УДК 519.6

С. М. Шахно

МЕТОД ҐАУССА – НЬЮТОНА – КУРЧАТОВА ДЛЯ РОЗВ'ЯЗУВАННЯ НЕЛІНІЙНИХ ЗАДАЧ ПРО НАЙМЕНШІ КВАДРАТИ

Запропоновано та досліджено ітераційний метод розв'язування нелінійної задачі про найменші квадрати з недиференційовним оператором, у якому замість матриці Якобі використовується сума похідної від диференційовної частини оператора та поділеної різниці зі спеціальним вибором вузлів від недиференційовної частини оператора. Доведено теорему, яка обґрунтовує збіжність і встановлює швидкість збіжності запропонованого методу. Наведено результати чисельного експерименту на тестових задачах з недиференційовним оператором.

Вступ. Розглянемо задачу знаходження наближеного розв'язку нелінійної задачі про найменші квадрати

$$\min_{x\in\mathbb{R}^n} \frac{1}{2} F(x)^\top F(x), \tag{1}$$

де функція відхилу $F : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $m \ge n$, – нелінійна за x, двічі неперервно диференційовна функція і D – відкрита опукла множина $\mathbb{R}^n$.

Нелінійні задачі про найменші квадрати найчастіше виникають при оцінюванні параметрів фізичних процесів за результатами вимірювань, побудови нелінійних регресійних моделей, розв'язування інженерних проблем, розв'язування перевизначених систем рівнянь тощо. Для розв'язування задачі (1) застосовують відомі методи типу Ґаусса – Ньютона [1, 2, 11, 21], які мають у своїх ітераційних формулах похідні від функції F. Однак часто на практиці виникають проблеми з обчисленням похідних. У такому випадку можна використовувати ітераційно-різницеві методи [11, 16, 17, 19, 20], які не потребують обчислення матриці похідних і часто не поступаються методу Ґаусса – Ньютона за швидкістю збіжності та кількістю ітерацій. Але інколи нелінійна функція складається з диференційовної і недиференційовної частин. Тоді виникає нелінійна задача про найменші квадрати

$$\min_{x\in\mathbb{R}^n} \frac{1}{2} (F(x) + G(x))^\top (F(x) + G(x)), \tag{2}$$

де функція відхилу $F + G : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $m \ge n$, – нелінійна за x, F – неперервно диференційовна функція, G – неперервна функція, диференційовності якої, взагалі кажучи, не вимагається, і D – відкрита опукла множина в $\mathbb{R}^n$. Хоча для розв'язування нелінійної задачі (2) можна застосовувати різницеві ітераційні методи [11, 16, 17, 19, 20], проте можна побудувати ітераційні методи, які враховують декомпозицію функції відхилу. У подібному випадку при розв'язуванні нелінійних рівнянь $F(x) + G(x) = 0$ ефективними виявилися методи [6–8, 12, 15, 16, 18], побудовані як комбінації методу Ньютона [1, 2, 11, 13] та ітераційно-різницевих методів хорд і Курчатова [1, 2, 4, 9, 11, 17, 16, 20].

У цій праці пропонуємо метод для розв'язування нелінійної задачі про найменші квадрати з недиференційовним оператором (2), побудований на базі методу Ґаусса – Ньютона [1, 2] і методу типу Курчатова [7, 17, 20], проводимо обґрунтування його локальної збіжності, на тестових задачах показуємо його ефективність порівняно з іншими методами.

Для знаходження розв'язку задачі (2) пропонуємо метод Ґаусса – Ньютона – Курчатова:

$$\begin{aligned} x_{n+1} &= x_n - (A_n^\top A_n)^{-1} A_n^\top (F(x_n) + G(x_n)), \\ A_n &= F'(x_n) + G(2x_n - x_{n-1}, x_{n-1}), \qquad n = 0, 1, \ldots, \end{aligned} \tag{3}$$

де $F'(x_n)$ – матриця Якобі від $F(x)$; $G(2x_n - x_{n-1}, x_{n-1})$ – поділена різниця першого порядку функції $G(x)$ [3] за точками $2x_n - x_{n-1}, x_{n-1}$; x_0, x_{-1} – задані початкові наближення. Метод (3) є комбінацією методу Ґаусса – Ньютона [1, 2] і методу типу Курчатова [7, 17, 20].

Якщо $m = n$, то метод (3) зводиться до методу Ньютона – Курчатова для розв'язування нелінійного рівняння $F(x) + G(x) = 0$ [8, 14, 15, 18]:

$$x_{n+1} = x_n - A_n^{-1}(F(x_n) + G(x_n)),$$

$$A_n = F'(x_n) + G(2x_n - x_{n-1}, x_{n-1}), \quad n = 0,1,\ldots. \tag{4}$$

Поклавши в (3) $A_n = F'(x_n) + G(x_n, x_{n-1})$, отримаємо комбінацію методу Ґаусса – Ньютона [1, 2] і методу типу хорд [16, 20] вигляду [10]

$$x_{n+1} = x_n - A_n^{-1}(F(x_n) + G(x_n)),$$

$$A_n = F'(x_n) + G(x_n, x_{n-1}), \quad n = 0,1,\ldots. \tag{5}$$

1. Обґрунтування збіжності. Достатні умови і швидкість локальної збіжності ітераційного процесу (3) визначені в такій теоремі.

Теорема 1. *Нехай функція* $F + G : \mathbb{R}^n \to \mathbb{R}^m$ *– неперервна в області* $D \subseteq \mathbb{R}^n$*, причому* F *– неперервно диференційовна в цій області,* G *– неперервна функція. Припустимо, що задача* (1) *має розв'язок* x^* *в області* D *та існує обернений оператор*

$$(A_*^\top A_*)^{-1} = \left[(F'(x^*) + G(x^*, x^*))^\top (F'(x^*) + G(x^*, x^*))\right]^{-1}$$

і

$$\left\|(A_*^\top A_*)^{-1}\right\| \le B.$$

Нехай в області D *похідна Фреше* F' *задовольняє умову Ліпшиця*

$$\|F'(x) - F'(y)\| \le L\|x - y\|, \tag{6}$$

а функція G *має поділені різниці першого і другого порядків, які задовольняють умови*

$$\|G(x, y) - G(u, v)\| \le M(\|x - u\| + \|y - v\|), \tag{7}$$

$$\|G(u, x, y) - G(v, x, y)\| \le N\|u - v\|, \tag{8}$$

де $G(\cdot, \cdot, \cdot)$ *– поділена різниця другого порядку* [2, 3]. *Крім того,*

$$\left\|F(x^*) + G(x^*)\right\| \le \eta, \qquad \left\|F'(x^*) + G(x^*, x^*)\right\| \le \alpha, \tag{9}$$

$$B(L + 2M)\eta < 1 \tag{10}$$

і

$$\Omega(x^*, 3r_*) = \left\{x : \left\|x - x^*\right\| < 3r_*\right\} \subseteq D, \tag{11}$$

де r_* *– єдиний додатний корінь полінома* q*, заданого рівністю*

$$q(r) = B\left[(\alpha + (L + 2M)r + 4Nr^2)((L/2 + M)r + 4Nr^2) + (L + 2M + 4Nr)\eta\right] +$$

$$+ B\left[2\alpha + (L + 2M)r + 4Nr^2\right]\left[(L + 2M)r + 4Nr^2\right] - 1. \tag{12}$$

Тоді для $x_0, x_{-1} \in \Omega(x^*, r_*)$ *ітераційний процес* (3) *є коректно визначеним і генерована ним послідовність* $\{x_n\}$, $n = 0,1,\ldots$, *міститься у від-*

критій області $\Omega(x^*, r_*)$ *та збігається до розв'язку* x^*, *причому справджується оцінка*

$$\left\|x_{n+1} - x^*\right\| \le C_1 \left\|x_n - x^*\right\| + C_2 \left\|x_n - x_{n-1}\right\|^2 + C_3 \left\|x_n - x^*\right\|^2 + \\ + C_4 \left\|x_{n-1} - x^*\right\|^2 \left\|x_n - x^*\right\|, \tag{13}$$

де

$$g(r) = B\left[1 - B\left(2\alpha + (L + 2M)r + 4Nr^2\right)\left((L + 2M)r + 4Nr^2\right)\right]^{-1},$$

$$C_1 = g(r_*)(L + 2M)\eta, \qquad C_2 = g(r_*)N\eta,$$

$$C_3 = g(r_*)(L/2 + M)\left(\alpha + (L + 2M)r_* + 4Nr_*^2\right),$$

$$C_4 = g(r_*)N\left(\alpha + (L + 2M)r_* + 4Nr_*^2\right). \tag{14}$$

Д о в е д е н н я. Згідно з теоремою про середнє значення на $[0, r]$ для достатньо великого r і при виконанні (10) поліном q має додатний корінь, позначений через r_*. Але $q'(r) \ge 0$ для $r \ge 0$. Отже, цей корінь єдиний на $[0, r]$.

За припущенням $x_0, x_{-1} \in \Omega(x^*, r_*)$. Тоді маємо

$$\left\|2x_0 - x_{-1} - x^*\right\| \le \left\|x_0 - x^*\right\| + \left\|x_0 - x_{-1}\right\| \le \\ \le \left\|x_0 - x^*\right\| + \left\|x_0 - x^*\right\| + \left\|x_{-1} - x^*\right\| < 3r_*.$$

Отже, $2x_0 - x_{-1} \in \Omega(x^*, 3r_*)$.

Позначимо $A_n = F'(x_n) + G(2x_n - x_{n-1}, x_{n-1})$. Покладемо $n = 0$ і зробимо таку оцінку:

$$\left\|I - (A_*^\top A_*)^{-1} A_0^\top A_0\right\| = \left\|(A_*^\top A_*)^{-1}(A_*^\top A_* - A_0^\top A_0)\right\| = \\ = \left\|(A_*^\top A_*)^{-1}\left(A_*^\top (A_* - A_0) + (A_*^\top - A_0^\top)(A_0 - A_*) + \right.\right. \\ \left.\left. + (A_*^\top - A_0^\top)A_*\right)\right\| \le \left\|(A_*^\top A_*)^{-1}\right\|\left(\left\|A_*^\top\right\|\left\|A_* - A_0\right\| + \right. \\ \left. + \left\|A_*^\top - A_0^\top\right\|\left\|A_0 - A_*\right\| + \left\|A_*^\top - A_0^\top\right\|\left\|A_*\right\|\right) \le \\ \le B\left(\alpha\left\|A_* - A_0\right\| + \left\|A_*^\top - A_0^\top\right\|\left\|A_0 - A_*\right\| + \alpha\left\|A_*^\top - A_0^\top\right\|\right). \tag{15}$$

Використовуючи (8), одержимо

$$\left\|G(2x_0 - x_{-1}, x_{-1}) - G(x_0, x_0)\right\| = \\ = \left\|G(2x_0 - x_{-1}, x_{-1}) - G(x_0, x_{-1}) + G(x_0, x_{-1}) - G(x_0, x_0)\right\| = \\ = \left\|G(2x_0 - x_{-1}, x_{-1}, x_0)(x_0 - x_{-1}) - G(x_0, x_{-1}, x_0)(x_0 - x_{-1})\right\| \le \\ \le N\left\|x_0 - x_{-1}\right\|^2 \tag{16}$$

та

$$\left\|G(2x_0 - x_{-1}, x_{-1}) - G(x_0, x^*)\right\| = \\ = \left\|G(2x_0 - x_{-1}, x_{-1}) - G(x_0, x_0) + G(x_0, x_0) - G(x_0, x^*)\right\| \le \\ \le N\left\|x_0 - x_{-1}\right\|^2 + M\left\|x_0 - x^*\right\|. \tag{17}$$

Врахуємо нерівності (7), (16), (17):

$$\|A_0 - A_*\| = \|(F'(x_0) + G(2x_0 - x_{-1}, x_{-1})) - (F'(x^*) + G(x^*, x^*))\| =$$

$$= \|F'(x_0) - F'(x^*) + G(2x_0 - x_{-1}, x_{-1}) -$$

$$- G(x_0, x^*) + G(x_0, x^*) - G(x^*, x^*)\| \le$$

$$\le L\|x_0 - x^*\| + N\|x_0 - x_{-1}\|^2 + 2M\|x_0 - x^*\| =$$

$$= (L + 2M)\|x_0 - x^*\| + N\|x_0 - x_{-1}\|^2. \qquad (18)$$

Тоді

$$\|A_0\| \le \|A_0\| \| + \|A_0 - A_*\| \le \alpha + (L + 2M)\|x_0 - x^*\| + N\|x_0 - x_{-1}\|^2. \qquad (19)$$

Далі з нерівності (15) та означення r_* (12) отримаємо

$$\|I - (A_*^\top A_*)^{-1} A_0^\top A_0\| \le B\left[2\alpha + (L + 2M)\|x_0 - x^*\| + N\|x_0 - x_{-1}\|^2\right] \times$$

$$\times \left[(L + 2M)\|x_0 - x^*\| + N\|x_0 - x_{-1}\|^2\right] \le$$

$$\le B[2\alpha + (L + 2M)r_* + 4Nr_*^2][(L + 2M)r_* + 4Nr_*^2] =$$

$$= 1 - B\{[\alpha + (L + 2M)r_* + 4Nr_*^2][(L/2 + M)r_* + 4Nr_*^2] +$$

$$+ (L + 2M + 4Nr_*)\eta\} < 1. \qquad (20)$$

За теоремою Банаха про обернений оператор [2] $(A_0^\top A_0)^{-1}$ існує і з (20) маємо

$$\|(A_0^\top A_0)^{-1}\| \le g_0 = B\{1 - B\left[2\alpha + (L + 2M)\|x_0 - x^*\| + N\|x_0 - x_{-1}\|^2\right] \times$$

$$\times \left[(L + 2M)\|x_0 - x^*\| + N\|x_0 - x_{-1}\|^2\right]\}^{-1} \le g(r_*) =$$

$$= B\{1 - B\left[2\alpha + (L + 2M)r_* + 4Nr_*^2\right]\left[(L + 2M)r_* + 4Nr_*^2\right]\}^{-1}.$$

Отже, ітерація x_1 коректно визначена.

Далі покажемо, що $x_1 \in \Omega(x^*, r_*)$. Використовуючи рівність

$$A_*^\top (F(x^*) + G(x^*)) = 0,$$

отримаємо оцінку

$$\|x_1 - x^*\| =$$

$$= \|x_0 - x^* - (A_0^\top A_0)^{-1}(A_0^\top (F(x_0) + G(x_0)) - A_*^\top (F(x^*) + G(x^*))\| \le$$

$$\le \|-(A_0^\top A_0)^{-1}\| \left\| \left[-A_0^\top \left(A_0 - \int_0^1 F'(x^* + t(x_0 - x^*)\right) dt - \right.\right.$$

$$\left.\left. - G(x_0, x^*))(x_0 - x^*) + (A_0^\top - A_*^\top)(F(x^*) + G(x^*)) \right] \right\|.$$

Звідси з урахуванням (17), (19) і нерівностей

$$\left\| A_0 - \int_0^1 F'(x^* + t(x_0 - x^*))dt - G(x_0, x^*) \right\| =$$

$$= \left\| F'(x_0) - \int_0^1 F'(x^* + t(x_0 - x^*))\,dt + \right.$$

$$\left. + G(2x_0 - x_{-1}, x_{-1}) - G(x_0, x^*) \right\| =$$

$$= \left\| \int_0^1 (F'(x_0) - F'(x^* + t(x_0 - x^*))\,dt + \right.$$

$$\left. + G(2x_0 - x_{-1}, x_{-1}) - G(x_0, x^*) \right\| \le$$

$$\le \frac{1}{2} L \|x_0 - x^*\| + M \|x_0 - x^*\| + N \|x_0 - x_{-1}\|^2 \le$$

$$\le \frac{1}{2} L \|x_0 - x^*\| + M \|x_0 - x^*\| + N\left(\|x_0 - x^*\| + \|x_{-1} - x^*\|\right)^2$$

отримаємо

$$\|x_1 - x^*\| \le B\left\{ \left(\alpha + (L + 2M)\|x_0 - x^*\| + N\|x_0 - x_{-1}\|^2\right) \times \right.$$

$$\times \left(\frac{1}{2} L \|x_0 - x^*\| + M \|x_0 - x^*\| + N \|x_0 - x_{-1}\|^2 \right) \|x_0 - x^*\| +$$

$$\left. + \eta\left((L + 2M)\|x_0 - x^*\| + N\|x_0 - x_{-1}\|^2\right) \right\} \times$$

$$\times \left\{ 1 - B\left[2\alpha + (L + 2M)\|x_0 - x^*\| + N\|x_0 - x_{-1}\|^2\right] \times \right.$$

$$\left. \times \left((L + 2M)\|x_0 - x^*\| + N\|x_0 - x_{-1}\|^2\right) \right\}^{-1} =$$

$$= g_0 \left\{ \left(\alpha + (L + 2M)\|x_0 - x^*\| + N\|x_0 - x_{-1}\|^2\right) \times \right.$$

$$\times \left(\frac{1}{2} L \|x_0 - x^*\| + M \|x_0 - x^*\| + N \|x_0 - x_{-1}\|^2 \right) \|x_0 - x^*\| +$$

$$\left. + \eta\left((L + 2M)\|x_0 - x^*\| + N\|x_0 - x_{-1}\|^2\right) \right\} <$$

$$< g(r_*)\left[\left(\alpha + (L + 2M)r_* + 4Nr_*^2\right)\left((L/2 + M)r_* + 4Nr_*^2\right) + \right.$$

$$\left. + (L + 2M + 4Nr_*)\eta\right] r_* = p(r_*) r_* = r_*,$$

де

$$p(r) = g(r)\left[\left(\alpha + (L + 2M)r + 4Nr^2\right)\left((L/2 + M)r + 4Nr^2\right) + \right.$$

$$\left. + (L + 2M + 4Nr)\eta\right].$$

Отже, $x_1 \in \Omega(x^*, r_*)$, і нерівність (13) справджується для $n = 0$.

Припустимо, що $x_n \in \Omega(x^*, r_*)$ для $n = 0,1,\ldots,k$, і виконується оцінка (13) для $n = 0,1,\ldots,k-1$, де $k \geq 1$ − ціле число. Далі доведемо, що $x_{n+1} \in \in \Omega(x^*, r_*)$, і оцінка (13) справджується для $n = k$.

Визначимо

$$\begin{aligned}
\left\| I - (A_*^\top A_*^\top)^{-1} A_k^\top A_k \right\| &= \left\| (A_*^\top A_*)^{-1} (A_*^\top A_* - A_k^\top A_k) \right\| = \\
&= \left\| (A_*^\top A_*)^{-1} (A_*^\top (A_* - A_k) + (A_*^\top - A_k^\top)(A_k - A_*) + \right. \\
&\left. + (A_*^\top - A_k^\top) A_*) \right\| \leq \left\| (A_*^\top A_*)^{-1} \right\| \left(\left\| A_*^\top \right\| \left\| A_* - A_k \right\| + \right. \\
&\left. + \left\| A_*^\top - A_k^\top \right\| \left\| A_k - A_* \right\| + \left\| A_*^\top - A_k^\top \right\| \left\| A_* \right\| \right) \leq \\
&\leq B\left(\alpha \left\| A_* - A_k \right\| + \left\| A_*^\top - A_k^\top \right\| \left\| A_k - A_* \right\| + \alpha \left\| A_*^\top - A_k^\top \right\| \right) \leq \\
&\leq B\left[2\alpha + (L + 2M)\left\| x_k - x^* \right\| + N \left\| x_k - x_{k-1} \right\|^2 \right] \times \\
&\times \left[(L/2 + M)\left\| x_k - x^* \right\| + N \left\| x_k - x_{k-1} \right\|^2 \right] \leq \\
&\leq B\left[2\alpha + (L + 2M) r_* + 4N r_*^2 \right]\left[(L + 2M) r_* + 4N r_*^2 \right] = \\
&= 1 - B\left\{ \left[\alpha + (L + 2M) r_* + 4N r_*^2 \right]\left[(L/2 + M) r_* + 4N r_*^2 \right] + \right. \\
&\left. + (L + 2M + 4N r_*)\eta \right\} < 1 .
\end{aligned}$$

Отже, $(A_k^\top A_k)^{-1}$ існує і

$$\begin{aligned}
\left\| (A_{k+1}^\top A_{k+1})^{-1} \right\| \leq g_k = B\left\{ 1 - B\left[2\alpha + (L+2M)\left\| x_k - x^* \right\| + N \left\| x_k - x_{k-1} \right\|^2 \right] \times \right. \\
\left. \times \left[(L/2 + M)\left\| x_k - x^* \right\| + N \left\| x_k - x_{k-1} \right\|^2 \right] \right\}^{-1} \leq g(r_*) .
\end{aligned}$$

Тому ітерація x_{k+1} коректно визначена і справджується оцінка

$$\begin{aligned}
\left\| x_{k+1} - x^* \right\| &= \left\| x_k - x^* - (A_k^\top A_k)^{-1} (A_k^\top (F(x_k) + G(x_k)) - \right. \\
&\left. - A_*^\top (F(x^*) + G(x^*)) \right\| \leq \left\| -(A_k^\top A_k)^{-1} \right\| \times \\
&\times \left\| \left[-A_k^\top \left(A_k - \int_0^1 F'(x^* + t(x_k - x^*))\, dt - \right. \right. \right. \\
&\left. \left. \left. - G(x_k, x^*) \right)(x_k - x^*) + (A_k^\top - A_*^\top)(F(x^*) + G(x^*)) \right] \right\| \leq \\
&\leq g_k \left\{ \left[\alpha + (L + 2M)\left\| x_k - x^* \right\| + N \left\| x_k - x_{k-1} \right\|^2 \right] \times \right. \\
&\times \left[(L/2 + M)\left\| x_k - x^* \right\| + N \left\| x_k - x_{k-1} \right\|^2 \right] \left\| x_k - x^* \right\| + \\
&\left. + \eta\left((L + 2M)\left\| x_k - x^* \right\| + N \left\| x_k - x_{k-1} \right\|^2 \right) \right\} \leq \\
&\leq g(r_*)\left\{ \left[\alpha + (L + 2M)\left\| x_k - x^* \right\| + N \left\| x_k - x_{k-1} \right\|^2 \right] \times \right.
\end{aligned}$$

$$\times\big[(L/2+M)\big\|x_k-x^*\big\|+N\|x_k-x_{k-1}\|^2\big]\big\|x_k-x^*\big\|+$$

$$+\eta\big((L+2M)\big\|x_k-x^*\big\|+N\|x_k-x_{k-1}\|^2\big)\big\}<p(r_*)r_*=r_*,$$

тобто $x_{k+1}\in\Omega(x^*,r_*)$, і виконується оцінка (13) для $n=k$.

Отже, ітераційний процес (3) коректно визначений, $x_n\in\Omega(x^*,r_*)$ для всіх $n\ge 0$, і справджується оцінка (13) для всіх $n\ge 0$.

Доведемо, що $x_n\to x^*$ при $n\to\infty$. Означимо функції a і b на $[0,r_*]$:

$$a(r)=g(r)((L+2M+3Nr)\eta+h(r)((L/2+M)r+4Nr^2)),$$

$$b(r)=g(r)Nr\eta, \tag{21}$$

де $h(r)=\alpha+(L+2M)r+4Nr^2$.

Згідно з вибором r_* отримаємо

$$a(r_*)\ge 0,\qquad b(r_*)\ge 0,\qquad a(r_*)+b(r_*)=1. \tag{22}$$

Використовуючи оцінку (13), означення сталих C_i, $i=1,2,3,4$, а також функцій a і b, для $n\ge 0$ отримаємо

$$\big\|x_{n+1}-x^*\big\|\le(C_1+C_3r+4C_4r_*^2)\big\|x_n-x^*\big\|+C_2\big(\big\|x_n-x^*\big\|^2+$$

$$+2\big\|x_{n-1}-x^*\big\|\big\|x_n-x^*\big\|+\big\|x_{n-1}-x^*\big\|^2\big)<(C_1+3C_2r_*+$$

$$+C_3r_*+4C_4r_*^2)\big\|x_n-x^*\big\|+C_2r_*\big\|x_{n-1}-x^*\big\|=$$

$$=a(r_*)\big\|x_n-x^*\big\|+b(r_*)\big\|x_{n-1}-x^*\big\|. \tag{23}$$

Аналогічно до [17] доведемо, що за умов (21), (22) послідовність $\{x_n\}$ при $n\to\infty$ збігається до x^*.

Насамперед, для дійсного числа $r_*>0$ і початкових точок $x_0,x_{-1}\in\in\Omega(x^*,r_*)$ існує дійсне число r' таке, що $0<r'<r_*$, $x_0,x_{-1}\in\Omega(x^*,r')$. Тоді всі отримані вище оцінки для послідовності $\{x_n\}$ виконуються і тоді, якщо r_* замінити на r'. Зокрема, з (23) для $n\ge 0$ отримуємо

$$\big\|x_{n+1}-x^*\big\|\le a\big\|x_n-x^*\big\|+b\big\|x_{n-1}-x^*\big\|, \tag{24}$$

де $a=a(r')$, $b=b(r')$.

Очевидно, також матимемо

$$a\ge 0,\qquad b\ge 0,\qquad a+b<a(r_*)\big\|x_n-x^*\big\|+b(r_*)\big\|x_{n-1}-x^*\big\|<1.$$

Послідовності $\{\theta_n\}$, $\{\rho_n\}$ означимо так:

$$\theta_n=\frac{\big\|x_n-x^*\big\|}{r'},\qquad n=-1,0,1,\dots,$$

$$\rho_{-1}=\theta_{-1},\quad \rho_0=\theta_0,\quad \rho_{n+1}=a\rho_n+b\rho_{n-1},\quad n=0,1,2,\dots. \tag{25}$$

Поділивши обидві частини нерівності (24) на r', отримаємо

$$\theta_{n+1}=a\theta_n+b\theta_{n-1},\qquad n=0,1,2,\dots.$$

За означенням послідовності $\{\rho_n\}$ маємо

$$0 \le \theta_n \le \rho_n, \quad n = -1, 0, 1, \dots . \tag{26}$$

Для послідовності $\{\rho_n\}$ відомі явні формули

$$\rho_n = \omega_1 \lambda_1^n + \omega_2 \lambda_2^n, \quad n = -1, 0, 1, \dots , \tag{27}$$

де

$$\lambda_1 = \frac{a - \sqrt{a^2 + 4b}}{2}, \qquad \lambda_2 = \frac{a + \sqrt{a^2 + 4b}}{2}$$

і

$$\omega_1 = \frac{\lambda_2^{-1}\rho_0 - \rho_{-1}}{\lambda_2^{-1} - \lambda_1^{-1}}, \qquad \omega_2 = \frac{\rho_{-1} - \lambda_1^{-1}\rho_0}{\lambda_2^{-1} - \lambda_1^{-1}}.$$

Зауважимо, що

$$0 \le |\lambda_1| \le |\lambda_2| < \frac{a + \sqrt{a^2 + 4(1-a)}}{2} = \frac{a + 2 - a}{2} = 1.$$

Зважаючи на (26) і (27), робимо висновок, що $\{\theta_n\} \to 0$ при $n \to \infty$, а, отже, $x_n \to x^*$ при $n \to \infty$.

Доведення теореми 1 завершено. ♦

Наслідок 1. *У випадку нульового відхилу порядок збіжності ітераційного процесу* (3) *є квадратичним.*

Коли $\eta = 0$, маємо нелінійну задачу найменших квадратів з нульовим відхилом у розв'язку. Тоді сталі $C_1 = 0$ і $C_2 = 0$ і оцінка (13) набуде вигляду

$$\|x_{n+1} - x^*\| \le C_3 \|x_n - x^*\|^2 + C_4 \|x_n - x_{n-1}\|^2 \|x_n - x^*\|. \tag{28}$$

З нерівності (28) видно, що порядок збіжності (3) не є вищим від квадратичного. Отже, існують стала $C_5 \ge 0$ і натуральне N такі, що для всіх $n \ge N$

$$\|x_n - x^*\| \ge C_5 \|x_{n-1} - x^*\|^2.$$

Оскільки

$$\|x_n - x^*\| \le \|x_{n-1} - x^*\|,$$

то

$$\|x_n - x_{n-1}\|^2 \le \left(\|x_n - x^*\| + \|x_{n-1} - x^*\|\right)^2 \le 4\|x_{n-1} - x^*\|^2,$$

і з (28) маємо

$$\|x_{n+1} - x^*\| \le C_3 \|x_n - x^*\|^2 + 4C_4 \|x_{n-1} - x^*\|^2 \|x_n - x^*\| \le$$

$$\le C_3 \|x_n - x^*\|^2 + 4\frac{C_4}{C_5}\|x_n - x^*\|^2 = C_6 \|x_n - x^*\|^2.$$

Отже, порядок збіжності ітераційного процесу (3) є квадратичним.

Як бачимо з оцінок (13) і (14), збіжність ітераційного процесу (3) суттєво залежить від доданків, які містять величини η, α, L, M та N. Для задач з нульовим відхилом у розв'язку ($\eta = 0$) встановлено квадратичну збіжність ітераційного процесу (3). Для задач з малим відхилом у розв'язку (η –«мале») і зі слабкою нелінійністю (α, L, M та N – «малі») збіжність ітераційного процесу є лінійною. У випадку великих відхилів (η – «велике») або для сильно нелінійних задач (α, L, M та N – «великі») ітераційний процес (3) може взагалі не збігатися.

2. Результати числового експерименту. На декількох тестових прикладах порівняємо швидкості збіжності методу Ґаусса – Ньютона – Курчатова (3), методу Ґаусса – Ньютона – хорд (5) і різницевих методу типу хорд [16, 20]

$$x_{n+1} = x_n - \left(A_n^\top A_n\right)^{-1} A_n^\top (F(x_n) + G(x_n)),$$

$$A_n = F(x_n, x_{n-1}) + G(x_n, x_{n-1}), \qquad n = 0,1,\dots, \tag{29}$$

і методу типу Курчатова [17, 20]

$$x_{n+1} = x_n - \left(A_n^\top A_n\right)^{-1} A_n^\top (F(x_n) + G(x_n)),$$

$$A_n = F(2x_n - x_{n-1}, x_{n-1}) + G(2x_n - x_{n-1}, x_{n-1}), \qquad n = 0,1,\dots. \tag{30}$$

Тестування проводили на нелінійних системах з недиференційовним оператором як з нульовим, так і з ненульовим відхилами. Класичні метод Ґаусса – Ньютона і метод Ньютона для їх розв'язування незастосовні. Результати шукали з точністю $\varepsilon = 10^{-8}$. Додаткове наближення x_{-1} вибирали так: $x_{-1} = x_0 - 10^{-4}$. Обчислення здійснювали до виконання умов

$$\left\| x_{n+1} - x_n \right\| \le \varepsilon \qquad \text{і} \qquad \left\| A_n^\top (F(x_n) + G(x_n)) \right\| \le \varepsilon,$$

при цьому $f(x) = \min\limits_{x \in \mathbb{R}^n} \frac{1}{2}(F(x) + G(x))^\top (F(x) + G(x))$.

Приклад 1 [6, 11, 12]:

$$\begin{cases} 3x^2 y + y^2 - 1 + |x - 1| = 0, \\ x^4 + xy^3 - 1 + |y| = 0, \end{cases}$$

$$(x^*, y^*) \approx (0.89465537,\ 0.32782652), \qquad f(x^*) = 0.$$

Приклад 2. $n = 2$, $m = 3$:

$$\begin{cases} 3x^2 y + y^2 - 1 + |x - 1| = 0, \\ x^4 + xy^3 - 1 + |y| = 0, \\ \left| x^2 - y \right| = 0, \end{cases}$$

$$(x^*, y^*) \approx (0.74862800,\ 0.43039151), \qquad f(x^*) \approx 4.0469349 \cdot 10^{-2}.$$

У табл. 1 наведено результати числового експерименту, зокрема, порівнюються досліджувані методи за кількістю ітерацій, виконаних для знаходження розв'язку з заданою точністю. Зауважимо, що методи для розв'язування систем лінійних алгебричних рівнянь, які виникають на кожній ітерації нелінійних алгоритмів (3), (5), (29), (30), наведено в [5, 13].

Таблиця 1. Кількість ітерацій, за які отримано розв'язки тестових задач.

Номер прикладу	Початкове наближення (x_0, y_0)	Метод			
		типу Курчатова (30)	Ґаусса – Ньютона – Курчатова (3)	типу хорд (29)	Ґаусса – Ньютона – хорд (5)
1	(1, 0.1)	6	5	6	5
	(3, 1)	12	9	11	10
	(0.5, 0.5)	12	10	18	10
2	(1, 0.1)	16	14	21	11
	(3, 1)	21	18	25	15
	(0.5, 0.5)	16	14	19	13

Висновки. Отже, на підставі теоретичних досліджень, практичних розрахунків і порівняння отриманих результатів можна стверджувати, що комбіновані диференціально-різницеві методи (3) і (5) збігаються швидше, ніж різницеві метод типу Курчатова (30) і метод типу хорд (29). Оскільки, як доведено, у випадку нульового відхилу метод (3) має квадратичний порядок збіжності та не потребує обчислення похідних від недиференційовної частини оператора, то розглянутий метод (3), як і метод (5), є ефективним методом розв'язування нелінійної задачі про найменші квадрати з недиференційовним оператором.

1. *Дэннис Дж., Шнабель Р.* Численные методы безусловной оптимизации и решения нелинейных уравнений. – Москва: Мир, 1988. – 440 с.
   Те саме: *Dennis J. E.* (*Jr.*)*, Schnabel R. B.* Numerical methods for unconstrained optimization and nonlinear equations. – SIAM, 1996. – 378 p.
2. *Ортега Дж., Рейнболдт В.* Итерационные методы решения нелинейных систем уравнений со многими неизвестными. – Москва: Мир, 1975. – 558 с.
   Те саме: *Ortega J. M., Rheinboldt W. C.* Iterative solution of nonlinear equations in several variables. – New York etc.: Acad. Press, 1970. – xx+572 p.
3. *Ульм С. Ю.* Об обобщенных разделенных разностях. I // Изв. АН ЭССР. Сер. физ.-мат. и техн. наук. – 1967. – **16**, № 1. – С. 13–26.
4. *Шахно С. М.* Про різницевий метод з квадратичною збіжністю для розв'язування нелінійних операторних рівнянь // Мат. студії. – 2006. – **26**, № 1. – С. 105–110.
5. *Шахно С. М.* Чисельні методи лінійної алгебри. – Львів: Видавн. центр ЛНУ ім. І. Франка, 2007. – 245 с.
6. *Шахно С. М., Мельник І. В., Ярмола Г. П.* Аналіз збіжності комбінованого методу для розв'язування нелінійних рівнянь // Мат. методи та фіз.-мех. поля. – 2013. – **56,** № 1. – С. 31–39.
   Те саме: *Shakhno S. M., Mel'nyk I. V., Yarmola H. P.* Analysis of the convergence of a combined method for the solution of nonlinear equations // J. Math. Sci. – 2014. – **201**, No. 1. – P. 32–43. – https://doi.org/10.1007/s10958-014-1971-3.
7. *Шахно С. М., Ярмола Г. П.* Двоточковий метод для розв'язування нелінійних рівнянь з недиференційовним оператором // Мат. студії. – 2011. – **36**, № 2. – С. 213–220.
8. *Шахно С. М., Ярмола Г. П.* Про збіжність методу Ньютона-Курчатова за класичних умов Ліпшиця // Журн. обчисл. та прикл. математики. – 2016. – № 1 (121). – С. 89–97.
9. *Шахно С.* Метод хорд при узагальнених умовах Ліпшиця для розділених різниць першого порядку // Мат. вісн. НТШ. – 2007. – **4**. – С. 296–303.
10. *Шахно С., Шунькін Ю.* Про комбінований метод для розв'язування нелінійних задач про найменші квадрати // Вісн. Львів. ун-ту. Сер. Прикл. математика та інформатика. – 2017. – Вип. 25. – С. 38–48.
11. *Argyros I. K.* Convergence and applications of Newton-type iterations. –NewYork: Springer-Verlag, 2008. – xvi+506 p. – DOI 10.1007/978-0-387-72743-1.
12. *Cătinaş E.* On some iterative methods for solving nonlinear equations // Revue d'Analyse Numér. Théor. de l'Appr. – 1994. – **23**, No. 1. – P. 47–53.
13. *Deuflhard P.* Newton methods for nonlinear problems. Affine invariance and adaptive algorithms. – Berlin: Springer-Verlag, 2004. – xii+424 p.
14. *Hernández-Verón M. A., Rubio M. J.* On the local convergence of Newton–Kurchatov-type method for non-differentiable operators // Appl. Math. Comp. – 2017. – **304**. – P. 1–9. – https://doi.org/10.1016/j.amc.2017.01.010.
15. *Iakymchuk R., Shakhno S., Yarmola H.* Combined Newton-Kurchatov method for solving nonlinear operator equations // Proc. Appl. Math. Mech. – 2016. – **16**, No. 1. – P. 719–720. – https://doi.org/10.1002/pamm.201610348.
16. *Ren H., Argyros I. K.* Local convergence of a secant type method for solving least squares problems // Appl. Math. Comp. – 2010. – **217**, No. 8. – P. 3816–3824. – https://doi.org/10.1016/j.amc.2010.09.040.
17. *Ren H., Argyros I. K., Hilout S.* A derivative free iterative method for solving least squares problems // Numer. Algorithms. – 2011. – **58**, No. 4. – P. 555–571. https://doi.org/10.1007/s11075-011-9470-9.
18. *Shakhno S. M.* Combined Newton-Kurchatov method under the generalized Lipschitz conditions for the derivatives and divided differences // Журн. обчисл. та прикл. математики. – 2015. – № 2 (119). – С. 78–89.

19. *Shakhno S. M., Gnatyshyn O. P.* Iterative-difference methods for solving nonlinear least-squares problem // In: Progress in Industrial Mathematics at ECMI 98 / Ed. by L. Arkeryd, J. Bergh, P. Brenner et al. – Stuttgart: B. G. Teubner, 1999. – P. 287–294.
20. *Shakhno S. M., Gnatyshyn O. P.* On an iterative algorithm of order 1.839... for solving the nonlinear least squares problems // Appl. Math. Comp. –2005. – **161**, No. 1. – P. 253–264. – https://doi.org/10.1016/j.amc.2003.12.025.
21. *Shakhno S.* Some numerical methods for nonlinear least squares problems // In: Symbolic-algebraic Methods and Verification Methods / Eds. G. Alefeld, J. Rohn, S. Rump, T. Yamamoto. – Vienna: Springer, 2001. – P. 235–243. https://link.springer.com/chapter/10.1007/978-3-7091-6280-4_22.

МЕТОД ГАУССА – НЬЮТОНА – КУРЧАТОВА ДЛЯ РЕШЕНИЯ НЕЛИНЕЙНЫХ ЗАДАЧ О НАИМЕНЬШИХ КВАДРАТАХ

Предложен и исследован итерационный метод решения нелинейной задачи о наименьших квадратах с недифференцируемым оператором, в котором вместо матрицы Якоби используется сумма производной от дифференцируемой части оператора и разделенной разности со специальным выбором узлов от недифференцируемой части оператора. Доказана теорема, обосновывающая сходимость, и определена скорость сходимости рассмотренного метода. Приведены результаты численного эксперимента на тестовых задачах с недифференцируемым оператором.

GAUSS – NEWTON – KURCHATOV METHOD FOR SOLVING NONLINEAR LEAST SQUARES PROBLEMS

An iterative method for solving nonlinear least squares problems with nondifferetiable operator, is proposed and examined. In this method instead the Jacobian matrix the sum of derivative of differentiable part of the operator and of divided difference with a special choice of nodes of nondifferentiable part of the operator is used. A theorem justifying convergence of the proposed method is proved and the rate of convergence is established. The results of a numerical experiment on test problems with a nondifferentiable operator are presented.

Львів. нац. ун-т ім. Івана Франка, Львів

Одержано
05.03.17