УДК 518.517.948

М.Я.Бартіш, С.М.Шахно

ДЕЯКІ МЕТОДИ РОЗВ"ЯЗУВАННЯ НЕЛІНІЙНОЇ ЗАДАЧІ НАЙМЕНШИХ КВАДРАТІВ

Нелінійні задачі найменших квадратів виникають при оцінюванні параметрів і перевірці гіпотез у математичній статистиці, параметрів фізичних процесів за результатами вимірювань, в керуванні різними об"єктами, процесами тощо.

У цій праці пропонується метод для розв"язування систем нелінійних рівнянь у сенсі найменших квадратів, тобто рівняння

$$F'(x)^T F(x) = 0, \qquad /1/$$

де $F: D \subset R^n \longrightarrow R^m \quad (m \geq n)$.
Ітераційна формула цього методу має вигляд

$$x_{k+1} = x_k - [F'(\bar{x}_k)^T F'(\bar{x}_k)]^{-1} F'(\bar{x}_k)^T F(\bar{x}_k),$$

$$\bar{x}_k = (1-\mu)x_k + \mu\varphi(x_k), \quad k = 0, 1, \ldots; \quad 0 \leq \mu \leq 1, \qquad /2/$$

де $\varphi: R^n \longrightarrow R^m$ – деякий допоміжний оператор, для якого

$$x^* = \varphi(x^*). \qquad /3/$$

Легко бачити, що при $\mu = 0$ отримуємо відомий метод Гауса-Ньютона [1]. За допомогою параметра μ можна вибрати метод із найбільшою швидкістю збіжності серед методів класу /2/.

Розглянемо деякі варіанти вибору оператора $\varphi(x)$.

1. Випадок з нульовою нев"язкою: $F(x_*) = x_* - \varphi(x_*) = 0$.

Умови і швидкість збіжності ітерацій /2/ встановлюються теоремою.

<u>Теорема 1.</u> Нехай $F: R^n \longrightarrow R^m (m \geq n)$ і функція $f(x) = \frac{1}{2} F(x)^T F(x)$ двічі неперервно диференційовна на відкритій випуклій множині $D \subset R^n$. Припустимо, що $F''(x) \in Lip_\gamma(D)$, $\|\varphi'(x)\| \leq \alpha$, $\|\varphi''(x)\| \leq L$ іх $x \in D$, а також що існують $x_* \in D$ і $\lambda \geq 0$ такі, що $F'(x_*)^T F(x_*) = 0$, 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 /2/, коректно визначена, збігається до x_* і задовольняє нерівності

$$r_n = \|x_n - x_*\| \leq (h_0 r_0)^{2^n - 1} r_0, \qquad /4/$$

де $h_0 = \frac{1+\alpha}{2\lambda}\left[\frac{N}{3} r_0 \mu^2 (1+\alpha)^2 + L\left(|2\mu - 1| + 2\mu\alpha\right)\right] < \frac{1}{r_0}$,

$$r_0 = \|x_0 - x_*\|, \ \Omega(x_*, r_0) = \{x : \|x - x_*\| < r_0\}, \ n = 0, 1, \dots .$$

<u>Доведення.</u> Виходячи з /2/, шляхом тотожних перетворень отримуємо

$$x_1 - x_* = x_0 - x_* - [F'(x_0)^T F'(x_0)]^{-1} F'(x_0)^T F(x_0) = -[F'(x_0)^T F'(\bar{x}_0)]^{-1} \times$$
$$\times F'(\bar{x}_0)^T \Big[\int_0^1 \big(F''(\bar{x}_0 + \tau(x_0 - \bar{x}_0)) - F''(\bar{x}_0 + \tau(x_* - \bar{x}_0))\big)(1-\tau)\,d\tau\,(x_0 - \bar{x}_0)^2 +$$
$$+ \int_0^1 F''(\bar{x}_0 + \tau(x_* - \bar{x}_0))(1-\tau)\,d\tau\,(x_0 - x_*)(x_0 - 2\bar{x}_0 + x_*)\Big]. \qquad /5/$$

Оцінюємо норми виразів $\|x_0 - \bar{x}_0\|$, $\|\bar{x}_0 - x_*\|$, $\|x_0 - 2\bar{x}_0 + x_*\|$:

$$\|x_0 - \bar{x}_0\| = \|x_0 - (1-\mu)x_0 - \mu\varphi(x_0)\| = \|\mu(\varphi(x_0) - x_0)\| =$$

$$= \|\mu(\varphi(x_0) - \varphi(x_*) + x_* - x_0)\| \leq \mu(1+\alpha)\|x_0 - x_*\|; \qquad /6/$$

$$\|x_0 - x_*\| \leq (1 - \mu + \mu\alpha)\|x_0 - x_*\|; \qquad /7/$$

$$\|x_0 - 2\bar{x}_0 + x_*\| \leq \left(|2\mu - 1| + 2\mu\alpha\right)\|x_0 - x_*\|. \qquad /8/$$

Із /5/ за допомогою оцінок /6/–/8/ записуємо

$$\|x_1 - x_*\| \leq \frac{1+\alpha}{2\lambda}\left(\frac{N}{3}\|x_0 - x_*\|^3 \mu^2 (1+\alpha)^2 + L\left(|2\mu - 1| + 2\mu\alpha\right)\|x_0 - x_*\|^2\right).$$

Далі шляхом математичної індукції отримуємо оцінку /4/. Теорема доведена.

Дослідимо вплив параметра μ на швидкість збіжності ітераційного процесу /2/. Швидкість збіжності можна підвищувати двома шляхами: збільшенням порядку збіжності і зменшенням знаменника збіжності $h_0 r_0$. У знаменник $h_0 r_0$ входить величина $\partial(\mu) = |2\mu - 1| + 2\mu\alpha$, найменше значення якої досягається при $\mu = 0{,}5$ у випадку $\alpha < 1$ і при $\mu = 0$ у протилежному випадку. Тоді ітераційний процес матиме найбільшу швидкість збіжності і найменш жорсткі умови на вибір початкового наближення x_0 /ширша область збіжності/. Як бачимо з /4/, при достатньо малих значеннях z_0 процес /2/ із значенням $\mu=0,5$ і $\alpha<1$ має вищу швидкість збіжності, ніж метод Гауса-Ньютона.

2. Випадок $F(x_*)\neq 0$. Умову /3/ задовольняє оператор

$$\varphi(x) = x-\beta F'(x)^T F(x), \quad \beta>0,$$

який відповідає градієнтному методу мінімізації функціоналу $f(x)=\frac{1}{2}F(x)^T F(x)$.

<u>Теорема 2.</u> Нехай виконуються умови теореми I і $\|F'(x)\|\le\alpha$,

$$\|(F'(x)-F'(x_*))^T F(x_*)\| \le \sigma\|x-x_*\| \qquad /9/$$

для всіх $x_0\in D$, причому $\lambda>\sigma(1+\delta)\ge 0$, $\delta=\mu\beta(\alpha^2+\sigma)$.

Тоді для довільного $c\in(1,\lambda/((1+\delta)\sigma))$ існує $\varepsilon>0$ таке, що для всіх $x_0\in\Omega(x_*,\varepsilon)$ послідовність /2/ коректно визначена, збігається до x_* і задовольняє нерівності

$$\|x_{k+1}-x_*\|\le\frac{c\sigma(1+\delta)}{\lambda}\|x_k-x_*\|+\frac{c\alpha}{2\lambda}\left[\frac{N}{3}\|x_k-x_*\|\delta^2+L(1+2\delta)\right]\|x_k-x_*\|^2; \qquad /10/$$

$$\|x_{k+1}-x_*\|\le\frac{c\sigma(1+\delta)+\lambda}{2\lambda}\|x_k-x_*\|<\|x_k-x_*\|. \qquad /11/$$

<u>Доведення.</u> Нехай c – фіксована константа з інтервалу $(1,\lambda/((1+\delta)\sigma))$. Тоді існує $\varepsilon_1>0$ таке, що

$$\|[F'(\bar{x}_0)^T F'(\bar{x}_0)]^{-1}\|\le\frac{c}{\lambda} \quad \text{для всіх} \quad x_0\in\Omega(x_*,\varepsilon_1). \qquad /12/$$

Нехай

$$\varepsilon=\min\left\{\varepsilon_1,\ \frac{\lambda-c\sigma(1+\delta)}{c\alpha\left[\frac{N}{3}\|x_0-x_*\|\delta^2+L(1+2\delta)\right]}\right\}. \qquad /13/$$

Тоді на першому кроці x_1 коректно визначене і справедлива рівність

$$x_{n+1}-x_*=[F'(\bar{x}_n)^T F'(x_n)]^{-1}F'(\bar{x}_n)^T\Big[\int_0^1\big(F''(\bar{x}_n+\tau(x_n-\bar{x}_n))-F''(x_n+\tau(x_*-x_n))\big)(1-\tau)\,d\tau\,(x_n-\bar{x}_n)^2+\int_0^1 F''(\bar{x}_n+\tau(x_*-\bar{x}_n))\times(1-\tau)\,d\tau\,(x_n-x_*)(x_n-2\bar{x}_n+x_*)\Big]+(F'(\bar{x}_n)-F'(x_*))^T F(x_*), \qquad /14/$$

де $n=0$.

Оскільки
$$F'(x)^T F(x) = F'(x)^T F(x) - F'(x_*)^T F(x_*) = F'(x)^T F'(\tilde{x})^T (x - x_*) +$$
$$+ (F'(x)^T - F'(x_*)^T) F(x_*), \quad (\tilde{x} = x + \theta(x - x_*), \quad 0 \leq \theta \leq 1),$$
то $\|F'(x)^T F(x)\| \leq (\alpha^2 + \sigma)\|x - x_*\|$ для всіх $x \in \Omega(x_*, \varepsilon)$.
Тепер отримаємо оцінки:
$$\|\bar{x}_n - x_n\| \leq \mu\beta(\alpha^2 + \sigma)\|x_n - x_*\| = \delta\|x_n - x_*\|; \qquad /15/$$
$$\|\bar{x}_n - x_*\| \leq (1 + \mu\beta(\alpha^2 + \sigma))\|x_n - x_*\| = (1 + \delta)\|x_n - x_*\|; \qquad /16/$$
$$\|x_n - 2\bar{x}_n + x_*\| \leq (1 + 2\mu\beta(\alpha^2 + \sigma))\|x_n - x_*\| = (1 + 2\delta)\|x_n - x_*\|. \qquad /17/$$

Із рівності /14/, враховуючи оцінки /12/, /15/–/17/, отримуємо для $n = 0$
$$\|x_1 - x_*\| \leq \frac{c\sigma(1+\delta)}{\lambda}\|x_0 - x_*\| + \frac{c\alpha}{2\lambda}\Big[\frac{N}{3}\|x_0 - x_*\|^3\delta^2 + L\|x_0 - x_*\|^2(1 + 2\delta)\Big], \qquad /18/$$
що доводить /10/ для $k = 0$.

Із /13/ і /18/ маємо
$$\|x_1 - x_*\| \leq \|x_0 - x_*\|\Big\{\frac{c\sigma(1+\delta)}{\lambda} + \frac{c\alpha}{2\lambda}\Big[\frac{N}{3}\|x_0 - x_*\|\delta^2 + L(1+2\delta)\Big]\|x_0 - x_*\|\Big\} \leq$$
$$\leq \|x_0 - x_*\|\Big[\frac{c\sigma(1+\delta)}{\lambda} + \frac{\lambda - c\sigma(1+\delta)}{2\lambda}\Big] = \frac{c\sigma(1+\delta) + \lambda}{2\lambda}\|x_0 - x_*\| < \|x_0 - x_*\|,$$
що доводить /11/ для $k = 0$. Далі доведення здійснюємо індукцією.

Наслідок. Нехай виконані умови теореми 2. Якщо $F(x_*) = 0$, то існує $\varepsilon > 0$ таке, що для всіх $x_0 \in \Omega(x_*, \varepsilon)$ послідовність $\{x_k\}$, породжена методом /2/, коректно визначена і збігається квадратично до x_*.

Доведення. Приймаючи у /9/ $\sigma = 0$, з /10/ отримуємо нерівність, яка свідчить про квадратичну збіжність послідовності $\{x_k\}$ до x_*.

З теореми 2 бачимо, що на багатьох задачах з невиродженою матрицею $F'(x_*)^T F'(x_*)$ метод /2/ має невисоку локальну збіжність, а на деяких з них він взагалі локально розбігається. Із /9/ бачимо, що параметр σ є абсолютною мірою нелінійності величини нев"язки в розв"язку задачі, а з /10/ – що $\frac{\sigma(1+\delta)}{\lambda}$ можна розглядати як відносну міру нелінійності і величини нев"язки в розв"язку задачі. З теореми /2/ випливає, що швидкість збіжності методу /2/ зменшується із зростанням відносної нелінійності чи відносної нев"язки в розв"язку задачі. Якщо одна з цих двох величин надто велика, то метод може взагалі розбігатися.

Недолік методу /2/ полягає в тому, що він не є коректно визначеним, якщо $F'(x_k)$ не має повного стовпцевого рангу /$F'(x_k)^T F'(x_k)$ невизначена/. У цьому випадку доцільно використовувати метод

$$x_{k+1} = x_k - [F'(\bar{x}_k)^T F'(\bar{x}_k) + \gamma_k I]^{-1} F'(\bar{x}_k)^T F(x_k),$$

$$\bar{x}_k = (1-\mu)x_k + \mu\varphi(x_k), \quad k = 0, 1, \ldots, \qquad /19/$$

частковим випадком якого при $\mu = 0$ є відомий метод Левенберга-Марквардта [1].

Властивості локальної збіжності методу /19/ аналогічні властивостям методу /2/ і наведені в теоремі 3.

Теорема 3. Нехай виконані умови теореми 2 і послідовність невід"ємних дійсних чисел обмежена зверху числом b. Якщо $\sigma(1+\delta) < \lambda$, то для довільного $c \in \left(1, \frac{\lambda+b}{\sigma(1+\delta)+b}\right)$ існує $\varepsilon > 0$ таке, що для всіх $\Omega(x_*, \varepsilon)$ послідовність, породжена методом /19/, коректно визначена і задовольняє співвідношення

$$\|x_{k+1} - x_*\| \le \frac{c(\sigma(1+\delta)+b)}{\lambda+b}\|x_k - x_*\| + \frac{c\alpha}{2(\lambda+b)}\left[\frac{N}{3}\|x_k - x_*\|\delta^2 + L(1+2\delta)\right]\|x_k - x_*\|^2$$

і $$\|x_{k+1} - x_*\| \le \frac{c(\sigma(1+\delta)+b)+(\lambda+b)}{2(\lambda+b)}\|x_k - x_*\| < \|x_k - x_*\|.$$

Якщо $F(x_*) = 0$ і $\gamma_k = O(\|F'(x_k)^T F(x_k)\|)$, то $\{x_k\}$ збігається квадратично до x_*.

Ітераційний процес /19/ можна розглядати як збурення процесу /2/. Тому доведення теореми легко здійснити за схемою доведення теореми 2 з використанням результатів праці [2].

1. Деннис Дж., Шнабель Р. Численные методы безусловной оптимизации и решения нелинейных уравнений. М., 1988. 440 с.
2. Шахно С.М. Построение и исследование некоторых методов типа Ньютона-Канторовича для решения нелинейных функциональных уравнений: Автореф. дис. ... канд. физ.-мат. наук. К., 1988. 17 с.

Стаття надійшла до редколегії 06.04.93