$$\varphi_k(x,y)=\cos\left(\frac{kn(y-g(x))}{h(x)-g(x)}\right), k=1,2,...$$

1. Варга Р. Функциональный анализ и теория аппроксимаций в численном анализе. М.: 1974. 2. Жук М.В. Дослідження швидкості збіжності методу Канторовича для нелінійних диференціальних рівнянь. //Укр. матем. журн. 1976. Т.28. №2. С.183–193. 3. Михлин С.Г. Численная реализация вариационных методов. М.: Наука, 1966.

Стаття надійшла до редколегії 3.09.97

УДК 519.6

Бартіш М.Я.,Чипурко А.І.

Дослідження одного рекурсивного методу розв'язування задачі про найменші квадрати

Розглянемо постановку задачі про найменші квадрати:

$$\text{знайти} \quad \min\frac{1}{2}F(x)^T F(x) \tag{1}$$

де $F(x):R^n \to R^m$ – нелінійна по x , $m \geq n$.

На практиці для розв'язування (1) найчастіше використовують метод Гаусса-Ньютона [1]. Даний метод на кожному кроці вимагає обчислення матриці Якобі функції $F(x)$, що у випадку складності функції є операцією трудомісткою.

В цій роботі пропонується рекурсивна модифікація методу Гаусса-Ньютона :

$$x_k^{\nu+1}=x_k^\nu-\left[F'(x_k)^T F'(x_k)\right]^{-1}F'(x_k)^T F(x_k^\nu), \tag{2}$$

$$x_{k+1}^0=x_{k+1}=x_k^t \ , \ x_0^0=x_0 \ ,$$

$$\nu=0,1,\dots,t-1, \ \mathrm{k}=0,1,\dots .$$

В методі (2) матриця $F'(x_k)$ обчислюється один раз на t кроків, що дозволяє значно зменшити кількість обчислень.

Умови збіжності та оцінки швидкості збіжності методу дає

© Бартіш М.Я., Чипурко А.І., 1997

Теорема 1. Нехай виконуються умови:

1) $F(x)$ – двічі неперервно-диференційовна в області $D \subset R^n$;

2) $\forall x \in D \quad \|F'(x)\| \le C$, $\|F''(x)\| \le M$;

3) $\forall x \in D$ існує обернений оператор $[F'(x)^T F'(x)]^{-1}$, причому $\|[F'(x)^T F'(x)]^{-1}\| \le B$;

4) в області D існує точка x^*, яка є розв'язком задачі (1), причому $F(x^*) = 0$;

5) існує константа $\eta_0 > 0$, така що $h_0 = l\eta_0 < 1$, де $l = \dfrac{3BCM}{2}$.

Тоді для будь-якого початкового наближення $x_0 \in \Omega_0 = \left\{x : \|x - x^*\| < \eta_0\right\} \subset D$ послідовність $\{x_k\}$, породжена методом (2), коректно визначена та збігається до x^* і має місце оцінка

$$\|x_k - x^*\| \le h_0^{(t+1)^k - 1}\eta_0. \tag{3}$$

Доведення . Покажемо справедливість (3) для $k = 1$. Застосуєм метод математичної індукції по $\nu = 0,1,\ldots,t-1$.

Оскільки $x_k^0 = x_k \quad \forall k = 0,1,\ldots$, то кожен крок при $\nu = 0$ методу (2) співпадає з кроком методу Гаусса-Ньютона. В даному випадку для $\nu = 0$ неважко отримати оцінку

$$\|x_0^1 - x^*\| \le \frac{BCM}{2}\|x_0 - x^*\|^2 \le l\eta_0^2 = h_0\eta_0. \tag{4}$$

Для $\nu = 1$ запишемо тотожність:

$$x_0^2 - x^* = x_0^1 - x^* - \left[F'(x_0)^T F'(x_0)\right]^{-1} F'(x_0)^T (F(x_0^1) - F(x^*)) =$$

$$= \left[F'(x_0)^T F'(x_0)\right]^{-1} F'(x_0)^T \left\{F'(x_0)(x_0^1 - x^*) - (F(x_0^1) - F(x^*))\right\}.$$

Використавши розклад $F(x)$ в околі точки x^* у ряд Тейлора, отримаєм

$$x_0^2 - x^* = \left[F'(x_0)^T F'(x_0)\right]^{-1} F'(x_0)^T \left\{(F'(x_0) - F'(x^*))(x_0^1 - x^*) - \right.$$

$$\left. - \int_0^1 F''(x^* + \tau(x_0^1 - x^*))(1-\tau)d\tau \ (x_0^1 - x^*)^2\right\}.$$

Звідси маємо

$$\|x_0^2 - x^*\| \le BCM\|x_0 - x^*\|\|x_0^1 - x^*\| + \frac{BCM}{2}\|x_0^1 - x^*\|^2 \le$$

$$\le \frac{3BCM}{2}\|x_0 - x^*\|\|x_0^1 - x^*\| \le l\eta_0 h_0 \eta_0 = h_0^2\eta_0. \tag{5}$$

Нехай $\|x_0^{t-1} - x^*\| \le h_0^{t-1}\eta_0$, тоді для $\nu = t-1$, аналогічно (5) отримаєм

$$\|x_0^t - x^*\| \le \frac{3BCM}{2}\|x_0 - x^*\|\|x_0^{t-1} - x^*\| \le l\eta_0 h_0^{t-1}\eta_0 = h_0^t\eta_0, \tag{6}$$

або $\|x_1 - x^*\| \le h_0^t\eta_0$.

Припустимо, що твердження теореми справедливі для послідовності наближень $x_2, x_3, \dots x_k$. Тоді має місце оцінка

$$\|x_k - x^*\| \le h_0^{(t+1)^k - 1}\eta_0$$

і при $\nu = 0$ отримаєм

$$\|x_k^1 - x^*\| \le \frac{BCM}{2}\|x_k - x^*\|^2 \le l h_0^{2(t+1)^k - 2}\eta_0^2 = h_0^{2(t+1)^k - 1}\eta_0.$$

При $\nu = 1$, аналогічно (5), маємо

$$\|x_k^2 - x^*\| \le \frac{3BCM}{2}\|x_k - x^*\|\|x_k^1 - x^*\| \le l h_0^{(t+1)^k - 1}\eta_0\, h_0^{2(t+1)^k - 1}\eta_0 = h_0^{3(t+1)^k - 1}\eta_0.$$

Неважко показати, що $\|x_k^t - x^*\| = \|x_{k+1} - x^*\| \le h_0^{(t+1)^{k+1} - 1}\eta_0$.

Теорема доведена.

Розглянемо різницевий аналог методу (2)

$$x_k^{\nu+1} = x_k^\nu - \left[F(x_k, \bar{x}_k)^T F(x_k, \bar{x}_k)\right]^{-1} F(x_k, \bar{x}_k)^T F(x_k^\nu), \tag{7}$$

$x_{k+1}^0 = x_{k+1} = x_k^t$, $x_0^0 = x_0$, $\bar{x}_k$ – близьке до x_k,

$\nu = 0, 1, \dots, t-1$, $k = 0, 1, \dots$.

Тут $F(x, y)$ – перша поділена різниця [2].

Теорема2. Нехай виконуються умови:

1) $F(x)$ – неперервна в області $D \subset R^n$;

2) $\forall x, y \in D$ $\|F(x, y)\| \le C$;

3) $\forall x, y, z \in D$ $\|F(x, y, z)\| \le M$, де $F(x, y, z)$ – друга поділена різниця ;

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

5) в області D існує точка x^*, яка є розв'язком задачі (1), причому $F(x^*)=0$;

6) $\|\bar{x}_k - x_k\| < \|x_k - x^*\| \quad \forall k = 0,1,\ldots$;

7) існує константа $\eta_0 > 0$, така що $h_0 = l\eta_0 < 1$, де $l = 4BCM$.

Тоді $\forall\ x_0 \in \Omega_0 = \left\{x: \|x - x^*\| < \eta_0\right\} \subset D$ послідовність $\{x_k\}$, породжена методом (7), коректно визначена та збігається до x^* і має місце оцінка:

$$\|x_k - x^*\| \le h_0^{(t+1)^k - 1}\eta_0 . \tag{8}$$

Доведення. Перевіримо справедливість твердження для $k=1$ Застосуєм метод математичної індукції для $\nu = 0,1,\ldots,t-1$.

Використовуючи (7), запишемо тотожність:

$$x_0^1 - x^* = \left[F(x_0,\bar{x}_0)^T F(x_0,\bar{x}_0)\right]^{-1} F(x_0,\bar{x}_0)^T \left\{F(x_0,\bar{x}_0)(x_0 - x^*) - F(x_0)\right\}.$$

Скориставшись аналогом інтерполяційної формули Ньютона [2], враховуючи умову 5 теореми 2, можемо записати

$$F(x_0) = F(x_0,x^*)(x_0 - x^*) .$$

Тоді отримаєм

$$x_0^1 - x^* = \left[F(x_0,\bar{x}_0)^T F(x_0,\bar{x}_0)\right]^{-1} F(x_0,\bar{x}_0)^T (F(x_0,\bar{x}_0) - F(x_0,x^*)) \times$$
$$\times (x_0 - x^*) = \left[F(x_0,\bar{x}_0)^T F(x_0,\bar{x}_0)\right]^{-1} F(x_0,\bar{x}_0)^T F(x_0,\bar{x}_0,x^*)(\bar{x}_0 - x^*) \times$$
$$\times (x_0 - x^*) \text{ і}$$

$$\|x_0^1 - x^*\| \le BCM\|\bar{x}_0 - x_0 + x_0 - x^*\|\|x_0 - x^*\| \le 2BCM\eta_0^2 = h_0\eta_0 .$$

Для $\nu = 1$ маємо

$$x_0^2 - x^* = \left[F(x_0,\bar{x}_0)^T F(x_0,\bar{x}_0)\right]^{-1} F(x_0,\bar{x}_0)^T (F(x_0,\bar{x}_0) - F(x_0,x^*) +$$
$$+ F(x_0,x^*) - F(x_0^1,x^*))(x_0^1 - x^*) = \left[F(x_0,\bar{x}_0)^T F(x_0,\bar{x}_0)\right]^{-1} F(x_0,\bar{x}_0)^T \times$$
$$\times (F(x_0,\bar{x}_0,x^*)(\bar{x}_0 - x^*) + F(x_0,x_0^1,x^*)(x_0 - x_0^1))(x_0^1 - x^*) .$$

Звідси

$$\|x_0^2 - x^*\| \le BCM(\|\bar{x}_0 - x_0 + x_0 - x^*\| + \|x_0 - x^*\| + \|x_0^1 - x^*\|)\|x_0^1 - x^*\| \le$$
$$\le 4BCM\|x_0 - x^*\|\|x_0^1 - x^*\| \le l\eta_0 h_0 \eta_0 = h_0^2 \eta_0 .$$

Легко бачити, що для $\nu = t-1$ $\|x_0^t - x^*\| \le h_0^t\eta_0$ тобто $\|x_1 - x^*\| \le h_0^t\eta_0$.

Аналогічно попередній теоремі неважко показати, що

$$\left\| x_{k+1} - x^* \right\| \le h_0^{(t+1)^{k+1}-1} \eta_0 .$$

Теорема доведена.

Використовуючи результати теорем, можна розглянути питання про вибір оптимальної глибини рекурсії t_{opt} з метою отримання результату з заданою точністю ε за мінімальне число обчислень.

Позначимо через K_1 кількість обчислень необхідних для переходу від x_k^0 до x_k^1, $k = 0,1,\dots$, а через K_2 – кількість обчислень необхідних для переходу від x_k^{ν} до $x_k^{\nu+1}$, $\nu = 1,\dots,t-1$, $k = 0,1,\dots$. Очевидно, що $K_1 >> K_2$. Таким чином, на кожній ітерації методу (2) виконується $K_1 + (t-1)K_2$ обчислень, а для того, щоб отримати k-те наближення потрібно виконати $(K_1 + (t-1)K_2)k$ обчислень.

Отже, для визначення оптимальної глибини рекурсії маємо задачу

$$(K_1 + (t-1)K_2)k \to \min_t$$

при обмеженнях $h_0^{(t+1)^k-1} \eta_0 \le \varepsilon$.

Для знаходження розв'язку даної задачі маємо рівняння

$$(1+t^*)\ln(1+t^*) = \frac{K_1}{K_2} + t^* - 1 ,$$

а за оптимальну глибину рекурсії вибираєм число $t_{opt} = [t^* + 0.5]$, де []- ціла частина числа.

Відзначимо, що K_1 та K_2 можна замінити величинами T_1 та T_2, які є часом, необхідним для обчислення x_k^1 при заданому x_k^0 та $x_k^{\nu+1}$ при заданому x_k^{ν}, $\nu = 1,\dots,t-1$ відповідно.

1. Дэннис Дж., Шнабель Р. Численные методы безусловной оптимизации и решения нелинейных уравнений . – М. Мир, 1988. – 440 с. 2. Ульм С.Ю. Об обобщенных разделенных разностях. I. Изд. АН ЭССР. Физика., Матем., 1967, 16, N1, С.13-26.

Стаття надійшла до редколегії 26.11.97