УДК 519.6

М.Я.Бартіш, О.В.Николайчук, А.І.Чипурко

Про деякі різницеві модифікації методу Гаусса-Ньютона для нелінійної задачі найменших квадратів з малою нормою нев'язки

В даній статті ми продовжуємо дослідження різницевих схем комбінованих методів розв'язання нелінійної задачі про найменші квадрати

знайти $$\min \frac{1}{2} F(x)^T F(x) \qquad (1)$$

де $F(x): R^n \to R^m$ – нелінійна по x, $m \geq n$. Зокрема, в [1] було запропоновано загальну схему на базі різницевого аналога формули Гаусса-Ньютона

$$x_{k+1} = x_k - \left[F(x_k, \theta_k)^T F(x_k, \theta_k)\right]^{-1} F(x_k, \theta_k)^T F(x_k), \qquad (2)$$

$$\theta_k = \alpha x_k + (1-\alpha)\Phi(x_k),\ 0 \leq \alpha \leq 1, \qquad (3)$$

$$\mathrm{k} = 0, 1, \ldots,$$

де $F(x, y)$ – перша поділена різниця [4], а оператор $\Phi(x)$ породжує ітераційну формулу порядку $1 \leq \tau \leq 2$, тобто

$$\left\|\Phi(x) - x^*\right\| \leq L\left\|x - x^*\right\|^\tau,\ L = const > 0. \qquad (4)$$

Доведено [1], що для задач з нульовою нормою нев'язки (2),(3) має кращі швидкісні властивості в порівнянні з класичними методами, а саме збігається з порядком $\tau + 1$ при $\alpha = 0$ та порядком 2 при $\alpha \neq 0$.

В даній замітці ми розглянемо випадок наявності малої норми нев'язки. Наступна теорема дає умови Коші збіжності (2),(3).

Теорема 1. Нехай виконуються умови:

1) $\Phi(x)$ – неперервний оператор, який породжує ітераційну формулу порядку τ, $1 \leq \tau \leq 2$;

2) $F(x)$ – неперервно-диференційовна в області $D \subseteq R^n$;

3) $\forall x, y \in D\ \left\|F(x, y)\right\| \leq C$, під $F(x, x)$ розуміємо $F'(x)$ [4];

4) $\forall x, y, z \in D\ \left\|F(x, y, z)\right\| \leq M$, де $F(x, y, z)$ – друга поділена різниця [4];

5) $\forall x, y \in D$ існує обернений оператор $[F(x,y)^T F(x,y)]^{-1}$, причому $\left\|[F(x,y)^T F(x,y)]^{-1}\right\| \le B$;

6) в області D існує точка x^*, яка є розв'язком задачі (1), причому $\left\|F(x^*)\right\| \le \varepsilon$, де ε – достатньо мале число;

7) початкове наближення x_0 вибрано так, що $x_0 \in \Omega_0 = \left\{x : \left\|x - x^*\right\| < \eta_0\right\} \subset D$, $\eta_0 = const > 0$, $L\eta_0^{\tau-1} \le 1$;

8) $h_0 = K_1\eta_0 < 1$, де $K_1 = BCM + 2BM\gamma\,\eta_0^{\tau-1}$.

Тоді послідовність наближень, породжена (2),(3), коректно означена, збігається до розв'язку задачі (1) і справедливі наступні оцінки

$$\left\|x_k - x^*\right\| \le h_0^{2^k - 1}\eta_0,\ k = 1,2,\ldots,m+1, \tag{5}$$

$$\left\|x_{m+p} - x^*\right\| \le h_0^{(p+1)2^m - 1}\eta_0,\ p = 2,3,\ldots, \tag{6}$$

Тут m визначаємо з умови

$$\gamma\left\|x_{m+1} - x^*\right\|^\tau < \varepsilon \le \gamma\left\|x_m - x^*\right\|^\tau,$$

де x_m, x_{m+1} – два послідовні наближення, отримані по (2),(3), $\gamma = const > 0$.

Якщо $\alpha = 0$ і виконується умова

8') $h_0^\tau = K_\tau\eta_0^\tau < 1$, де $K_\tau = BCML + 2BM\gamma$,

то будемо мати

$$\left\|x_k - x^*\right\| \le h_0^{(\tau+1)^k - 1}\eta_0,\ k = 1,2,\ldots,m+1 \tag{7}$$

$$\left\|x_{m+p} - x^*\right\| \le h_0^{(p\tau+1)(\tau+1)^m - 1}\eta_0,\ p = 2,3,\ldots \tag{8}$$

Доведення **.** Отримаємо спочатку загальну оцінку для деякого $k+1$ наближення схеми (2),(3). При цьому, будемо вважати, що всі попередні наближення не вийшли за межі околу Ω_0, що буде показано пізніше. З (2), при використанні властивостей скінчених різниць [4], можемо записати

$$x_{k+1} - x^* = \left[F(x_k,\theta_k)^T F(x_k,\theta_k)\right]^{-1} F(x_k,\theta_k)^T\big(F(x_k,\theta_k)(x_k - x^*)$$
$$- F(x_k)\big) = \left[F(x_k,\theta_k)^T F(x_k,\theta_k)\right]^{-1} F(x_k,\theta_k)^T\big((F(x_k,\theta_k) - F(x_k,x^*))\times$$
$$\times(x_k - x^*) - F(x^*)\big) = \left[F(x_k,\theta_k)^T F(x_k,\theta_k)\right]^{-1} F(x_k,\theta_k)^T F(x_k,\theta_k,x^*)\times$$
$$\times(\theta_k - x^*)(x_k - x^*) - \left[F(x_k,\theta_k)^T F(x_k,\theta_k)\right]^{-1}\big(F(x_k,\theta_k,x^*)\ (\theta_k - x^*) +$$

$$+F(x_k,x^*)-F'(x^*)\Big)^T F(x^*).$$

Звідси, з огляду на умови 3)-6) теореми, буде справедливою нерівність

$$\left\|x_{k+1}-x^*\right\| \le BCM\left\|\theta_k-x^*\right\|\left\|x_k-x^*\right\|+BM\varepsilon\left(\left\|\theta_k-x^*\right\|+\left\|x_k-x^*\right\|\right). \quad (9)$$

Далі, використовуючи (3) та умову 7) теореми, отримаємо оцінку

$$\left\|\theta_k-x^*\right\| \le (\alpha+(1-\alpha)L\eta_0^{\tau-1})\left\|x_k-x^*\right\| \le \left\|x_k-x^*\right\|. \quad (10)$$

З (9), тепер можемо записати

$$\left\|x_{k+1}-x^*\right\| \le BCM\left(\alpha\left\|x_k-x^*\right\|^2+(1-\alpha)L\left\|x_k-x^*\right\|^{\tau+1}\right)+$$

$$+2BM\varepsilon\left\|x_k-x^*\right\|. \quad (11)$$

Для $\alpha \ne 0$, використовуючи метод математичної індукції та умови теореми маємо

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

$$+B\gamma(M+\beta)\left\|x_0-x^*\right\|^{\tau-1}\left\|x_0-x^*\right\|^2 \le (BCM+2BM\gamma\eta_0^{\tau-1})\left\|x_0-x^*\right\|^2 \le$$

$$\le K_1\eta_0\eta_0=h_0\eta_0=h_0^{2^1-1}\eta_0. \quad (12)$$

Для k-го наближення ($k \le m$), аналогічно (12) отримаємо

$$\left\|x_{k+1}-x^*\right\| \le (BCM+2BM\gamma\eta_0^{\tau-1})\left\|x_k-x^*\right\|^2 \le K_1 h_0^{2^{k+1}-2}\eta_0^2=$$

$$=h_0^{2^{k+1}-1}\eta_0. \quad (13)$$

Таким чином, послідовність $\{x_k\}\in\Omega_0,\ k=1,2,..m+1$ і (5) є справедливою для цих наближень.

Покажемо тепер виконання (6) для наближень з номерами від $m+2$ і дальше.

$$\left\|x_{m+2}-x^*\right\| \le BCM\left(\alpha\left\|x_{m+1}-x^*\right\|^2+(1-\alpha)L\left\|x_{m+1}-x^*\right\|^{\tau+1}\right)+$$

$$+2BM\varepsilon\left\|x_{m+1}-x^*\right\| \le BCM\left(\alpha\left\|x_m-x^*\right\|\left\|x_{m+1}-x^*\right\|+\right.$$

$$\left.+(1-\alpha)L\left\|x_m-x^*\right\|^\tau\left\|x_{m+1}-x^*\right\|\right)+2BM\gamma\left\|x_m-x^*\right\|^\tau\left\|x_{m+1}-x^*\right\| \le$$

$$\leq \left(BCM + 2BM\gamma \left\| x_m - x^* \right\|^{\tau-1} \right) \left\| x_m - x^* \right\| \left\| x_{m+1} - x^* \right\| \leq K_1 \, h_0^{2^m-1} \eta_0 \leq$$

$$\leq h_0^{2^{m+1}-1} \eta_0 \leq h_0^{(2+1)2^m-1} \eta_0 . \tag{14}$$

Припустимо тепер, що (6) виконується для $m+p$ - го наближення ($p>2$). Тоді, як і в (14)

$$\left\| x_{m+p+1} - x^* \right\| \leq \left(BCM + 2BM \left\| x_m - x^* \right\|^{\tau-1} \right) \left\| x_m - x^* \right\| \times$$

$$\times \left\| x_{m+p} - x^* \right\| \leq K_1 \, h_0^{2^m-1} \eta_0 \, h_0^{(p+1))2^m-1} \eta_0 \leq h_0^{(p+2))2^m-1} \eta_0 . \tag{15}$$

Спрямувавши в (15) p до нескінченості, отримаємо

$$\lim_{p \to \infty} \left\| x_{m+p+1} - x^* \right\| = 0 .$$

Отже, ми показали, що в загальному послідовність $\{x_k\}$, $k = 1,2,...,$ отримана по (2), (3) при умовах теореми збігається до точки x^* – розв'язку задачі (1), мають місце оцінки (5), (6) і кожне наближення схеми лежить в околі Ω_0 .

Розглянемо тепер випадок, коли $\alpha = 0$ і виконується умова 10' теореми. Проведемо індукцію для послідовності наближень з номерами від 1 до $m+1$. Використовуючи (11), отримаємо

$$\left\| x_1 - x^* \right\| \leq K_\tau \eta_0^\tau \eta_0 = h_0^{(\tau+1)^1-1} \eta_0 . \tag{16}$$

Припустимо, що оцінка (7) справедлива для k -го наближення ($k \leq m$), тоді

$$\left\| x_{k+1} - x^* \right\| \leq (BCML + 2BM\gamma) \left\| x_k - x^* \right\|^{\tau+1} \leq$$

$$\leq K_\tau h_0^{((\tau+1)^k-1)(\tau+1)} \eta_0^{\tau+1} \eta_0 = h_0^\tau h_0^{(\tau+1)^{k+1}-\tau-1} \eta_0 = h_0^{(\tau+1)^{k+1}-1} \eta_0 . \tag{17}$$

Аналогічно попередньому, для $k > m+1$, маємо таку оцінку

$$\left\| x_{m+p+1} - x^* \right\| \leq (BCML + 2BM\gamma) \left\| x_m - x^* \right\|^\tau \left\| x_{m+p} - x^* \right\| \leq$$

$$\leq K_\tau h_0^{\tau(\tau+1)^m-\tau} \eta_0^\tau h_0^{(p\tau+1)(\tau+1)^m-1} \eta_0 = h_0^{((p+1)\tau+1))(\tau+1)^m-1} \eta_0 . \tag{18}$$

Формули (17) та (18) показують збіжність (2),(3) для $\alpha = 0$, при умові 8' та справедливість оцінок (7), (8).

Теорему доведено.

Як зазначалось вище, теорема показує збіжність процесу для нелінійних задач про найменші квадрати з ненульовими, але достатньо малими нормами нев'язок. Зокрема важливою в цьому плані є умова 9 теореми. За цієї умови, в загальному випадку метод буде генерувати $m+1$ наближення з квадратичною швидкістю збіжності (формула (13)). Далі швидкість збіжності буде тільки лінійною (формула (15)).

Більш ефективної роботи методу можна сподіватись при $\alpha = 0$. З (17) видно, що до $m+1$ наближення включно, порядок швидкості збіжності становитиме $\tau+1$. Для наближень з номерами від $m+2$ і далі будемо мати надлінійний порядок (формула (18)).

Досліджуючи збіжність загальної схеми [1] та її різницевого аналога (2), (3) для випадку задачі (1) з малою нормою нев'язки, були отримані оцінки швидкості різних порядків. А саме, для довільних $\alpha \in [0,1]$ мали місце оцінки типу

$$\left\|x_k - x^*\right\| \le K_1 \left\|x_{k-1} - x^*\right\|^2 \le h_0^{2^k-1}\eta_0, \tag{19}$$

де $k = 1,2,\dots,m+1$ та

$$\left\|x_{m+p} - x^*\right\| \le K_1 \left\|x_m - x^*\right\| \left\|x_{m+p-1} - x^*\right\| \le h_0^{(p+1)2^m-1}\eta_0, \tag{20}$$

де $p = 2,3,\dots$. $K_1 > 0$ – константа, окремо отримана для кожного методу. Якщо з (19) будемо мати квадратичний порядок збіжності для наближень від 1 до $m+1$, то оцінка (20) вказує на змінний порядок ρ для всіх наступних наближень, який отримаємо з рівняння

$$\rho = 1 + \frac{1}{\rho^{p-1}}. \tag{21}$$

Зокрема, для $p = 2$, $\rho \approx 1.61$, для $p = 3$, $\rho \approx 1.46$.

Формула (23) показує, що з ростом p, порядок збіжності ρ буде прямувати до одиниці, а значить схеми збігаються і швидкість збіжності переходить в лінійну.

Для нерізницевої схеми [1] при $\alpha = \frac{1}{2}$ та схеми (2), (3) при $\alpha = 0$, отримані оцінки типу

$$\left\|x_k - x^*\right\| \le K_\tau \left\|x_{k-1} - x^*\right\|^{\tau+1} \le h_0^{(\tau+1)^k-1}\eta_0, \tag{22}$$

де $k = 1,2,\dots,m+1$

$$\left\|x_{m+p} - x^*\right\| \le K_\tau \left\|x_m - x^*\right\|^\tau \left\|x_{m+p-1} - x^*\right\| \le h_0^{(p\tau+1)(\tau+1)^m-1}\eta_0, \tag{23}$$

де $p = 2,3,\ldots$ $K_\tau > 0$ – константа, окремо отримана для кожного методу. В даному випадку (22) вказує на порядок збіжності рівний $\tau + 1$ для наближень до $m+1$-го. Далі для визначення змінного порядку ρ, що залежить від p, маємо рівняння

$$\rho = 1 + \frac{\tau}{\rho^{p-1}}.$$

Література.

1. М.Я.Бартіш, А.І.Чипурко. Про одну модифікацію методу Гаусса - Ньютона, Математичні Студії. - 1998. - Т.10, №1. -85-92. 2. Дэннис Дж., Шнабель Р. Численные методы безусловной оптимизации и решения нелинейных уравнений.- М. Мир, 1988.- 440 с. 3. Ортега Дж., Рейнболдт В. Итерационные методы решения нелинейных систем уравнений со многими неизвестными. - М. Мир, 1975.- с. 4. Ульм С.Ю. Об обобщенных разделенных разностях. I. Изд. АН ЭССР. Физика. Матем., 1967, 16, N1, C.13-26.

Bartish M.Ja., Nykolajchuk O.V., Chypurko A.I.

About Some Difference Modifications of Gauss-Newton Method for a Nonlinear Least Squares Problem with Small Residual Norm

In this paper we study the difference analogue of the generalized scheme of the modifications of Gauss-Newton method, proposed in [1]. There are presented the conditions of convergence and estimates of the accuracy of the method for solving Nonlinear Least Squares Problem in the case when the residual is sufficiently small but not equal zero. We determine the influence of valid parameter on rate of convergence of process.

Стаття надійшла до редколегії 8.12.1998

УДК 539.3

П.П.Вагін, Р.Б.Малець, Г.А.Шинкаренко

Квазілінеаризація задачі термопружності для гнучких оболонок з деформівною нормаллю

Дослідження нелінійної теорії оболонок типу Тимошенка з врахуванням обтиску нормального елемента можна знайти в працях