ISSN 2078–5097. Вісн. Львів. ун-ту. Сер. прикл. матем. та інф. 2016. Вип. 24. С. 7–13
Visnyk of the Lviv University. Series Appl. Math. and Informatics. Issue 24. P. 7–13

ОБЧИСЛЮВАЛЬНА МАТЕМАТИКА

УДК 519.6

ЗБУРЕНИЙ АНАЛОГ ТРИКРОКОВОГО МЕТОДУ НЬЮТОНА

В. Бартіш, Н. Огородник

Львівський національний університет імені Івана Франка,
вул. Університетська, 1, Львів, 79000, e-mail: ktop@lnu.edu.ua

Розглянуто збурений аналог трикрокового методу Ньютона розв'язування задач безумовної мінімізації. Доведено можливість побудови різних варіантів трикрокових методів на підставі збуреного аналогу. Доведено збіжність трикрокових методів за єдиною схемою, використовуючи матрицю збурення. З'ясовано можливість побудови трикрокових методів розв'язування систем нелінійних рівнянь, використовуючи матрицю збурення.

Ключові слова: безумовна мінімізація, метод Ньютона, матриця збурення, трикрокові методи, ітераційно-різницеві методи, системи нелінійних рівнянь.

1. ВСТУП

На практиці часто виникають задачі, математичні моделі яких призводять до задач оптимізації. Незважаючи на те, що моделювання практичних задач, які містять природні обмеження на ті чи інші параметри, призводить до появи таких обмежень у математичних моделях, багато задач можна якісно розв'язувати методами безумовної оптимізації. В літературі наведено декілька класів таких методів, а також низка модифікацій базових методів [1, 2, 3]. Для різних класів задач досліджуються методи, які найбільш якісно їх розв'язують. Тому важливо запропонувати метод, який для конкретної задачі знайде розв'язок з найменшими обчислювальними затратами та прийнятною точністю.

У зв'язку з тим, що при вимірюванні чи оцінюванні різних величин часто виникають похибки, у числові методи вносять збурення для дослідження їхньої стійкості. Ми узагальнили трикрокові методи мінімізації функцій з використанням матриці збурення. Доведено збіжність збуреного аналогу трикрокового методу в загальному випадку. З'ясовано як з використанням матриці збурення можна будувати різні варіанти трикрокових методів і вплив матриці збурення на збіжність трикрокового методу. Для розв'язування систем нелінійних рівнянь також доведено можливість побудови трикрокових методів з використанням матриці збурення.

2. ФОРМУЛЮВАННЯ ЗАДАЧІ

Розглянемо задачу безумовної мінімізації функції багатьох змінння

$$f(x) \to \min, \tag{1}$$

де $x \in R^n$, $f(x) \in C^2(R^n)$.

Для розв'язування задачі (1) можна використовувати різні методи [1,2,3]. У [4,5,6] розглядається низка трикрокових методів розв'язування задачі (1), які виявили свою ефективність на певних типах задач. Наша мета – узагальнити деякі трикрокові методи, використовуючи матрицю збурення. Використання збурених аналогів на практиці зумовлено похибками обчислень, апроксимацією тощо.

У загальному випадку вигляд трикрокового методу, в якому використовується матриця збурення, такий:

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

$$v_k = x_k - \beta_k h_k, \tag{3}$$

$$x_{k+1} = \arg\min_{\gamma} f\left(u_k + \gamma(v_k - u_k)\right),\ k = 0,1,\ldots. \tag{4}$$

де

$$H(x_k) = f''(x_k) + S(x_k), \tag{5}$$

і

$$\left(h_k, f'(x_k)\right) > 0. \tag{6}$$

Параметр $\beta_k > 0$ має забезпечувати монотонне спадання функції $f(x)$.

3. ОБҐРУНТУВАННЯ ЗБІЖНОСТІ

Теорема. Нехай виконуються умови:

1) $f(x) \in C^2(D),\ D = \{x \in R^n : f(x) \le f(x_0)\}$;

2) $\forall x, y \in D$, $f'''(x)$ задовольняє умову Ліпшиця

$$\|f'''(x) - f'''(y)\| \le L\|x - y\|,$$

де $0 < L < \infty$;

3) $\forall x \in D$, $y \in R^n$

$$m\|y\|^2 \le \left(f''(x)y, y\right) \le M\|y\|^2,$$

де $0 < m \le M < \infty$;

4) матриця $S(x_k)$ задовольняє умову

$$\|S(x_k)\| \le \alpha \|f'(x_k)\|,$$

де $0 < \alpha < \infty$;

5) початкове наближення x_0 вибирають так, що виконуються умови

$$\alpha \frac{M}{m}\|x_0 - x_*\| \le \eta < 1;$$

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

де

$$C = 2M\left(\frac{L + \alpha M}{m^2(1-\eta)}\right)^2.$$

Тоді послідовність $\{x_k\}$, породжена методом (2) – (6), коректно визначена та збігається до розв’язку x_* задачі (1) і справджується оцінка

$$f(x_k) - f(x_*) \le \left(\prod_{i=0}^{k-1} \mu_{k-i}^{2^i}\right) q^{2^k - 1}\left(f(x_0) - f(x_*)\right), \tag{7}$$

де $\mu_k \in (0,1]$, $k = 1,2,\ldots$.

Доведення. З умови 3) теореми випливає сильна опуклість функції $f(x)$, отже, розв’язок задачі (1) x_* існує і єдиний. Аналогічно до результатів [2] можемо записати

$$\left\|\left(f''(x)\right)^{-1}\right\| \le \frac{1}{m}. \tag{8}$$

Нехай відоме деяке наближення $x_k \in D$, $k > 0$ до розв'язку задачі (1), тоді, використовуючи квадратичну частину приросту розкладу функції $f(x)$ в точці x_* і умову 3) теореми, отримаємо оцінку

$$f(u_k) - f(x_*) = \left(f'(x_*), u_k - x_*\right) + \frac{1}{2}\left(f''(x_* + \tau(u_k - x_*))(u_k - x_*), u_k - x_*\right) =$$

$$= \frac{1}{2}\left(f''(x_* + \tau(u_k - x_*))(u_k - x_*), u_k - x_*\right) \le \frac{M}{2}\|u_k - x_*\|^2. \tag{9}$$

Тут і надалі $\tau \in (0,1)$.

Використовуючи (2), одержуємо

$$\begin{aligned} u_k - x_* &= x_k - x_* - \left(f''(x_k) + S(x_k)\right)^{-1} f'(x_k) = \\ &= \left(f''(x_k) + S(x_k)\right)^{-1}\left(\left(f''(x_k) + S(x_k)\right)(x_k - x_*) - f'(x_k)\right) = \\ &= \left(f''(x_k) + S(x_k)\right)^{-1}\left(f''(x_k) + S(x_k) - f''(x_* + \tau(x_k - x_*))\right)(x_k - x_*) = \\ &= \left(f''(x_k) + S(x_k)\right)^{-1}\left(\left(f''(x_k) - f''(x_* + \tau(x_k - x_*))\right)(x_k - x_*) + S(x_k)(x_k - x_*)\right). \end{aligned} \tag{10}$$

Використавши умови теореми, одержимо

$$\left\|\left(f''(x_k) + S(x_k)\right)^{-1}\right\| = \left\|f''(x_k)\left(I + \left(f''(x_k)\right)^{-1} S(x_k)\right)^{-1}\right\| \le$$

$$\le \left\|\left(I + \left(f''(x_k)\right)^{-1} S(x_k)\right)^{-1}\right\| \left\|\left(f''(x_k)\right)^{-1}\right\|.$$

Використовуючи умови 3) і 4) теореми

$$\|S(x_k)\| \le \alpha \|f'(x_k) - f'(x_*)\| = \alpha \|f''(x_* + \tau(x_k - x_*))(x_k - x_*)\| \le \alpha M \|x_k - x_*\|. \tag{11}$$

Згідно з умовою 5) теореми і нерівності (8) матимемо

$$\left\|\left(f''(x_k)\right)^{-1} S(x_k)\right\| \le \left\|\left(f''(x_k)\right)^{-1}\right\| \|S(x_k)\| \le \alpha \frac{M}{m}\|x_k - x_*\| \le \eta < 1.$$

Тоді одержимо оцінку

$$\left\|\left(f''(x_k) + S(x_k)\right)^{-1}\right\| \le \frac{1}{m(1-\eta)}. \tag{12}$$

Використовуючи (11), (12), умову 2) теореми , оцінимо (10)

$$\|u_k - x_*\| = \left\|\left(f''(x_k) + S(x_k)\right)^{-1}\right\| \left(\left\|\left(f''(x_k) - f''(x_* + \tau(x_k - x_*))\right)(x_k - x_*)\right\| +\right.$$

$$\left. + \|S(x_k)(x_k - x_*)\|\right) \le \frac{1}{m(1-\eta)}(L + \alpha M)\|x_k - x_*\|^2.$$

З властивостей опуклих функцій, умови 3) теореми одержуємо таку оцінку:

$$f(x_k) - f(x_*) = \frac{1}{2}\left(f''(x_k + \tau(x_* - x_k))(x_k - x_*), (x_k - x_*)\right) \ge \frac{m}{2}\|x_k - x_*\|^2.$$

$$\|x_k - x_*\|^2 \le \frac{2}{m}\left(f(x_k) - f(x_*)\right).$$

Продовжимо оцінку (9), враховуючи наведені оцінки

$$f(u_k) - f(x_*) \le \frac{M}{2}\left(\frac{L + \alpha M}{m(1-\eta)}\|x_k - x_*\|^2\right)^2 \le \frac{M}{2}\left(\frac{L + \alpha M}{m(1-\eta)}\right)^2 \left(\frac{2}{m}\left(f(x_k) - f(x_*)\right)\right)^2 =$$

$$= 2M\left(\frac{L+\alpha M}{m^2(1-\eta)}\right)^2 \left(f(x_k) - f(x_*)\right)^2 = C\left(f(x_k) - f(x_*)\right)^2 .$$

З нерівності (6) випливає, що для обчислення v_k використовують деякий варіант методу спуску, тому $f(v_k) < f(x_k)$. Тоді з врахуванням (4), матимемо

$$f(x_{k+1}) - f(x_*) \le \mu_{k+1} \min\{f(u_k) - f(x_*), f(v_k) - f(x_*)\} \le \mu_{k+1}\left(f(u_k) - f(x_*)\right),$$

де $\mu_k \in (0,1]$.

Отож,

$$f(x_{k+1}) - f(x_*) \le C\mu_k \left(f(x_k) - f(x_*)\right)^2 . \tag{13}$$

За допомогою методу математичної індукції, враховуючи (13), доведемо, що оцінка (7) виконується для довільного k. При $k = 1$ одержимо

$$f(x_1) - f(x_*) \le C\mu_1\left(f(x_0) - f(x_*)\right)^2 = \mu_1 q^{2^1-1}\left(f(x_0) - f(x_*)\right) =$$

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

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

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

Доведемо, що оцінка правильна для $k+1$

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

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

Отже, послідовність наближень $\{x_k\}$, побудованих за формулами (2) – (6), збігається до точки розв'язку x_*. Теорему доведено.

Отож, при накладанні відповідних умов на матрицю збурення отримаємо квадратичну збіжність збуреного аналога методу Ньютона. Якщо матрицю $S(x_k)$ вибрати нульовою, то отримаємо звичайний трикроковий метод на підставі методу Ньютона.

4. ТРИКРОКОВІ МЕТОДИ РІЗНИЦЕВОГО ТИПУ

На практиці часто виникають проблеми з обчисленням похідних, особливо другого порядку. Це пов'язано з тим, що математичні моделі задач описуються досить складними формулами або значення функцій відомо лише на певній дискретній множині точок. У цьому випадку для обчислення других похідних використовують деякі наближені формули, наприклад, ітераційно-різницеві.

Запис трикрокового алгоритму у вигляді (2)-(6) дає змогу узагальнити трикрокові ітераційно-різницеві методи, які були розглянуті у [5]. Розглянемо такі часткові випадки алгоритму (2)-(6).

Якщо вибрати матрицю $H(x_k)$ так:

$$H(x_k) = f'(x_k, x_k - \lambda f'(x_k)),$$

то отримаємо трикроковий метод на підставі методу Стеффенсена, де $f'(x,y)$ – поділена різниця першого порядку вектор функції $f'(x)$ [1]. Доведемо, що матрицю $H(x_k)$ можна подати у вигляді (5). Використовуючи властивість поділених різниць, отримаємо

$$H(x_k) = f'(x_k, x_k - \lambda f'(x_k)) = f'(x_k, x_k) + f'(x_k, x_k - \lambda f'(x_k)) - f'(x_k, x_k) =$$
$$= f'(x_k, x_k) + f'(x_k, x_k - \lambda f'(x_k), x_k)(-\lambda f'(x_k)).$$

Отже, у нашому випадку $S(x_k) = -\lambda f'(x_k, x_k - \lambda f'(x_k), x_k) f'(x_k)$, $h_k = f'(x_k)$.

Накладемо обмеження на поділені різниці третього порядку, а саме

$$\| f'(x,y,z) \| \le M_3,$$

де $0 < M_3 < \infty$, $x, y \in D$. Тоді матрицю $S(x_k)$ можна оцінити так:

$$\| S(x_k) \| \le \lambda M_3 \| f'(x_k) \|.$$

Якщо прийняти, що $\lambda M_3 \le \alpha$, то умова 5) теореми виконується, а отже, трикроковий алгоритм на підставі методу Стеффенсена має другий порядок збіжності.

Аналогічно до трикрокового методу на підставі методу Стеффенсена можна розглянути випадки трикрокових методів на базі методів лінійної інтерполяції та хорд.

5. РОЗВ'ЯЗУВАННЯ СИСТЕМ НЕЛІНІЙНИХ РІВНЯНЬ

Становить інтерес застосування методу (2)-(6) до розв'язування систем нелінійних рівнянь.

Нехай

$$P(x) = 0, \tag{14}$$

де $P : R^n \to R^m$, $m \ge n$.

Розглянемо функцію

$$f(x) = \frac{1}{2}(P(x), P(x)).$$

Задача $f(x) \to \min$, $x \in R^n$, еквівалентна задачі (14). Розглянемо випадок, коли розв'язок системи (14) існує, тобто існує x_*, для якого $P(x_*) = 0$. У такому випадку, якщо $P(x) \in C^2(R^n)$, то одержуємо

$$f''(x) = P''(x)P(x) + (P'(x))^T P'(x),$$
$$f'(x) = (P'(x))^T P'(x).$$

У цьому випадку алгоритм (2)-(6) для розв'язування задачі (14) набуде вигляду

$$u_k = x_k - \left((P'(x))^T P'(x)\right)^{-1} (P'(x))^T P(x), \tag{15}$$

$$v_k = x_k - \beta_k (P'(x))^T P(x), \tag{16}$$

$$x_{k+1} = \arg\min_{\gamma} f(u_k + \gamma(v_k - u_k)), \; k = 0,1,\dots. \tag{17}$$

Отже, ми вибрали

$$S(x_k) = P''(x)P(x),$$
$$h_k = (P'(x_k))^T P(x_k).$$

У випадку нульової нев'язки $P(x_*)=0$ існує окіл точки x_*, для якого

$$\|P''(x_0)P(x_0)\| < \varepsilon ,$$

де $\varepsilon > 0$ – деяке достатньо мале число, а це означає, що у разі виконання відповідних умов послідовність $\{x_k\}$, яку отримали за (15)-(17), буде збігатися квадратично до x_* розв'язку задачі (14).

6. ВИСНОВКИ

Розглянуто збурений аналог трикрокового методу Ньютона розв'язування задач безумовної мінімізації. Доведено квадратичну збіжність збуреного аналогу у випадку певних обмежень на матрицю збурення. Виявлено можливість побудови різних варіантів трикрокових методів при різному виборі матриці збурення. Також доведено залежність збіжності методу від матриці збурення. Доведено застосування збуреного аналогу трикрокового методу Ньютона для розв'язування систем нелінійних рівнянь.

Розглянуті у статті трикрокові методи було апробовано на низці тестових прикладів [5]. Вони виявили кращі результати в сенсі кількості обчислень, ніж методи, на підставі яких вони побудовані, особливо для функцій типу яру та функцій з виродженою матрицею Гессе. Також трикрокові методи працюють ефективно для задач великих розмірностей.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. *Васильев Ф.П.* Методы оптимизации / Ф.П. Васильев. – М.: Факториал Пресс, 2002. – 824 с.
2. *Пшеничный Б.Н.* Численные методы в экстремальных задачах / Б.Н. Пшеничный, Ю.М. Данилин. – М.: Наука, 1975. – 320 с.
3. *Chong E.K.P.* An introduction to Optimization / Edwin K.P. Chong, Stanislav H. Zak. – New-York: Wiley, 2001. – 477 p.
4. *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.
5. *Огородник Н.П.* Трикрокові ітераційні методи мінімізаційні функцій: автореф. дис. на здобуття. наук. ступеня канд. фіз.-мат. наук: спец. 01.01.07 "Обчислювальна математика" / Н.П. Огородник. – Л., 2012. – 22 с.
6. *Бартіш В.Я.* Трикроковий ітераційно-різницевий метод з порядком збіжності $1+\sqrt{2}$ / В.Я. Бартіш, Н.П. Огородник // Мат. студії. – 2015. – Т. 43, № 2. – С. 220-224.

Стаття: надійшла до редколегії 02.09.2015
доопрацьована 14.10.2015
прийнята до друку 28.10.2015

SOME PERTURBED THREE-STEP NEWTON METHOD

V. Bartish, N. Ogorodnyk

Ivan Franko National University of Lviv,
Universytetska Str., 1, Lviv, 79000, e-mail: ktop@lnu.edu.ua

We consider the perturbed three-step analogue of Newton's method solving unconstrained minimization. The possibility of constructing different options based on the three-step methods perturbed analogue. We prove the convergence of three-step method using the same scheme using matrix of perturbation. The possibility of building a three-step methods for solving systems of nonlinear equations using matrix perturbation.

Key words: unconstrained optimization, Newton method, matrix of perturbation, three-step methods, iterative-difference methods, systems of nonlinear equations.