ISSN 2078–5097. Вісн. Львів. ун-ту. Сер. прикл. матем. та інф. 2017. Вип. 25. С. 38–48
Visnyk of the Lviv University. Series Appl. Math. and Informatics. Issue 25. P. 38–48

УДК 519.21:519.61

ПРО КОМБІНОВАНИЙ МЕТОД ДЛЯ РОЗВ'ЯЗУВАННЯ НЕЛІНІЙНИХ ЗАДАЧ ПРО НАЙМЕНШІ КВАДРАТИ

С. Шахно, Ю. Шунькін

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

Запропоновано та досліджено ітераційний диференціально-різницевий метод розв'язування нелінійної задачі про найменші квадрати, який використовує замість матриці Якобі суму похідної від диференційовної частини оператора і поділену різницю від недиференційовної частини, та метод, який використовує замість якобіана тільки похідну від диференційовної частини оператора. Доведено теореми, які обґрунтовують їхню локальну збіжність, визначено швидкість збіжності розглянутих методів. Наведено результати числового експерименту.

Ключові слова: нелінійна задача про найменші квадрати, диференціально-різницевий метод, поділені різниці, швидкість збіжності, відхил.

1. ВСТУП

Нелінійні задачі про найменші квадрати найчастіше виникають у разі розв'язування перевизначених систем рівнянь, оцінювання параметрів фізичних процесів за результатами вимірювань, побудови нелінійних регресійних моделей, розв'язування інженерних проблем тощо. Ефективними методами розв'язування нелінійної задачі про найменші квадрати є метод Гаусса – Ньютона [1, 2, 6] та його модифікації [1, 5, 6, 8, 9]. Однак часто на практиці маємо проблеми з обчисленням похідних. У такому випадку можна використовувати ітераційно-різницеві методи, які не потребують обчислення матриці похідних і водночас не поступаються методу Гаусса – Ньютона за швидкістю збіжності та близькі до нього за кількістю обчислень. Інколи нелінійна функція складається з диференційовної і недиференційовної частин. Хоч можна застосовувати для розв'язування таких задач різницеві ітераційні методи [5, 6, 8, 9], проте можна побудувати ітераційні методи, які враховують цю специфіку. Зокрема, можна використовувати замість повної матриці Якобі (яка не існує) тільки похідну від диференційовної частини оператора. Однак отримані так методи збігаються досить повільно. Ефективнішими виявляються методи, які замість повної матриці Якобі використовують суму похідної від диференційовної частини оператора та поділеної різниці від недиференційовної частини. Зауважимо, що такий підхід добре зарекомендував себе для розв'язування нелінійних рівнянь [4, 6, 7].

Вперше пропонуємо комбінований метод для розв'язування нелінійної задачі про найменші квадрати, побудований на підставі методів Гаусса-Ньютона та хорд, а також метод, який використовує тільки похідну від диференційовної частини оператора, проводимо обґрунтування їхньої локальної збіжності, на тестових задачах з'ясовуємо їхню ефективність порівняно з методом типу хорд [8, 9].

2. ФОРМУЛЮВАННЯ ЗАДАЧІ

Розглянемо нелінійну задачу про найменші квадрати вигляду

$$\min_{x\in R^p} \frac{1}{2}(F(x)+G(x))^T (F(x)+G(x)), \qquad (1)$$

де функція відхилу $F+G: R^p \to R^m$ ($m \geq p$) – нелінійна по x; F – неперервно диференційовна функція; G – неперервна функція, диференційовності якої загалом не вимагається.

Для знаходження розв'язку задачі (1) пропонуємо таку модифікацію методу Гаусса-Ньютона:

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

Тут $A_n = F'(x_n) + G(x_n, x_{n-1})$, $F'(x_n)$ – якобіан від $F(x)$; $G(x_n, x_{n-1})$ – поділена різниця першого порядку функції $G(x)$ [3] за точками x_n, x_{n-1}; x_0, x_{-1} – задані початкові наближення.

Прийнявши $A_n = F'(x_n)$, ми отримаємо з (2) для розв'язування задачі (1) ітераційний метод типу Гаусса-Ньютона

$$x_{n+1} = x_n - (F'(x_n)^T F'(x_n))^{-1} F'(x_n)^T (F(x_n)+G(x_n)),\ n = 0,1,... \qquad (3)$$

Частковим випадком (1) при $m = p$ є нелінійна система рівнянь

$$F(x)+G(x) = 0. \qquad (4)$$

У цьому разі метод (2) вироджується у комбінований метод Ньютона-хорд [4,7]

$$x_{n+1} = x_n - (F'(x_n) + G(x_n, x_{n-1}))^{-1}(F(x_n)+G(x_n)),\ n = 0,1,..., \qquad (5)$$

а метод (3) – у метод типу Ньютона для розв'язування нелінійних рівнянь [10]

$$x_{n+1} = x_n - (F'(x_n))^{-1}(F(x_n)+G(x_n)),\ n = 0,1,... .$$

3. ОБҐРУНТУВАННЯ ЗБІЖНОСТІ МЕТОДУ (2)

Достатні умови та швидкість локальної збіжності ітераційного процесу (2) визначені в такій теоремі.

Теорема 1. Нехай $F+G: R^p \to R^m$ – неперервна в області $D \subseteq R^p$, 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 $(A_*^T A_*)^{-1} = [(F'(x^*)+G(x^*,x^*))^T (F'(x^*)+G(x^*,x^*))]^{-1}$ і $\| (A_*^T A_*)^{-1} \| \leq B$.

В області D похідна Фреше F' задовольняє умову Ліпшиця

$$\left\| F'(x) - F'(x^*) \right\| \leq L \left\| x - x^* \right\|, \qquad (6)$$

а функція G має поділену різницю першого порядку і

$$\| G(x,y) - G(u,v) \| \leq M(\| x-u \| + \| y-v \|). \qquad (7)$$

Крім того,

$$\| F(x^*) + G(x^*) \| \leq \eta, \qquad \| F'(x^*) + G(x^*,x^*) \| \leq \alpha; \qquad (8)$$

$$B(L+2M)\eta < 1, \qquad (9)$$

і $\Omega = \Omega(x^*, r_*) = \{x : \| x - x^* \| < r_*\} \subseteq D,$ де r_* – єдиний додатний корінь полінома q, заданого рівністю

$$\begin{aligned} q(r) = B[(\alpha + (L+2M)r)(L+2M)r/2 + (L+2M)\eta] + \\ + B[2\alpha + (L+2M)r](L+2M)r - 1. \end{aligned} \qquad (10)$$

Тоді для $x_0, x_{-1} \in \Omega$ ітераційний процес (2) коректно визначений і генерована ним послідовність $\{x_n\}$, $n = 0,1,...,$ міститься у відкритій області Ω та збігається до розв'язку x^*, де справджується оцінка

$$\begin{aligned} \| x_{n+1} - x^* \| \le C_1 \| x_{n-1} - x^* \| + C_2 \| x_n - x^* \| + C_3 \| x_{n-1} - x^* \| \| x_n - x^* \| + \\ + C_4 \| x_{n-1} - x^* \|^2 \| x_n - x^* \| + C_5 \| x_n - x^* \|^2 + \\ + C_6 \| x_{n-1} - x^* \| \| x_n - x^* \|^2 + C_7 \| x_n - x^* \|^3 . \end{aligned} \tag{11}$$

Тут

$$\begin{gathered} g(r) = B[1 - B(2\alpha + (L+2M)r)(L+2M)r]^{-1}; \\ C_1 = g(r_*)M\eta;\ \ C_2 = g(r_*)(L+M)\eta;\ \ C_3 = g(r_*)\alpha M;\ \ C_4 = g(r_*)M^2; \quad (12) \\ C_5 = g(r_*)\alpha \frac{L}{2};\ \ C_6 = g(r_*)\left(\frac{3}{2}LM + M^2\right);\ \ C_7 = g(r_*)\left(\frac{L^2}{2} + \frac{LM}{2}\right). \end{gathered}$$

Доведення. Згідно з теоремою про середнє значення на $[0, r]$ для досить великого r і за виконання (9) поліном q має додатний корінь, позначений через r_*. Але $q'(r) \ge 0$ для $r \ge 0$. Отже, цей корінь єдиний на $[0, r]$.

Позначимо $A_n = F'(x_n) + G(x_n, x_{n-1})$. Приймемо $n = 0$. За припущенням, $x_0, x_{-1} \in \Omega$. Зробимо таку оцінку:

$$\begin{aligned} &\| I - (A_*^T A_*)^{-1} A_0^T A_0 \| = \| (A_*^T A_*)^{-1} (A_*^T A_* - A_0^T A_0) \| = \\ &= \| (A_*^T A_*)^{-1} (A_*^T (A_* - A_0) + (A_*^T - A_0^T)(A_0 - A_*) + (A_*^T - A_0^T) A_*) \| \le \\ &\le \| (A_*^T A_*)^{-1} \| (\| A_*^T \| \| A_* - A_0 \| + \| A_*^T - A_0^T \| \| A_0 - A_* \| + \| A_*^T - A_0^T \| \| A_* \|) \le \\ &\le B(\alpha \| A_* - A_0 \| + \| A_*^T - A_0^T \| \| A_0 - A_* \| + \alpha \| A_*^T - A_0^T \|). \end{aligned} \tag{13}$$

Вважатимемо, що

$$\begin{aligned} &\| A_0 - A_* \| = \| (F'(x_0) + G(x_0, x_{-1})) - (F'(x^*) + G(x^*, x^*)) \| = \\ &= \| F'(x_0) - F'(x_*) + G(x_0, x_{-1}) - G(x^*, x^*) \| \le \\ &\le \| F'(x_0) - F'(x^*) \| + \| G(x_0, x_{-1}) - G(x^*, x^*) \| \le \\ &\le L \| x_0 - x^* \| + M(\| x_0 - x^* \| + \| x_{-1} - x^* \|) \end{aligned} \tag{14}$$

і для евклідової норми $\| A_* - A_0 \| = \| A_*^T - A_0^T \|$. Тоді з нерівності (13), означення r_* (10) отримаємо

$$\begin{aligned} \| I - (A_*^T A_*)^{-1} A_0^T A_0 \| \le B[2\alpha + L \| x_0 - x_* \| + M(\| x_0 - x^* \| + \| x_{-1} - x^* \|)] \times \\ \times [L \| x_0 - x^* \| + M(\| x_0 - x^* \| + \| x_{-1} - x^* \|)] \le \\ \le B[2\alpha + (L+2M)r_*](L+2M)r_* = \\ = 1 - B\{\frac{1}{2}[\alpha + (L+2M)r_*](L+2M)r_* + (L+2M)\eta\} < 1. \end{aligned} \tag{15}$$

Далі за теоремою Банаха про обернений оператор [2] з (13), (14) і (15) матимемо

$$\begin{aligned} \| (A_0^T A_0)^{-1} \| \le g_0 = B\{1 - B[2\alpha + (L \| x_0 - x^* \| + M(\| x_0 - x^* \| + \| x_{-1} - x^* \|)] \times \\ \times (L \| x_0 - x^* \| + M(\| x_0 - x^* \| + \| x_{-1} - x^* \|)\}^{-1} \le \\ \le g(r_*) = B\{1 - B[2\alpha + (L+2M)r_*](L+2M)r_*\}^{-1}. \end{aligned}$$

Отже, ітерація x_1 коректно визначена.

Далі з'ясуємо, що $x_1 \in \Omega(x^*, r_*)$. Насамперед отримаємо оцінку

$$\| x_1 - x^* \| = \| x_0 - x^* - (A_0^T A_0)^{-1}(A_0^T (F(x_0) + G(x_0)) - A_*^T (F(x^*) + G(x^*)) \| \le$$

$$\le \| -(A_0^T A_0)^{-1} \| \, \| [-A_0^T (A_0 - \int_0^1 F'(x_* + t(x_0 - x^*))dt -$$

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

Звідси з урахуванням нерівностей

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

$$= \| F'(x_0) - \int_0^1 F'(x^* + t(x_0 - x^*))dt + G(x_0, x_{-1}) - G(x_0, x^*) \| =$$

$$= \| \int_0^1 (F'(x_0) - F'(x^* + t(x_0 - x^*))dt + G(x_0, x_{-1}) - G(x_0, x^*) \| \le$$

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

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

отримаємо

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

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

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

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

$$\times (L \| x_0 - x^* \| + M(\| x_0 - x^* \| + \| x_{-1} - x^* \|))\} \le$$

$$\le g_0 \{(\alpha + L \| x_0 - x^* \| + M(\| x_0 - x^* \| + \| x_{-1} - x^* \|)) \times$$

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

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

$$< g(r_*)[(\alpha + (L + 2M) r_*)(L + 2M) r_* / 2 + (L + 2M)\eta] r_* =$$

$$= p(r_*) r_* = r_*,$$

де $p(r) = g(r)[(\alpha + (L + 2M) r)(L + 2M) r / 2 + (L + 2M)\eta],$ тобто $x_1 \in \Omega(x^*, r_*)$ і справджується нерівність (11) для $n = 0$.

Припустимо, що $x_n \in \Omega(x^*, r_*)$ для $n = 0, 1, ..., k$ і виконується оцінка (11) для $n = 0, 1, ..., k-1$, де $k \ge 1$ – ціле число. Далі доведемо, що $x_{n+1} \in \Omega$ і справджується оцінка (11) для $n = k$. Визначимо

$$\| I - (A_*^T A_*^T)^{-1} A_k^T A_k \| = \| (A_*^T A_*)^{-1} (A_*^T A_* - A_k^T A_k) \| =$$

$$= \| (A_*^T A_*)^{-1} (A_*^T (A_* - A_k) + (A_*^T - A_k^T)(A_k - A_*) + (A_*^T - A_k^T) A_*) \| \le$$

$$\le \| (A_*^T A_*)^{-1} \| (\| A_*^T \| \| A_* - A_k \| + \| A_*^T - A_k^T \| \| A_k - A_* \| + \| A_*^T - A_k^T \| \| A_* \|) \le$$
$$\le B(\alpha \| A_* - A_k \| + \| A_*^T - A_k^T \| \| A_k - A_* \| + \alpha \| A_*^T - A_k^T \|) \le$$
$$\le B[2\alpha + L \| x_k - x^* \| + M(\| x_k - x^* \| + \| x_{k-1} - x^* \|)] \times$$
$$\times [L \| x_k - x^* \| + M(\| x_k - x^* \| + \| x_{k-1} - x^* \|)] \le$$
$$\le B[2\alpha + (L + 2M) r_*](L + 2M) r_* < 1.$$

Отже, $\left(A_k^T A_k\right)^{-1}$ існує і

$$\| (A_{k+1}^T A_{k+1})^{-1} \| \le g_k = B\{1 - B[2\alpha + (L \| x_k - x^* \| + M(\| x_k - x^* \| + \| x_{k-1} - x^* \|))] \times$$
$$\times (L \| x_k - x^* \| + M(\| x_k - x^* \| + \| x_{k-1} - x^* \|))\}^{-1} \le g(r_*).$$

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

$$\| x_{k+1} - x_* \| = \| x_k - x^* - (A_k^T A_k)^{-1} (A_k^T (F(x_k) + G(x_k)) - A_*^T (F(x^*) + G(x^*)) \| \le$$
$$\le \| -(A_k^T A_k)^{-1} \| \| ([-A_k^T (A_k - \int_0^1 F'(x^* + t(x_k - x^*)) dt -$$
$$- G(x_k, x_*))(x_k - x^*) + (A_k^T - A_*^T)(F(x^*) + G(x^*))] \| \le$$
$$\le g_k \{ \left(\alpha + L \| x_k - x^* \| + M(\| x_k - x^* \| + \| x_{k-1} - x^* \|) \right) \times$$
$$\times \frac{1}{2} \left(L \left\| x_k - x^* \right\| + 2M \left\| x_{k-1} - x^* \right\| \right) \| x_k - x^* \| +$$
$$+ \eta \left(L \| x_k - x^* \| + M(\| x_k - x^* \| + \| x_{k-1} - x^* \|) \right) \} \le$$
$$\le g(r_*) \{ \left(\alpha + L \| x_k - x^* \| + M(\| x_k - x^* \| + \| x_{k-1} - x^* \|) \right) \times$$
$$\times \frac{1}{2} \left(L \left\| x_k - x^* \right\| + 2M \left\| x_{k-1} - x^* \right\| \right) \| x_k - x^* \| +$$
$$+ \eta \left(L \| x_k - x^* \| + M(\| x_k - x^* \| + \| x_{k-1} - x^* \|) \right) \} \le$$
$$< p(r_*) r_* = r_*,$$

що доводить $x_{k+1} \in \Omega(x^*, r_*)$ і оцінку (11) для $n = k$.

Отже, ітераційний процес (2) коректно визначений, $x_n \in \Omega(x^*, r_*)$ для всіх $n \ge 0$ і справджується оцінка (11) для всіх $n \ge 0$. Залишилось довести, що $x_n \to x^*$ при $n \to \infty$.

Визначимо функції a і b на $[0, r_*]$

$$a(r) = g(r)((L + M)\eta + \alpha L r / 2 + L(L + M) r^2 / 2);$$
$$b(r) = g(r)(M\eta + \alpha M r + (2M^2 + 3ML / 2) r^2). \tag{16}$$

Згідно з вибором r_* ми отримаємо

$$a(r_*) \ge 0, \ \ b(r_*) \ge 0, \ \ a(r_*) + b(r_*) = 1. \tag{17}$$

Використовуючи оцінку (11), означення сталих C_i, $i = 1, 2, ..., 7$, і функцій a і b, ми отримаємо для $n \ge 0$

$$\| x_{n+1} - x^* \| \le C_1 \| x_{n-1} - x^* \| + C_2 \| x_n - x^* \| + C_3 r_* \| x_{n-1} - x^* \| +$$
$$+ C_4 r_*^2 \| x_{n-1} - x^* \| + C_5 r_* \| x_n - x^* \| + C_6 r_*^2 \| x_{n-1} - x^* \| + C_7 r_*^2 \| x_n - x^* \| = \tag{18}$$
$$= a(r_*) \| x_n - x^* \| + b(r_*) \| x_{n-1} - x^* \| .$$

Як доведено в [6], за умов (16) – (18) послідовність $\{ x_n \}$ збігається до x^* при $n \to \infty$. Доведення теореми 1 завершено.

Наслідок 1. Порядок збіжності ітераційного процесу (2) у випадку нульового відхилу дорівнює $\frac{1+\sqrt{5}}{2}$.

У випадку $\eta = 0$ маємо нелінійну задачу найменших квадратів з нульовим відхилом у розв'язку. Тоді сталі $C_1 = 0$ і $C_2 = 0$ і оцінка (11) набуде вигляду

$$\| x_{n+1} - x^* \| \le (C_3 + C_4 r_*) \| x_{n-1} - x^* \| \| x_n - x^* \| + (C_5 + C_6 r_* + C_7 r_*) \| x_n - x^* \|^2,$$

тобто ітераційний процес (2) збігається з порядком $\frac{1+\sqrt{5}}{2}$.

Нехай в (1) $G(x) \equiv 0$. Тоді з теореми 1 отримаємо такий наслідок.

Наслідок 2. Порядок збіжності ітераційного процесу (2) у випадку нульового відхилу квадратичний.

Справді, якщо $G(x) \equiv 0$, то в умовах теореми стала $M = 0$. Тоді $C_3 = 0$ і $C_4 = 0$ і оцінка (11) стане такою:

$$\| x_{n+1} - x^* \| \le (C_5 + C_6 r_* + C_7 r_*) \| x_n - x^* \|^2,$$

що свідчить про квадратичну швидкість збіжності процесу (2).

4. ОБҐРУНТУВАННЯ ЗБІЖНОСТІ МЕТОДУ (3)

Достатні умови та швидкість локальної збіжності ітераційного процесу (3) визначені в такій теоремі.

Теорема 2. Нехай $F + G : R^p \to R^m$ – неперервна в області $D \subseteq R^p$, причому F неперервно диференційовна в цій області, G – неперервна функція. Припустимо, що задача (1) має розв'язок x^* в області D, причому $F(x^*) + G(x^*) = 0$, та існує обернений оператор $(A_*^T A_*)^{-1} = [F'(x^*)^T F'(x^*)]^{-1}$ і $\| (A_*^T A_*)^{-1} \| \le B$. В області Ω похідна Фреше F' і функція G задовольняють умови Ліпшиця

$$\left\| F'(x) - F'(x^*) \right\| \le L \left\| x - x^* \right\|, \tag{19}$$

$$\| G(x) - G(x^*) \| \le M \| x - x^* \| . \tag{20}$$

Крім того,

$$\| F'(x^*) \| \le \alpha , \tag{21}$$

$$BM\alpha < 1 \tag{22}$$

і $\Omega = \Omega(x^*, r_*) = \{ x : \| x - x^* \| < r_* \} \subseteq D,$ де r_* – єдиний додатний корінь полінома q, заданого рівністю

$$q(r) = B[(\alpha + Lr)(Lr + 2M) / 2 + L\eta] + B[2\alpha + Lr]Lr - 1. \tag{23}$$

Тоді для $x_0 \in \Omega$ ітераційний процес (3) коректно визначений і генерована ним послідовність $\{x_n\}$, $n = 0,1,...$ міститься у відкритій області Ω та збігається до розв'язку x^*, де справджується оцінка

$$\| x_{n+1} - x^* \| \le C_1 \| x_n - x^* \| + C_2 \| x_n - x^* \|^2 + C_3 \| x_n - x^* \|^3 . \tag{24}$$

Тут

$$g(r) = B[1 - B(2\alpha + Lr)(Lr + 2M)]^{-1};$$
$$C_1 = g(r_*) M\alpha; \quad C_2 = g(r_*)\left(LM + \frac{\alpha L}{2} \right); \quad C_3 = g(r_*) \frac{L^2}{2}. \tag{25}$$

Доведення. Згідно з теоремою про середнє значення на $[0, r]$ для досить великого r і за виконання (22) поліном q має додатний корінь, позначений через r_*, У цьому випадку $q'(r) \ge 0$ для $r \ge 0$. Тому цей корінь єдиний на $[0, r]$.

Позначимо $A_n = F'(x_n)$. Приймемо $n = 0$. За припущенням $x_0, x_{-1} \in \Omega$. Аналогічно, як в (13) теореми 1, отримуємо

$$\| I - (A_*^T A_*)^{-1} A_0^T A_0 \| \le B(\alpha \| A_* - A_0 \| + \| A_*^T - A_0^T \| \| \| A_0 - A_* \| + \alpha \| A_*^T - A_0^T \|). \tag{26}$$

Враховуючи, що

$$\| A_0 - A_* \| = \| F'(x_0) - F'(x^*) \| \| \le L \| x_0 - x^* \| \tag{27}$$

та рівність $\| A_* - A_0 \| = \| A_*^T - A_0^T \|$, правильну принаймні для евклідової норми, з нерівності (26), означення r_* (23) отримаємо

$$\| I - (A_*^T A_*)^{-1} A_0^T A_0 \| \le B[2\alpha + L \| x_0 - x^* \|] L \| x_0 - x^* \| \le$$
$$\le B[2\alpha + Lr_*] Lr_* = 1 - B\{\frac{1}{2}[\alpha + Lr_*] Lr_* + 2M\} < 1. \tag{28}$$

Далі за теоремою Банаха про обернений оператор [2] з (26) – (28) матимемо

$$\| (A_0^T A_0)^{-1} \| \le g_0 = B\{1 - B[2\alpha + L \| x_0 - x^* \|](L \| x_0 - x^* \|\}^{-1} \le$$
$$\le g(r_*) = B\{1 - B[2\alpha + Lr_*] Lr_*\}^{-1}.$$

Отже, ітерація x_1 коректно визначена.

Далі з'ясуємо, що $x_1 \in \Omega(x^*, r_*)$. Насамперед отримаємо оцінку

$$\| x_1 - x^* \| = \| x_0 - x^* - (A_0^T A_0)^{-1} (A_0^T (F(x_0) + G(x_0)) - A_*^T (F(x^*) + G(x^*)) \| \le$$
$$\le \| -(A_0^T A_0)^{-1} \| \| [-A_0^T (A_0 - \int_0^1 F'(x_* + t(x_0 - x^*)) dt (x_0 - x^*) -$$
$$- (G(x_0) - G(x^*))) + (A_0^T - A_*^T)(F(x^*) + G(x^*))] \| .$$

Звідси з урахуванням нерівностей

$$\| A_0 - \int_0^1 F'(x_* + t(x_0 - x^*)) dt (x_0 - x^*) - (G(x_0) - G(x^*)) \| =$$
$$= \| \int_0^1 (F'(x_0) - F'(x^* + t(x_0 - x^*)) dt \, (x_0 - x^*) - (G(x_0) - G(x^*)) \| \le$$
$$\le \left(\frac{1}{2} L \| x_0 - x^* \| + M \right) \| x_0 - x^* \| = \frac{1}{2}\left(L \| x_0 - x^* \| + 2M \right) \| x_0 - x^* \|,$$

$$\| A_0 \| \le \| A_* \| + \| A_0 - A_* \| \le \alpha + L \| x_0 - x^* \|$$

одержимо

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

$$+\eta\left(L \| x_0 - x^* \|\right)\} / \{1 - B[2\alpha + L \| x_0 - x^* \|]L \| x_0 - x^* \|\} \le$$

$$\le g_0\{\left(\alpha + L \| x_0 - x^* \|\right)\frac{1}{2}\left(L\left\|x_0 - x^*\right\| + 2M\right)\| x_0 - x^* \|\} \le$$

$$\le g(r_*)\{\left(\alpha + L \| x_0 - x^* \|\right)\frac{1}{2}\left(L\left\|x_0 - x^*\right\| + 2M\right)\| x_0 - x^* \|\} <$$

$$< g(r_*)[(\alpha + Lr_*)(Lr_* + 2M)/2]r_* = r_*,$$

тобто $x_1 \in \Omega(x^*, r_*)$ і справджується нерівність (24) для $n = 0$.

Припустимо, що $x_n \in \Omega(x^*, r_*)$ для $n = 0,1,...,k$ і виконується оцінка (24) для $n = 0,1,...,k-1$, де $k \ge 1$ – ціле число. Далі доведемо, що $x_{n+1} \in \Omega$ і справджується оцінка (24) для $n = k$.

Визначимо

$$\| I - (A_*^T A_*^T)^{-1} A_k^T A_k \| = \| (A_*^T A_*)^{-1}(A_*^T A_* - A_k^T A_k) \| =$$

$$= \| (A_*^T A_*)^{-1}(A_*^T (A_* - A_k) + (A_*^T - A_k^T)(A_k - A_*) + (A_*^T - A_k^T)A_*) \| \le$$

$$\le \| (A_*^T A_*)^{-1} \| (\| A_*^T \| \| A_* - A_k \| + \| A_*^T - A_k^T \| \| A_k - A_* \| + \| A_*^T - A_k^T \| \| A_* \|) \le$$

$$\le B(\alpha \| A_* - A_k \| + \| A_*^T - A_k^T \| \| A_k - A_* \| + \alpha \| A_*^T - A_k^T \|) \le$$

$$\le B[2\alpha + L \| x_k - x^* \|]L \| x_k - x^* \| \le B[2\alpha + Lr_*]Lr_* < 1.$$

Отже, $\left(A_k^T A_k\right)^{-1}$ існує і

$$\| (A_{k+1}^T A_{k+1})^{-1} \| \le g_k = B\{1 - B[2\alpha + L \| x_k - x^* \|]L \| x_k - x^* \|\}^{-1} \le g(r_*).$$

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

$$\| x_{k+1} - x^* \| = \| x_k - x^* - (A_k^T A_k)^{-1}(A_k^T (F(x_k) + G(x_k)) - A_*^T (F(x^*) + G(x^*)) \| \le$$

$$\le \| -(A_k^T A_k)^{-1} \| \| ([-A_k^T (A_k - \int_0^1 F'(x^* + t(x_k - x^*))dt -$$

$$- G(x_k, x_*))(x_k - x^*) + (A_k^T - A_*^T)(F(x^*) + G(x^*))] \| \le$$

$$\le g_k\{\left(\alpha + L \| x_k - x^* \|\right)\frac{1}{2}\left(L\left\|x_k - x^*\right\| + 2M\right)\| x_k - x^* \|\} \le$$

$$\le g(r_*)\{\left(\alpha + L \| x_k - x^* \|\right)\frac{1}{2}\left(L\left\|x_k - x^*\right\| + 2M\right)\| x_k - x^* \|\} < r_*.$$

що доводить $x_k \in \Omega(x^*, r_*)$ і оцінку (24) для $n = k$.

Отже, ітераційний процес (3) коректно визначений, $x_n \in \Omega(x^*, r_*)$ для всіх $n \ge 0$ і справджується оцінка (24) для всіх $n \ge 0$.

Визначимо функцію a на $[0, r_*]$

$$a(r) = g(r)(M\alpha + (\alpha L/2 + LM)r + L^2 r^2/2). \qquad (29)$$

Використовуючи оцінку (24), означення сталих C_i, $i=1,2,3$ і функції a, ми отримаємо для $n \ge 0$

$$\begin{aligned} \| x_{n+1} - x^* \| &\le C_1 \| x_n - x^* \| + C_2 r_* \| x_n - x^* \| + C_3 r_*^2 \| x_n - x^* \| = \\ &= a(r_*) \| x_n - x^* \| . \end{aligned} \tag{30}$$

Для довільного $r_* > 0$ і початкової точки $x_0 \in \Omega(x^*, r_*)$ існує r' таке, що $0 < r' < r_*$, $x_0 \in \Omega(x^*, r')$. Отже, оцінка (30) справджується, якщо r_* замінити на r'. Зокрема, ми отримаємо з (30) для $n \ge 0$

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

де $a = a(r')$. Очевидно, що $a \ge 0$, $a < a(r_*) = 1$. Тоді

$$\| x_{n+1} - x^* \| \le a \| x_n - x^* \| \le \ldots \le a^{n+1} \| x_0 - x^* \|,$$

де $a^{n+1} \to 0$ при $n \to \infty$. Отже, послідовність $\{ x_n \}$ збігається до x^* при $n \to \infty$, причому зі швидкістю геометричної прогресії. Теорему доведено.

Як бачимо з оцінок (24) і (25), збіжність ітераційного процесу (3) суттєво залежить від доданків, які містять величини α, L та M. Для задач зі слабкою нелінійністю (α, L і M – "малі") збіжність ітераційного процесу лінійна. У разі в сильно нелінійних задач (α, L і/або M – "великі") ітераційний процес (3) може взагалі не збігатися.

5. РЕЗУЛЬТАТИ ЧИСЛОВОГО ЕКСПЕРИМЕНТУ

На кількох тестових прикладах порівняємо швидкості збіжності комбінованого методу (2), методу типу Гаусса-Ньютона (3) та різницевого методу типу хорд для нелінійних задач про найменші квадрати [8, 9]

$$\begin{aligned} x_{n+1} &= x_n - \left(A_n^T A_n \right)^{-1} A_n^T (F(x_n) + G(x_n)), \\ A_n &= F(x_n, x_{n-1}) + G(x_n, x_{n-1}), \quad n = 0,1,\ldots \end{aligned} \tag{31}$$

Тестування проводили на нелінійних системах з недиференційовним оператором з нульовим і ненульовим відхилами. Класичні метод Гаусса-Ньютона та Ньютона для їхнього розв'язування незастосовні. Результати шукали з точністю $\varepsilon = 10^{-8}$. Обчислення виконували до виконання умов

$$\| x_{n+1} - x_n \| \le \varepsilon \ \text{ та } \ \| A_n^T (F(x_n) + G(x_n)) \| \le \varepsilon ,$$

у цьому випадку $f(x) = \min\limits_{x \in R^n} \frac{1}{2} (F(x) + G(x))^T (F(x) + G(x)).$

Приклад 1. [4, 7]

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

$$(x^*, y^*) \approx (0{,}89465537, \quad 0{,}32782652), \quad f(x^*) = 0 .$$

Приклад 2. $n = 2$, $m = 3$;

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

$$(x^*, y^*) \approx (0,74862800, \quad 0,43039151), \quad f(x^*) \approx 4.0469349 \cdot 10^{-2}.$$

У таблиці наведено результати числового експерименту, зокрема, порівнюємо досліджувані методи за кількістю ітерацій, зроблених для знаходження розв'язку з заданою точністю. У прикладі 1 всі методи збіглися до одного розв'язку. Зауважимо, що у прикладі 2 метод типу Гаусса – Ньютона (3) збігся до першої точки $(x^*, y^*) \approx (0,89465537, \quad 0,32782652)$ з відхилом $f(x^*) \approx 1.11666739 \cdot 10^{-1}$, за тієї самої кількості ітерацій. Ці ітерації позначені в таблиці символом *. Інші методи знайшли точку $(x^*, y^*) \approx (0,74862800, \quad 0,43039151)$ з меншим відхилом $f(x^*) \approx 4.0469349 \cdot 10^{-2}$.

Кількість ітерацій, за які отримано розв'язок тестових задач

Номер прикладу	Початкове наближення (x_0, y_0)	Метод		
		типу Гаусса – Ньютона (3)	типу хорд (31)	комбінований метод (2)
1	(1,0)	19	7	7
	(3,1)	22	11	10
	(0,5, 0,5)	21	18	10
2	(1,0)	19*	22	12
	(3,1)	22*	25	15
	(0,5, 0,5)	21*	19	13

6. ВИСНОВКИ

Отже, на підставі теоретичних досліджень, практичних розрахунків і порівняння отриманих результатів можна стверджувати, що комбінований диференціально-різницевий метод (2) збігається швидше, ніж метод типу Гаусса–Ньютона (3) та метод типу хорд (31). Оскільки, як доведено, у випадку нульового відхилу метод має високий порядок збіжності ($(1+\sqrt{5}$) та не потребує обчислення похідних від недиференційовної частини оператора, то розглянутий метод (2) є ефективним методом розв'язування нелінійної задачі про найменші квадрати з недиференційовним оператором.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. *Дэннис Дж.* Численные методы безусловной оптимизации и решения нелинейных уравнений / Дж. Дэннис, Р. Шнабель. – Москва: Мир, 1988. – 440 с.
2. *Ортега Дж.* Итерационные методы решения нелинейных систем уравнений со многими неизвестными / Дж. Ортега, В. Рейнболдт. – Москва: – Мир, 1975. – 558 с.
3. *Ульм С.* Об обобщенных разделенных разностях / С. Ульм // Известия АН ЭССР. Физика. Математика. – 1967. – Т. 16. – С. 13–26.
4. *Шахно С. М.* Аналіз збіжності комбінованого методу для розв'язування нелінійних рівнянь / С. М. Шахно, І. В. Мельник, Г. П. Ярмола // Мат. методи та фіз.-мех. поля. – 2013. – 56, № 1. – С. 11–22.

Те саме: *Shakhno S. M.* Convergence analysis of combined method for solving nonlinear equations / S. M. Shakhno, I. V. Mel'nyk, H. P. Yarmola // J. Math. Sci. – 2016. – 212, № 1. – P. 16–26.

5. *Argyros I. K.* A derivative free iterative method method for solving least squares problems / I. K. Argyros, H. Ren H. // Numerical Algorithms. – 2011. – Vol. 58, № 4. – P. 555–571.
6. *Argyros I. K.* Convergence and Applications of Newton-type Iterations / I. K. Argyros. – Springer-Verlag: New York, 2008. – 506 p.
7. *Cătinas E.* On some iterative methods for solving nonlinear equations / E. Catinas // Revue d'Analyse Numérique et de Theorie de l'Approximation. – 1994. – Vol. 23, № 1. – P. 47–53.
8. *Ren H.* Local convergence of a secant type method for solving least squares problems / H. Ren, I. K. Argyros // AMC (Appl. Math. Comp.). – 2010. – Vol. 217. – P. 3816–3824.
9. *Shakhno S. M.* On an iterative algorithm of order 1.839... for solving the nonlinear least squares problems / S. M. Shakhno, O. P. Gnatyshyn // AMC (Appl. Math. Comp.). – 2005. – Vol. 161. – P. 253–264.
10. *Zabrejko P. P.* The majorant method in the theory of Newton-Kantorovich approximations and the Pták error estimates / P. P. Zabrejko, D. F. Nguen // Numer. Funct. Anal. Optim. – 1987. – Vol. 9. – P. 671–686.

Стаття: надійшла до редколегії 14.12.2016
доопрацьована 19.04.2017
прийнята до друку 14.06.2017

ONE COMBINED METHOD FOR SOLVING NONLINEAR LEAST SQUARES PROBLEMS

S. Shakhno, Yu. Shunkin

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

One iterative differential-difference method for solving nonlinear least squares problems, which uses instead Jacobian the sum of derivative of differentiable parts of operator and divided difference of nondifferentiable parts, and a method, which uses instead Jacobian the derivative of differentiable parts, are proposed and studied. We have proved theorems where convergence of the proposed methods are justified and rate of convergence are established. Numerical results are presented.

Key words: nonlinear least squares problem, differential-difference method, divided differences, rate of convergence, residual.