УДК 519.61

ДОСЛІДЖЕННЯ РІЗНИЦЕВИХ АНАЛОГІВ МЕТОДУ ГАУСА-НЬЮТОНА

С.М. Шахно

S.M. Shachno. *Investigation of difference analogous of Gauss-Newton methods*, Matematychni Studii, **10**(1998) 219–222.

It is proposed the difference analogous of the Gauss-Newton method which don't demand the evaluation of derivatives. It is received the conditions and estimates of the rate of convergence of the proposed methods.

С.М. Шахно. *Исследование разничных аналогов метода Гаусса-Ньютона* // Математичні Студії. – 1998. – Т.10, № 2. – С.219–222.

Для решения нелинейной задачи о наименьших квадратах предложены разностные аналоги метода Гаусса-Ньютона, не требующие вычисления производных. Исследованы условия и порядок сходимости предложенных методов.

1. Постановка задачі. Найбільш поширеними і ефективними методами для розв'язування нелінійної задачі про найменші квадрати: Знайти

$$\min_{x\in R^n} \frac{1}{2}F(x)^T F(x) = \frac{1}{2}\sum_{i=1}^{m} {F_i}^2(x), \qquad m \geq n, \tag{1}$$

де $F\colon R^n \to R^m$ нелінійна по x функція, є метод Гауса-Ньютона та його модифікації [1,2]. Однак всі вони потребують обчислення оператора $F'(x)$, що небажано, наприклад, при складному аналітичному виразі $F'(x)$. Розглянуті в статті різницеві методи мають порядок збіжності не нижчий, ніж в методі Гауса-Ньютона, однак вони не потребують обчислення $F'(x)$, і їх можна застосовувати навіть тоді, коли відсутній аналітичний вираз функції, а заданий лише машинний алгоритм обчислення функції (тоді метод Гауса-Ньютона просто незастосовний). Тому для розв'язування (1) пропонуємо ітераційно-різницевий метод

$$x_{n+1} = x_n - [H_n^T H_n]^{-1} H_n^T F(x_n), \qquad n = 1, 2, \dots, \tag{2}$$

де лінійна операція $H_n = H(x_{n-1}, x_n)$ залежить тільки від елементів x_{n-1}, x_n, причому

$$\|F'(x_n) - H(x_{n-1}, x_n)\| = \|F'_n - H_n\| \leq \tfrac{1}{6}\|x_n - x_{n-1}\|^2 \|F'''(x)\|_R, \tag{3}$$

а x_* — єдиний розв'язок задачі (1) в деякій області $R = \{x : \|x - x_*\| \leq \eta\}$, $\eta = \text{const} > 0$. Метод (2) простий в реалізації, володіє такою ж швидкістю

1991 *Mathematics Subject Classification.* 65H10, 65J15.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

збіжності, як і метод Гауса-Ньютона, проте, на відміну від останнього, не містить оператора $F'(x)$. У випадку $m = n$ метод (2) перетворюється в метод лінійної інтерполяції для розв'язування нелінійного функціонального рівняння $F(x) = 0$, дослідженим в [3].

2. Дослідження збіжності методу (2). Збіжність і оцінка швидкості збіжності процесу (2) встановлюється наступною теоремою.

Теорема 1. *Нехай: 1) Для початкового наближення x_0 існує обернена операція $\Gamma_0 = [F'(x_0)^T F'(x_0)]^{-1}$; 2) $\|F'(x_0)\| \le \alpha$; 3) $\|(F'(x) - F'(x_*))^T F(x_*)\| \le \sigma\|x - x_*\|$; $\|F''(x)\| \le M$; $\|F'''(x)\| \le N$; в області R; 4) $\|F(x_*) \le \eta$.*

Тоді для x_0, x_1, достатньо близьких до розв'язку задачі (1), ітераційний процес (2) збігається до розв'язку зі швидкістю, яка характеризується нерівністю

$$\|x_{n+1} - x_*\| \le C_1\|x_n - x_*\| + C_2\|x_n - x_*\|^2, \tag{4}$$

де C_1, C_2 — деякі обмежені додатні постійні.

Доведення. Доведемо спочатку, аналогічно [3], що для x_{n-1}, x_n, достатньо близьких до x_0, існує обернена операція $A_n^{-1} = [H_n^T H_n]^{-1}$. Позначимо $\Gamma_0 = [F'(x_0)^T F'(x_0)]^{-1} = [{F'_0}^T F'_0]^{-1}$. Тоді для x_{n-1}, x_n, достатньо близьких до x_0 , отримаємо

$$\|\Gamma_0({F'_0}^T F'_0 - H_n^T H_n)\| \le \|\Gamma_0\|\|{F'_0}^T F'_0 - H_n^T H_n)\| =$$
$$= \|\Gamma_0\|\|{F'_0}^T (F'_0 - H_n) + ({F'_0}^T - H_n^T)(H_n - F'_0 + F'_0)\| \le \|\Gamma_0\|[(M\|x_n - x_0\| +$$
$$+ \tfrac{1}{6}\|x_{n-1} - x_n\|^2)(2\alpha + M\|x_n - x_0\| + \tfrac{N}{6}\|x_{n-1} - x_n\|^2)] = h_n < 1, \tag{5}$$

де використано оцінку $\|F'_0 - H_n\| = \|F'_0 - F'_n + F'_n - H_n\| \le M\|x_n - x_0\| + \frac{N}{6}\|x_{n-1} - x_n\|^2$.

Згідно теореми Банаха випливає існування оберненої операції $\Gamma_n = [I - ({F'_0}^T F'_0)^{-1}({F'_0}^T F'_0 - H_n^T H_n)]^{-1}$ і $\|\Gamma_n\| \le \frac{1}{1-h_n}$. Таким чином, операція $\Gamma_n \Gamma_0 = (H_n^T H_n)^{-1}$ існує і

$$\|A_n^{-1}\| = \|(H_n^T H_n)^{-1}\| \le \|\Gamma_0\|/(1 - h_n). \tag{6}$$

Тепер можна встановити збіжність досліджуваного методу і отримати оцінку швидкості збіжності. Шляхом тотожних перетворень отримаємо $x_{n+1} - x_* = x_n - x_* - A_n^{-1}[H_n^T F(x_n) - F(x_*, x_*)^T F(x_*)] = x_n - x_* - A_n^{-1}\{H_n^T F(x_n, x_*)(x_n - x_*) + [H_n^T - F(x_n, x_n)^T + F(x_n, x_n)^T - F(x_*, x_*)^T]F(x_*)\}$, де $F(u, v)$ — перша поділена різниця оператора $F(x)$, причому $F(u, u) = F'(u)$ [4]. Тоді

$$\|x_{n+1} - x_*\| \le \|x_n - x_*\|\|I - A_n^{-1} H_n^T F(x_n, x_*) + \|A_n^{-1}\|\|H_n^T -$$
$$- F(x_n, x_n)^T\|\|F(x_*)\| + \|A_n^{-1}\|\|F(x_n, x_n)^T - F(x_*, x_*)^T\|\|F(x_*)\|. \tag{7}$$

Оцінимо $\|I - A_n^{-1} H_n^T F(x_n, x_*)\| \le \|A_n^{-1}\|\|H_n^T H_n - H_n^T F(x_n, x_*)\| \le$

$$\le \|A_n^{-1}\|\|H_n^T(\|H_n - F(x_n, x_n)\| + \|F(x_n, x_n) - F(x_n, x_*)\|). \tag{8}$$

Тепер із (7) з врахуванням (8),(3) та умов теореми, отримаємо

$$\|x_{n+1} - x_*\| \le \|x_n - x_*\|\|A_n^{-1}\|[\|H_n^T\|\|(H_n - F'_n)\| + \|F(x_n, x_n) -$$
$$- F(x_n, x_*)\| + \|A_n^{-1}\|[\|H_n^T - {F'_n}^T\|\|F(x_*)\| + \|({F'_n}^T - {F'_*}^T)F(x_*)\|] \le$$
$$\le \|x_n - x_*\|\|A_n^{-1}\|[(\tfrac{N}{6}\|x_n - x_{n-1}\|^2 + \alpha)(\tfrac{N}{6}\|x_n - x_{n-1}\|^2 + \tfrac{1}{2}M\|x_n - x_*\|)] +$$
$$+ \|A_n^{-1}\|[\tfrac{N}{6}\|x_n - x_{n-1}\|^2\eta + \sigma\|x_n - x_*\|]. \tag{9}$$

З нерівності (9) випливає, що швидкість досліджуваного методу, взагалі кажучи, не перевищить квадратичної. Тому

$$\|x_n - x_{n-1}\|^2 \leq C\|x_n - x_*\| \tag{10}$$

Враховуючи (10), з (9) отримаємо

$$\|x_{n+1} - x_*\| \leq \|x_n - x_*\|\|A_n^{-1}\|[(\tfrac{N}{6}C\|x_n - x_*\| + \alpha)(\tfrac{N}{6}C + \tfrac{M}{2})\|x_n - x_*\| + \tfrac{N}{6}\eta + \sigma], \tag{11}$$

звідки після відповідних перепозначень отримаємо нерівність (4). Теорему доведено.

З оцінки (11) випливає, що у випадку нульової нев'язки в розв'язку (тоді $\eta = 0$, $\sigma = 0$) та лінійної функції $F(x)$ (тоді $N = 0$, $\sigma = 0$) ітераційний процес (2) збігається з квадратичною швидкістю. Проте, очевидно, для задач з ненульовою нев'язкою та нелінійних задач збіжність буде лише лінійною, а для сильно нелінійних задач та задач з великою нев'язкою в розв'язку метод (2) може взагалі розбігатися. Таким чином, різницевий метод володіє як перевагами, так і недоліками методу Гауса-Ньютона.

Легко переконатися, що умову (3) задовольняє операція $H(x_{n-1}, x_n) = F(2x_n - x_{n-1}, x_{n-1})$, яка і використовується на практиці. Тоді процес (2)–(3) прийме вигляд

$$x_{n+1} = x_n - \frac{F(2x_n - x_{n-1}, x_{n-1})^T F(x_n)}{F(2x_n - x_{n-1}, x_{n-1})^T F(2x_n - x_{n-1}, x_{n-1})}, \tag{12}$$

$n = 1, 2, \ldots$ Тут $F(u, v)$ — перша поділена різниця для функції $F(x)$. Цей процес володіє такою ж швидкістю збіжності, як і в методі Гауса-Ньютона, однак, на противагу останньому, не вимагає обчислення похідних.

3. Інші різницеві аналоги методу Гауса-Ньютона. Точніші оцінки отримаємо у випадку нульової нев'язки в розв'язку для функції $F(x)$, тобто якщо $F(x_*) = 0$. Нехай $\varphi\colon R^n \to R^n$ — деякий допоміжний оператор, для якого $x_* = \varphi(x_*)$. Оператор $\varphi(x)$ можна отримати з довільних n рівнянь $F_i(x) = 0$, $i = 1, 2, \ldots, m$. Наприклад, можна взяти $\varphi_i(x) = x_i - F_i(x)$, $i = 1, \ldots, n$. Тоді різницевий аналог методу Гауса-Ньютона доцільно записати у вигляді

$$x_{n+1} = x_n - [F(x_n, \overline{x}_n)^T F(x_n, \overline{x}_n)]^{-1} F(x_n, \overline{x}_n)^T F(x_n), \quad n = 0, 1, \ldots. \tag{13}$$

Ітераційна формула (13) не вимагає обчислення похідних і може бути використана навіть тоді, коли відсутній аналітичний вираз для $F(x)$, а відомий лише машинний алгоритм обчислення оператора $F(x)$. Зауважимо, що в методі типу Стеффенсена для безумовної мінімізації, розглянутому в [5], використані поділені різниці для похідної $f'(x)$, щоб уникнути обчислення другої похідної $f''(x)$.

Розглянемо деякі варіанти $\overline{x}_n$.

а). Виберемо $\overline{x}_n = \varphi(x_n)$. Через $F(x, y, z)$ позначимо другу поділену різницю оператора $F(x)$. Тоді для ітераційного процесу (13) справедлива теорема.

Теорема 2. *Нехай $F\colon R^n \to R^m$ $(m \geq n)$ і $f(x) = \frac{1}{2}F(x)^T F(x)$ двічі неперервно диференційовна на відкритій опуклій множині $D \subset R^n$. Припустимо, що $\|F(x, y)\| \leq \beta$, $\|\varphi'(x) \leq \alpha < 1$, $\|F(x, y, x_*)\| \leq \frac{1}{2}L$ для всіх $x, y \in D$, а також що існують $x_* \in D$ і $\lambda \geq 0$ такі, що $F(x_*) = 0$, а λ — найменше власне число для $F'(x_*)^T F'(x_*)$. Тоді існує r_0 таке, що для всіх $x_0 \in \Omega(x_*, r_0)$ послідовність,* *породжена методом (13), коректно визначена, збігається до розв'язку x_* і задовольняє нерівності*

$$\|x_n - x_*\| \leq \frac{\beta}{2\lambda} L\alpha \|x_{n-1} - x_*\|^2, \quad \|x_n - x_*\| \leq (h_0 r_0)^{2^n - 1} r_0, \quad n = 1, 2, \ldots, \tag{14}$$

де $h_0 = \frac{\beta}{2\lambda} L\alpha < \frac{1}{r_0}$, $r_0 = \|x_0 - x_*\|$, $\Omega(x_*, r_0) = \{x : \|x - x_*\| \leq r_0\}$.

Доведення теореми 2 здійснюється за схемою доведення теореми 1 [6].

б). Вибравши в (13) $\overline{x}_n = (1-\mu)x_n + \mu\varphi(x_n)$, де $0 \leq \mu \leq 1$ — числовий параметр, ми отримаємо різницевий аналог модифікації методу Гауса-Ньютона, яка досліджена в [6–8]. Для цього різницевого методу справедлива теорема, аналогічна теоремі 2. Зауважимо, що при $\mu = 0$ $\overline{x}_n = x_n$, і тоді (13) перетвориться у метод Гауса-Ньютона.

в). Виберемо $\overline{x}_n = x_{n-1}$. Тоді з (13) отримаємо ітераційний процес

$$x_{n+1} = x_n - [F(x_n, x_{n-1})^T F(x_n, x_{n-1})]^{-1} F(x_n, x_{n-1})^T F(x_n), \quad n = 1, 2, \ldots, \tag{15}$$

де x_0, x_1 — деякі початкові наближення. Ітераційний процес (15) придатний для розв'язування нелінійної задачі про найменші квадрати (1). Його можна розглядати як застосування методу січних [1] до задачі (1). У випадку нульової нев'язки $F(x_*) = 0$ ітераційний процес збігається до розв'язку x_* зі швидкістю, яка характеризується нерівністю

$$\|x_{n+1} - x_*\| \leq C\|x_n - x_*\|\|x_{n-1} - x_*\|, \; n = 1, 2, \ldots, \tag{16}$$

де C — обмежена додатна постійна. З нерівності (16) випливає, що порядок збіжності процесу (15) дорівнює $\frac{1+\sqrt{5}}{2} \approx 1,618$.

Проведені нами чисельні розрахунки показали, що за кількістю ітерацій і за загальною кількістю обчислень (часом рахунку) результати за різницевим аналогом з квадратичною збіжністю не поступаються відповідним результатам за класичним методом Гауса-Ньютона. Враховуючи, що запропоновані методи прості в реалізації і потребують лише обчислення значень функції, а також володіють швидкою збіжністю, то вони можуть широко використовуватися на практиці.

ЛІТЕРАТУРА

1. Дэннис Дж., мл., Шнабель Р. Численные методы безусловной оптимизации и решения нелинейных уравнений. – М.: Мир, 1988. – 440с.
2. Ортега Дж., Рейнболдт В. Итерационные методы решения нелинейных систем уравнений со многими неизвестными. – М.: Мир, 1975. – 558с.
3. Курчатов В.А. *Об одном методе линейной интерполяции решения функциональных уравнений* Докл. АН СССР. Сер. Математика. Физика. – 1971. – Т.198, №3. – С.524–526.
4. Ульм С. *Об обобщенных разделенных разностях. I, II* Известия АН ЭССР. Физика. Математика. – 1967. – Т.16. – С.13–26, С.146–156.
5. Bartish M.Ya., Shachno S.M. *On the Iterative Steffensen Type Methods* Zeitschrift für Angewandte Mathematik und Mechanik, Berlin. – 1996. – V.76, S1. – S.351–352.
6. Бартіш М.Я., Шахно С.М. *Деякі методи розв'язування нелінійної задачі найменших квадратів* Вісн. Львів. ун-ту. Сер. мех.-мат. – 1993. – Вип.39. – С.3–7.
7. Shachno S.M. *Numerical methods for solving nonlinear least squares problems* 9th Conference of the European Consortium for Mathematics in Industry. Technical University of Denmark Lyngby /Copenhagen, Denmark. – June 25-29, 1996. – Book of abstracts. – P.543–545.
8. Бартиш М.Я., Шахно С.М. *Исследование параметрических итерационных процессов для решения нелинейных уравнений* Проблемы управления и информатики. – 1997. – №2. – С.22–30.

Львівський ун-т, ф-т прикладної математики та інформатики

Надійшло 20.01.1998