УДК 519.853.6

Є. М. Страхов, А. Т. Яровий
Одеський національний університет імені І. І. Мечникова

АНАЛІЗ P-КРОКОВИХ МЕТОДІВ МІНІМІЗАЦІЇ ФУНКЦІЙ БАГАТЬОХ ЗМІННИХ

Розглядається задача багатовимірної мінімізації неперервно диференційовної функції при відсутності обмежень. Ітераційний алгоритм розв'язування такої задачі називається багатокроковим, якщо для знаходження наступного наближення до точки мінімуму використовуються значення функції або її градієнта у двох або більше попередніх точках. Так, алгоритм методу спряжених градієнтів належить до двокрокових. Описується узагальнений p-кроковий алгоритм, встановлені його властивості у випадку квадратичної цільової функції. Показано, що даний метод належить до методів спряжених напрямків. Метою обчислювального експерименту було порівняння результатів мінімізації у залежності від кількості доданків (кроків) p та виявлення «оптимального» значення для p. Наводяться результати обчислень для деяких відомих тестових функцій.
MSC: 90C30, 49M20.
Ключові слова: p-кроковий алгоритм, спряжені напрямки, задача безумовної оптимізації .

Вступ. Розглянемо задачу нелінійного програмування

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

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

Одним із класичних підходів до розв'язування задачі (1) є метод спряжених градієнтів. Його алгоритм має такий загальний вигляд:

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

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

Для обчислення напрямку спуску s^{k+1} у алгоритмі (2) використовується не лише інформація поточного k-го, але й попереднього, $(k-1)$-го кроку. Отже, метод (2) належить до двокрокових методів.

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

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

Так як розв'язування задачі (3) ускладнює процес пошуку точки мінімуму, можна піти іншим шляхом: ітеративно підбирати величину кроку так, щоб ця величина задовольняла певну умову. Однією із поширених є умова Вольфе:

$$\begin{gathered} \varphi\left(x^k + \beta_k s^k\right) - \varphi\left(x^k\right) \leqslant \delta \beta_k \left[\varphi'\left(x^k\right)\right]^\top s^k, \\ \left[\varphi'\left(x^k + \beta_k s^k\right)\right]^\top s^k \geqslant \sigma \left[\varphi'\left(x^k\right)\right]^\top s^k, \end{gathered} \tag{4}$$

Надійшла 16.02.2017

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

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

- $\gamma_{k-1} = \frac{\left\|\varphi'\left(x^k\right)\right\|^2}{\|\varphi'(x^{k-1})\|^2}$ (метод Флетчера — Рівза);
- $\gamma_{k-1} = \frac{\left(\varphi'\left(x^k\right), \varphi'\left(x^k\right) - \varphi'\left(x^{k-1}\right)\right)}{\|\varphi'(x^{k-1})\|^2}$ (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 великого виміру. Він у більшій мірі порівняно із однокроковими методами враховує геометричні властивості цільової функції. Недоліком методу є чутливість до похибок обчислень, особливо при великій кількості змінних.

У статті [4] нами було розглянуто трикроковий метод

$$x^{k+1} = x^k + \beta_k s^k, \quad k = 0{,}1, \dots,$$
$$s^k = \begin{cases} -\varphi'\left(x^k\right), & k = 0, \\ -\varphi'\left(x^k\right) + \xi_{k-1} s^{k-1}, & k = 1, \\ -\varphi'\left(x^k\right) + \xi_{k-1} s^{k-1} + \gamma_{k-2} s^{k-2}, & k = 2, 3, \dots \end{cases} \tag{5}$$

та показано його переваги у порівнянні з методом спряжених градієнтів на тестових функціях переважно із невеликою кількістю змінних ($n \leqslant 20$). Також слід відзначити, що використання ще одного доданку при визначенні напрямку спуску не збільшує кількість обчислень функції та її градієнта на одній ітерації.

Метою даної роботи є узагальнення алгоритму (5) на будь-яку кількість доданків p.

Основні результати

1. Теоретичні положення методу.

Отже, повернемося до задачі нелінійного програмування (1). Розглянемо узагальнений p-кроковий метод з таким алгоритмом:

$$x^{k+1} = x^k + \beta_k s^k,\ k = 0, 1, \dots,$$
$$s^k = \begin{cases} -\varphi'\left(x^k\right), & k = 0, \\ -\varphi'\left(x^k\right) + \gamma_{k-1} s^{k-1} + \dots + \gamma_{k-(p-1)} s^{k-(p-1)}, & k = 1, 2, \dots, \end{cases} \tag{6}$$

де, як і раніше, $x^0, \dots, x^k, \dots$ — послідовні наближення, $s^0, \dots, s^k, \dots$ — напрямки спуску, β_k, $\gamma_{k-1}, \dots, \gamma_{k-(p-1)}$ — числові параметри. Крок β_k будемо визначати з умови (3) або (4). Зазначимо, що при $p = 2$ отримаємо класичний метод спряжених градієнтів, а при $p = 3$ — метод (5).

Означення 1. [5] *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 $\left(As^i, s^j\right) = 0$, $i \neq j$, $0 \leqslant i, j \leqslant k$. Матриця A вважається симетричною і додатно означеною ($A > 0$).*

Розглянемо деякі властивості методу (6) за умови, що функція $\varphi(x)$ є квадратичною, тобто

$$\varphi(x) = \frac{1}{2}(Ax, x) + (b, x) + c. \tag{7}$$

Побудуємо систему взаємно спряжених напрямків за правилом (6).

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

Звідси

$$\gamma_{k-1} = \frac{\left(\varphi'\left(x^k\right), As^{k-1}\right)}{\left(s^{k-1}, As^{k-1}\right)}. \tag{8}$$

Знаменник у (8) не дорівнює нулеві, так як матриця A додатно означена.

Далі отримаємо

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

Враховуючи взаємну спряженість векторів s^j, отримаємо

$$\gamma_{k-2} = \frac{\left(\varphi'\left(x^k\right), As^{k-2}\right)}{\left(s^{k-2}, As^{k-2}\right)}. \tag{9}$$

Аналогічно можна отримати, що

$$\gamma_{k-i} = \frac{\left(\varphi'\left(x^k\right), As^{k-i}\right)}{\left(s^{k-i}, As^{k-i}\right)}, \quad i = 3, 4, \ldots, p-1. \tag{10}$$

Теорема 1. *Для квадратичної функції $\varphi(x)$ послідовність $\left\{x^k\right\}$, що побудована за алгоритмом (6), (3), (8), (9), (10), є такою, що*

$$\left(\varphi'\left(x^{k+1}\right), s^k\right) = 0, \quad k = 0, 1, \ldots. \tag{11}$$

Доведення. Враховуючи (6), отримаємо

$$Ax^{k+1} = Ax^k + \beta_k As^k.$$

Так як $\varphi'(x) = Ax + b$, то остаточно маємо

$$\varphi'\left(x^{k+1}\right) = \varphi'\left(x^k\right) + \beta_k As^k. \tag{12}$$

З (3) отримаємо, що при $\beta_k > 0$

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

а при $\beta_k = 0$

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

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

Доведення того, що співвідношення (11) справедливе і при $\beta_k = 0$, проводимо за індукцією.

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

звідки отримаємо рівність

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

Нехай $\left(\varphi'\left(x^k\right), s^{k-1}\right) = 0$. Доведемо, що $\left(\varphi'\left(x^{k+1}\right), s^k\right) = 0$ при $\beta_k = 0$. Так як $x^{k+1} = x^k$ $(\beta_k = 0)$, то, враховуючи (6), отримаємо

$$\begin{aligned}
0 \leqslant \frac{d}{d\beta}\varphi\left(x^k + \beta s^k\right)\Big|_{\beta=0} &= \left(\varphi'\left(x^{k+1}\right), s^k\right) = \left(\varphi'\left(x^k\right), s^k\right) = \\
&= \left(\varphi'\left(x^k\right), -\varphi'\left(x^k\right) + \gamma_{k-1}s^{k-1} + \ldots + \gamma_{k-(p-1)}s^{k-(p-1)}\right) = \\
&= -\left\|\varphi'\left(x^k\right)\right\|^2 + \gamma_{k-1}\left(\varphi'\left(x^k\right), s^{k-1}\right) + \ldots + \gamma_{k-i}\left(\varphi'\left(x^k\right), s^{k-i}\right) + \\
&\quad + \ldots + \gamma_{k-(p-1)}\left(\varphi'\left(x^k\right), s^{k-(p-1)}\right).
\end{aligned}$$

Враховуючи (12) і припущення індукції, маємо

$$\begin{aligned}
\left(\varphi'\left(x^k\right), s^{k-i}\right) &= \left(\varphi'\left(x^{k-1}\right) + \beta_{k-1}As^{k-1}, s^{k-i}\right) = \\
&= \left(\varphi'\left(x^{k-1}\right), s^{k-i}\right) + \beta_{k-1}\left(As^{k-1}, s^{k-i}\right) = \\
&= \left(\varphi'\left(x^{k-1}\right), s^{k-i}\right) = \ldots = \left(\varphi'\left(x^{k-i+1}\right), s^{k-i}\right) = 0.
\end{aligned}$$

Отже, ми отримали

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

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

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

Доведення. Відомо [3], що квадратична функція (7) досягає мінімального значення при

$$\beta_k = -\frac{\left(\varphi'\left(x^k\right), s^k\right)}{\left(As^k, s^k\right)}. \tag{13}$$

Тоді, враховуючи (12) і (13), отримаємо

$$\begin{aligned}
\left(\varphi'\left(x^{k+1}\right), \varphi'\left(x^k\right)\right) &= \left(\varphi'\left(x^k\right) + \beta_k As^k, \varphi'\left(x^k\right)\right) = \\
&= \left(\varphi'\left(x^k\right) - \frac{\left(\varphi'\left(x^k\right), s^k\right)}{\left(As^k, s^k\right)}As^k, \varphi'\left(x^k\right)\right) = \\
&= \left(\varphi'\left(x^k\right), \varphi'\left(x^k\right)\right) - \frac{\left(\varphi'\left(x^k\right), s^k\right)}{\left(As^k, s^k\right)}\left(As^k, \varphi'\left(x^k\right)\right).
\end{aligned}$$

Розглянемо $\left(\varphi'\left(x^k\right), s^k\right)$. У попередній теоремі ми довели, що

$$\left(\varphi'\left(x^k\right), s^k\right) = -\left(\varphi'\left(x^k\right), \varphi'\left(x^k\right)\right). \tag{14}$$

Далі,

$$\begin{gathered}\left(As^k, s^k\right) = \left(As^k, -\varphi'\left(x^k\right) + \gamma_{k-1}s^{k-1} + \ldots + \gamma_{k-(p-1)}s^{k-(p-1)}\right) = \\ = -\left(As^k, \varphi'\left(x^k\right)\right) + \gamma_{k-1}\left(As^k, s^{k-1}\right) + \ldots + \gamma_{k-(p-1)}\left(As^k, s^{k-(p-1)}\right) = \\ = -\left(As^k, \varphi'\left(x^k\right)\right).\end{gathered} \tag{15}$$

З урахуванням (14) і (15), отримаємо

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

Теорему доведено.

Теорема 3. *Нехай $x^0 \in \mathbb{R}^n$, точки $x^1, \ldots, x^{n-1}$ і вектори $s^0, s^1, \ldots, s^{n-1}$ отримані за формулами (6), (3), (8), (9), (10) 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 $\varphi'\left(x^0\right)$, $\varphi'\left(x^1\right)$, …, $\varphi'\left(x^{n-1}\right)$ взаємно ортогональні.*

Доведення. Теорему доведемо 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 теореми; вектор s^1 теж не дорівнює нулеві, так як $s^1 = -\varphi'\left(x^1\right) - \gamma_0\varphi'\left(x^0\right) = 0$, а це неможливо, враховуючи ортогональність $\varphi'\left(x^0\right)$ і $\varphi'\left(x^1\right)$. Спряженість s^0 і s^1 отримаємо з (8) і (9).

Припустимо, що $k \leqslant n-1$, вектори $s^0, s^1, \ldots, s^{k-1}$ взаємно спряжені, а градієнти $\varphi'\left(x^0\right)$, $\varphi'\left(x^1\right)$, …, $\varphi'\left(x^{k-1}\right)$ взаємно ортогональні. Тоді за теоремою 2 $\left(\varphi'\left(x^k\right), \varphi'\left(x^{k-1}\right)\right) = 0$. При $i \leqslant k-2$ маємо, використовуючи (12), індукцію та (6),

$$\begin{gathered}\left(\varphi'\left(x^k\right), \varphi'\left(x^i\right)\right) = \left(\varphi'\left(x^{k-1}\right), \varphi'\left(x^i\right)\right) + \beta_{k-1}\left(As^{k-1}, \varphi'\left(x^i\right)\right) = \\ = \beta_{k-1}\left(As^{k-1}, -s^i + \gamma_{i-1}s^{i-1} + \ldots + \gamma_{i-(p-1)}s^{i-(p-1)}\right) = 0.\end{gathered}$$

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

Вектор $s^k \neq 0$, інакше з (6) отримали б, що вектори $\varphi'\left(x^0\right), \varphi'\left(x^1\right), \ldots, \varphi'\left(x^k\right)$ лінійно залежні, що суперечить їх взаємній ортогональності.

Доведемо, що вектори $s^0, \ldots, s^k$ взаємно спряжені. Дійсно, $\left(s^k, As^{k-1}\right) = 0$ за (8). Враховуючи (13), отримаємо

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

і тому $\beta_i \neq 0$, $i \leqslant k$. Тоді з (12) отримаємо

$$As^i = \frac{\varphi'\left(x^{i+1}\right) - \varphi'\left(x^i\right)}{\beta_i}. \tag{16}$$

При $i \leqslant k-2$, використовуючи (6), індукцію, (16) та доведену взаємну ортогональність градієнтів, отримаємо

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

Теорему доведено.

Отже, розглянутий p-кроковий метод належить до методів спряжених напрямків.

Сформулюємо тепер p-кроковий метод для мінімізації неквадратичних функцій. Для цього перетворимо формулу (10) так, щоб до неї не входила матриця A:

$$\gamma_{k-i} = \frac{\left(\varphi'\left(x^k\right), As^{k-i}\right)}{\left(s^{k-i}, As^{k-i}\right)} = \frac{\left(\varphi'\left(x^k\right), \varphi'\left(x^{k-i+1}\right) - \varphi'\left(x^{k-i}\right)\right)}{\left(s^{k-i}, \varphi'\left(x^{k-i+1}\right) - \varphi'\left(x^{k-i}\right)\right)} =$$
$$= \frac{\left(\varphi'\left(x^k\right), \varphi'\left(x^{k-i+1}\right) - \varphi'\left(x^{k-i}\right)\right)}{-\left(\varphi'\left(x^{k-i}\right) - \gamma_{k-i-1}\varphi'\left(x^{k-i-1}\right) - \ldots, \varphi'\left(x^{k-i+1}\right) - \varphi'\left(x^{k-i}\right)\right)} = \tag{17}$$
$$= \frac{\left(\varphi'\left(x^k\right), \varphi'\left(x^{k-i+1}\right) - \varphi'\left(x^{k-i}\right)\right)}{\left\|\varphi'\left(x^{k-i}\right)\right\|^2}.$$

2. Обчислювальний експеримент.

Зробимо аналіз роботи описаного p-крокового методу при різних значеннях p. 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 p-кроковому методі не змінюється порівняно із класичним двокроковим методом спряжених градієнтів.

Обчислення проводилися на тестових функціях із точністю $\varepsilon = 10^{-6}$. Критерієм зупинки обох алгоритмів було одночасне виконання трьох умов:

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

Розрахунки проводилися за допомогою відкритої системи комп'ютерної математики Sage [2].

Задача 1. Функція [6, стор. 476], кількість змінних $n = 3$:

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

початкові точки $x_0^1 = (-1.2, 2, 0)^\top$, $\varphi(x_0^1) = 8.40$ та $x_0^2 = (-2, 2, 4)^\top$, $\varphi(x_0^2) = 1610$; точний розв'язок задачі $x^* = (1, 1, 1)^\top$, $\varphi(x^*) = 0$.

Початкова точка	Вибір β_k	Кількість доданків p	$\varphi\left(x^*\right)$	$\left\|\varphi'\left(x^*\right)\right\|$	Кількість ітерацій
x_0^1	умова min	2	2.70e-5	0.0097	148
		3	9.86e-8	0.0042	34
		4	2.52e-5	0.0074	33
		5	2.28e-7	0.0084	56
		7	1.23e-9	0.0004	82
		10	1.23e-9	0.0004	115
x_0^1	умова Вольфе	2	8.16e-6	0.0046	26
		3	6.97e-8	0.0082	20
		4	9.02e-8	0.0010	40
		5	1.87e-6	0.0022	37
		7	4.81e-7	0.0014	69
		10	3.63e-6	0.0042	77
x_0^2	умова min	2	3.11e-5	0.0095	93
		3	2.79e-7	0.0011	35
		4	6.78e-7	0.0017	44
		5	2.96e-6	0.0029	45
		7	4.23e-6	0.0028	61
		10	4.28e-6	0.0028	85
x_0^2	умова Вольфе	2	5.94e-8	0.0030	27
		3	3.49e-8	0.0026	41
		4	6.28e-6	0.0031	32
		5	1.84e-7	0.0049	56
		7	1.75e-5	0.0054	64
		10	9.75e-8	0.0037	131

Задача 2. Функція Пауелла [6, стор. 475], кількість змінних $n = 4$:

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

початкові точки $x_0^1 = (3, -1, 0, 1)^\top$, $\varphi(x_0^1) = 215$ та $x_0^2 = (1, 1, 1, 1)^\top$, $\varphi(x_0^1) = 125$; точний розв'язок задачі $x^* = (0, 0, 0, 0)^\top$, $\varphi(x^*) = 0$.

Початкова точка	Вибір β_k	Кількість доданків p	$\varphi\left(x^*\right)$	$\left\|\varphi'\left(x^*\right)\right\|$	Кількість ітерацій
x_0^1	умова min	2	1.25e-5	0.0053	46
		3	6.07e-7	0.0005	28
		4	2.92e-5	0.0067	34
		5	1.14e-5	0.0024	20
		7	1.11e-6	0.0005	45
		10	6.15e-6	0.0048	75

Початкова точка	Вибір β_k	Кількість доданків p	$\varphi\left(x^{*}\right)$	$\left\|\varphi'\left(x^{*}\right)\right\|$	Кількість ітерацій
x_0^2	умова min	2	8.11e-6	0.0021	25
		3	5.47e-7	0.0007	21
		4	1.99e-6	0.0010	33
		5	4.05e-6	0.0043	32
		7	3.62e-7	0.0065	46
		10	1.27e-5	0.0066	51

Задача 3. Узагальнена функція Розенброка, кількість змінних $n = 8$:

$$\varphi(x) = \sum_{i=1}^{4} \left[(1 - x_{2i-1})^2 + 100 \left(x_{2i} - x_{2i-1}^2 \right)^2 \right],$$

початкова точка $x^0 = (2, 4, 2, 4, \ldots)^\top$, $\varphi(x^0) = 58831$; точний розв'язок задачі $x^* = (1, 1, \ldots)^\top$, $\varphi(x^*) = 0$.

Початкова точка	Вибір β_k	Кількість доданків p	$\varphi\left(x^{*}\right)$	$\left\|\varphi'\left(x^{*}\right)\right\|$	Кількість ітерацій
x^0	умова min	2	2.47e-6	0.0098	152
		3	2.34e-6	0.0056	60
x^0	умова Вольфе	2	2.53e-7	0.0072	98
		3	1.49e-5	0.0052	77

Задача 4. Узагальнена функція Розенброка, кількість змінних $n = 20$:

$$\varphi(x) = \sum_{i=1}^{10} \left[(1 - x_{2i-1})^2 + 100 \left(x_{2i} - x_{2i-1}^2 \right)^2 \right],$$

початкові точки $x_0^1 = (-1.2, 1, -1.2, 1, \ldots)^\top$, $\varphi(x_0^1) = 4598$; $x_0^2 = (0, 0, \ldots)^\top$, $\varphi(x_0^2) = 19$; точний розв'язок задачі $x^* = (1, 1, \ldots)^\top$, $\varphi(x^*) = 0$.

Початкова точка	Вибір β_k	Кількість доданків p	$\varphi\left(x^{*}\right)$	$\left\|\varphi'\left(x^{*}\right)\right\|$	Кількість ітерацій
x_0^1	умова min	2	1.70e-6	0.0098	283
		3	2.07e-6	0.0098	268
x_0^2	умова min	2	4.67e-6	0.0094	105
		3	1.76e-6	0.0085	93

Задача 5. Узагальнена функція Біла (Beale) [1], $n = 100$:

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

початкова точка $x^0 = (1, 0.8 \ldots, 1, 0.8)^\top$.

Початкова точка	Вибір β_k	Кількість доданків p	$\varphi(x^*)$	$\|\varphi'(x^*)\|$	Кількість ітерацій
x^0	умова	2	4.85e-8	0.0020	11
	min	3	4.85e-8	0.0020	8

Задача 6. Узагальнена функція Манєвича (Manevich) [1], $n = 200$:

$$\varphi(x) = \sum_{i=1}^{n} \frac{(1 - x_i)^2}{2^i}$$

початкова точка $x^0 = (0, \ldots, 0)^\top$.

Початкова точка	Вибір β_k	Кількість доданків p	$\varphi(x^*)$	$\|\varphi'(x^*)\|$	Кількість ітерацій
x^0	умова	2	9.78e-4	0.0020	9
	min	3	9.78e-4	0.0020	9

В задачах 3–6 дані експериментів при $p > 3$ не наводяться, тому що вони подібні або гірші за результати при $p = 3$.

Висновки. У роботі розглянуто узагальнений p-кроковий метод для задач безумовної мінімізації, обґрунтована його належність до методів спряжених напрямків. Метою обчислювального експерименту було виявлення «оптимальної» кількості кроків (доданків) p для обчислення нового напрямку спуску.

1. Результати в цілому вказують на переваги методу при $p = 3$ у порівнянні з класичним методом спряжених градієнтів ($p = 2$).

2. Для тестових функцій із кількістю змінних $n \leqslant 20$ бачимо досить вагомі переваги трикрокового методу: алгоритм дає більш точний розв'язок при суттєво меншій кількості кроків.

3. Збільшення кількості доданків p у більшості випадків призводило до погіршення результатів: кількість ітерацій росте, точність зменшується.

4. Для задач великої розмірності ($n = 100$, 200) вказані переваги трикрокового методу спостерігаються в меншій мірі.

1. **Neculai Andrei.** An Unconstrained Optimization Test Functions Collection / Neculai Andrei // Advanced Modeling and Optimization. — Vol. 10, No. 1. — 2008. — pp. 147–161.
2. **SageMath**, the Sage Mathematics Software System (Version 7.5), The Sage Developers, 2017, http://www.sagemath.org.
3. **Карманов В. Г.** Математическое программирование / В. Г. Карманов. — М. : ФИЗМАТЛИТ, 2004. — 264 с.
4. **Страхов Є. М.** Трикроковий метод для задачі багатовимірної оптимізації / Страхов Є. М., Яровий А. Т. // Вісник Київського національного університету. Серія фізико-математичні науки. — 2015. — Вип. 3 — С. 121–126.
5. **Сухарев А. Г.** Курс методов оптимизации / А. Г. Сухарев, А. В. Тимохов, В. В. Федоров. — М. : ФИЗМАТЛИТ, 2005. — 368 с.
6. **Химмельблау Д.** Прикладное нелинейное программирование / Д. Химмельблау. — М. : Мир, 1975. — 536 с.

Страхов Е. М., Яровой А. Т.
Анализ p-шаговых методов минимизации функций многих переменных

Резюме

Рассматривается задача многомерной минимизации непрерывно дифференцируемой функции при отсутствии ограничений. Итерационный алгоритм решения такой задачи называется многошаговым, если для нахождения следующего приближения к точке минимума используются значения функции или ее градиента в двух или более предыдущих точках. Так, алгоритм метода сопряженных градиентов относится к двухшаговым. Описывается обобщенный p-шаговый алгоритм, установлены его свойства в случае квадратичной целевой функции. Показано, что данный метод относится к методам сопряженных направлений. Целью вычислительного эксперимента было сравнение результатов минимизации в зависимости от количества слагаемых (шагов) p и выявления «оптимального» значения для p. Приводятся результаты вычислений для некоторых известных тестовых функций.
Ключевые слова: p-шаговый алгоритм, сопряженные направления, задача безусловной оптимизации .

Strakhov Ye. M., Yaroviy A. T.
The analysis of p-term algorithms for unconstrained optimization

Summary

Consider the multidimensional unconstrained minimization problem in a case of continuously differentiable function. An iterative algorithm for solving such a problem is called a multi-term if in order to find the next approximation to the optimal point we need to compute values of the function or its gradient in two or more previous points. So that, conjugate gradient algorithm is a two-term algorithm. The aim of this paper is to study a generalized p-term method for unconstrained optimization. One substantiates the properties of this algorithm for quadratic functions and proves that it relates to conjugate direction methods. The goal of computational experiment was to compare the results of minimization with different number of terms p and find the “optimal” value for the p. The numerical results for some well-known test functions are given.
Key words: p-term algorithm, conjugate directions, unconstrained optimization.

References

1. Neculai Andrei (2008). An unconstrained optimization test functions collection. *Advanced Modeling and Optimization*, Vol. 10, № 1, P. 147–161.
2. SageMath, the Sage Mathematics Software System (Version 7.5), The Sage Developers, 2017, http://www.sagemath.org.
3. Karmanov, V. G. (2004). *Matematicheskoe programmirovanie [Mathematical programming]*. Moscow: Fizmatlit, 264 p.
4. Strakhov, Ye. M., Yaroviy, A. T. (2015). Trikrokoviy metod dlya zadachi bagatovimirnoyi optimizatsiyi [A three-term method for unconstrained optimization]. *Bulletin of Taras Shevchenko National University of Kyiv. Series Physics & Mathematics*, № 3, P. 121–126.
5. Sukharev, A. G., Timokhov, A. V., Ferorov, V. V. (2005). *Kurs metodov optimizatsii [A course of optimization methods]*. Moscow: Fizmatlit, 368 p.
6. Himmelblau, D. (1975). *Prikladnoe nelineynoe programmirovanie [Applied nonlinear programming]*. Moscow: Mir, 536 p.