - x_{2i}^2\right)\right)^2 + \left(2.625 - x_{2i-1}\left(1 - x_{2i}^3\right)\right)^2 \Big],$$

початкова точка $x_0 = (1, 0.8, \ldots, 1, 0.8)^T$.

Таблиця 5
Результати обчислень для задачі 4

Початкова точка x_0	МСГ	Трикроковий метод
Значення функції	$4{,}85 \cdot 10^{-8}$	$4{,}85 \cdot 10^{-8}$
Норма градієнту	0,00198	0,00198
К-ть ітерацій	11	8

Висновки

У роботі розглянуто трикроковий метод для задач безумовної мінімізації, обґрунтована його належність до методів спряжених напрямків. Результати обчислювальних експериментів для тестових функцій вказують на переваги запропонованого методу у порівнянні з класичним двокроковим методом спряжених градієнтів. При невеликій кількості змінних ($n \le 20$) трикроковий алгоритм дає більш точний розв'язок при суттєво меншій кількості кроків. Для задач великої розмірності ($n = 100, 200$) вказані переваги також спостерігаються, але в меншій мірі. Слід зазначити, що кількість обчислень функції та градієнту в обох методах однакова. Зважаючи на отримані результати, перспективним може бути дослідження властивостей загального *k*-крокового алгоритму.

Список використаних джерел

1. *Карманов В. Г.* Математическое программирование / В. Г. Карманов. — Москва: ФИЗМАТЛИТ, 2004. — 264 с.
2. *Сухарев А. Г.* Курс методов оптимизации / А. Г. Сухарев, А. В. Тимохов, В. В. Федоров. — Москва: ФИЗМАТЛИТ, 2005. — 368 с.
3. *Химмельблау Д.* Прикладное нелинейное программирование / Д. Химмельблау. — Москва: Мир, 1972. — 536 с.
4. *Andrei N.* An Unconstrained Optimization Test Functions Collection / N. Andrei // Advanced Modelling and Optimization. — 2008. — Vol. 10, No. 1. — P. 147–161.
5. *Sun Z., Tian Y., Li H.* Two Modified Three-Term Type Conjugate Gradient Methods and Their Global Convergence for Unconstrained Optimization [Електронний ресурс] / Zhongbo Sun, Yantao Tian, Hongyang Li // Mathematical Problems in Engineering. — 2014. — 9 P. – Режим доступу до журн.: http://dx.doi.org/10.1155/2014/394096.

References

1. KARMANOV, V. (2004) *Matematicheskoe programmirovaniye*. Moskva: Fizmatlit.
2. SUKHAREV A., TIMOKHOV A., FEDOROV V. (2005) *Kurs metodov optimizatsii*. Moskva: Fizmatlit.
3. HIMMELBLAU D. (1972) *Prikladnoye nelineinoye programmirovaniye.* Moskva: Mir.
4. ANDREI, N. (2008) An Unconstrained Optimization Test Functions Collection. *Advanced Modeling and Optimization.* 10 (1). P. 147–161.
5. SUN Z., TIAN Y., LI H. (2014) Two Modified Three-Term Type Conjugate Gradient Methods and Their Global Convergence for Unconstrained Optimization. *Mathematical Problems in Engineering* [Online]. — Available from: http://dx.doi.org/10.1155/2014/394096.

Надійшла до редколегії 30.04.15