ДОКЛАДЫ АКАДЕМИИ НАУК, 2014, том 455, № 5, с. 512–515

МАТЕМАТИКА

УДК 519.615

МЕТОДЫ p-ГО ПОРЯДКА ДЛЯ РЕШЕНИЯ СИСТЕМ НЕЛИНЕЙНЫХ УРАВНЕНИЙ

© 2014 г. Академик Ю. Г. Евтушенко, А. А. Третьяков

Поступило 10.12.2013 г.

DOI: 10.7868/S0869565214110061

Численным методам решения систем нелинейных уравнений посвящена обширная литература. Среди этих публикаций наиболее популярным является метод Ньютона, обладающий квадратичной скоростью сходимости. В статье П.Л. Чебышева, Опубликованной в 1838 г. был предложен метод, сходящийся с более высокой скоростью. В последние годы появились работы, в которых изучали методы, обладающие сверхквадратичной скоростью сходимости.

В данной работе предложены методы с p-м порядком скорости сходимости как по функционалу, так и по аргументу.

Дана система из n нелинейных уравнений с неизвестным n-мерным вектором x

$$F(x) = 0_n, \quad x \in \mathbb{R}^n, \tag{1}$$

где вектор-функция $F(x) = (f_1(x), f_2(x), \ldots, f_n(x))^{\mathrm{т}}$, $F \in C^{p+1}(\mathbb{R}^n)$, 0_n есть n-мерный нулевой вектор. Считаем, что множество решений этой системы X непусто.

Предположим, что существует точка x_*, являющаяся решением системы (1). Возьмем точку x, принадлежащую некоторой ε-окрестности $U_\varepsilon(x_*)$ точки x_*, расстояние между этими точками $\|x - x_*\|$ считаем величиной первого порядка малости $O(\varepsilon)$. Воспользуемся формулой Тейлора разложения функции $F(x)$ в окрестности точки x:

$$F(x_*) = 0_n = F(x) - F'(x)h + \frac{1}{2}F^{(2)}(x)[h]^2 + \ldots$$

$$\ldots + \frac{(-1)^p}{p!}F^{(p)}(x)[h]^p + O_n(\|h\|^{p+1}). \tag{2}$$

Здесь $h = x - x_*$, $[h]^p$ — мультивектор, $O_n(\|\cdot\|^{p+1})$ — n-мерный вектор, каждая компонента которого имеет $(p+1)$-й порядок малости, $F'(x)$ — матрица Якоби; ее ij-й элемент есть $\dfrac{\partial f_i}{\partial x^j}$.

Предполагаем, что матрица $F'(x_*)$ не вырождена. Тогда существует окрестность U_ε точки x_*, где матрица $F'(x)$ также не вырождена и где можно определить вектор-функцию $N_1(x) = [F'(x)]^{-1}F(x)$. Умножим (2) слева на матрицу $[F'(x)]^{-1}$, получим

$$x_* = x - N_1(x) - [F'(x)]^{-1}\Big[\frac{1}{2}F^{(2)}(x)[h]^2 + \ldots$$

$$\ldots + \frac{(-1)^p}{p!}F^{(p)}(x)[h]^p\Big] + O_n(\|h\|^{p+1}). \tag{3}$$

Здесь входящий в правую часть (3) вектор x можно рассматривать как параметр, принадлежащий некоторой окрестности точки x_*. Найдем точку x_* методом последовательных приближений, исходя из представления (3).

Введем функцию

$$x_1(x) = x - N_1(x). \tag{4}$$

Отсюда следует $x_1(x) - x_* = h(x) - (F'(x))^{-1}F(x)$. Подставим в (4) выражение для $F(x)$, взятое из (2), и отбросим члены второго порядка малости; получим

$$\|x_1(x) - x_*\| \approx O(\|x - x_*\|^2). \tag{5}$$

В качестве приближенного решения системы (1) возьмем функцию $x_1(x)$. Разлагая $F(x_1(x))$ в ряд в окрестности точки x, получим

$$F(x_1(x)) = F(x) - F'(x)N_1(x) + O_n(\|x - x_*\|^2) =$$

$$= O_n(\|x - x_*\|^2). \tag{6}$$

Таким образом, вектор $x_1(x)$ определяет приближенное решение системы (1) с погрешностью второго порядка малости как по норме невязки вектор-функции F, так и по норме разности векторов $x_1(x)$ и x_*.

Вычислительный центр им. А.А. Дородницына Российской Академии наук, Москва
System Research Institute, Polish Academy of Sciences, Warsaw
University of Podlasie, Siedlce, Poland

В правой части равенства (3) отбросим члены третьего порядка малости; получим

$$x_* = x - N_2(x) + O_n(\|x - x_*\|^3),$$
$$x_2(x) = x - N_2(x),$$

где $N_2(x) = N_1(x) + \frac{1}{2}[F'(x)]^{-1}F^{(2)}(x)[N_1(x)]^2$.

Если в качестве приближенного решения системы (2) взять $x_2(x) = x - N_2(x)$, то найдем оценки, аналогичные (5) и (6):

$$F(x_2(x)) = F(x) - F'(x)N_2(x) + \frac{1}{2}F^{(2)}(x)[N_2(x)]^2 +$$
$$+ O_n(\|x - x_*\|^3) = O_n(\|x - x_*\|^3),$$
$$\|x_2(x) - x_*\| \approx O(\|x - x_*\|^3).$$

Аналогично, отбрасывая в (3) члены четвертого порядка малости, находим

$$x - x_* = N_3(x) - O_n(\|x - x_*\|^4), \quad x_3(x) = x - N_3(x),$$

где $N_3 \triangleq N_1 + [F'(x)]^{-1}\left\{\frac{1}{2}F^{(2)}(x)[N_2]^2 - \frac{1}{3!}F^{(3)}(x)[N_2]^3\right\}$

и

$$\|F(x_3(x))\| \approx O(\|x - x_*\|^4),$$
$$\|x_3(x) - x_*\| \approx O(\|x - x_*\|^4).$$

Продолжая этот процесс, приходим к следующим формулам:

$$x - x_* = N_p(x) + O_n(\|x - x_*\|^{p+1}),$$
$$x_p(x) = x - N_p(x), \tag{7}$$

где

$$N_p(x) =$$
$$= N_1(x) + [F'(x)]^{-1}\left\{\sum_{k=2}^{p}\frac{(-1)^k}{k!}F^{(k)}(x)[N_{p-1}(x)]^k\right\}. \tag{8}$$

Обобщая все вышесказанное, можно утверждать, что верен следующий результат (см. также [2–5]).

Т е о р е м а 1. *Пусть $F \in C^{p+1}(\mathbb{R}^n, \mathbb{R}^n)$, x_* — решение системы* (1), *матрица $F'(x_*)$ не вырождена, $x \in U_\varepsilon(x_*)$, где $\varepsilon > 0$ достаточно малое. Тогда для точек $x_p(x)$, вычисленных по формулам* (7), (8), *имеют место оценки*

$$\|F(x_p(x))\| \le C_1\|x - x_*\|^{p+1},$$
$$\|x_p(x) - x_*\| \le C_2\|x - x_*\|^{p+1}, \tag{9}$$

где C_1, $C_2 > 0$ — независимые константы.

Замечательным свойством ряда (8) является то, что для получения формул произвольного порядка точности достаточно обратить только одну матрицу $F'(x)$. Особенностью метода является следующее: даже для функций с нулевыми высшими производными ряд (8), вообще говоря, является бесконечным, но сходящимся к решению x_*. Это иллюстрирует следующий

П р и м е р. Пусть $n = 1$, $F(x) = x + x^2$; тогда $F'(x) = 1 + 2x$, $(F'(x))^{-1} = \frac{1}{1 + 2x}$, $F^{(2)}(x) = 2$. Корнями уравнения (1) являются $\overset{1}{x}_* = 0$ и $\overset{2}{x}_* = -1$. Все производные функции F, начиная с третьей, равны нулю; они не дают вклада в формулу (2). Однако для $x \in U_\varepsilon(0)$, вычисляя x_1 по формуле (7) при $p = 3$, имеем

$$F'(x) = 1 + 2x, \quad (F'(x))^{-1} = \frac{1}{1 + 2x}, \quad N_1 = \frac{x + x^2}{1 + 2x},$$
$$N_2 = N_1 + \frac{(x + x^2)^2}{(1 + 2x)^3}, \quad N_3 = N_1 + \frac{1}{1 + 2x}N_2^2,$$
$$x_3 = x - N_3(x).$$

Здесь $x_i = x_i(x)$, $N_i = N_i(x)$, $i = 1, 2, 3$. Используя метод Ньютона (т.е. взяв $i = 1$), получим приближенный корень близкий к нулю $x_1 = \frac{x^2}{1 + 2x} \cong x^2 + O(x^3)$ и при $x \in U_\varepsilon(-1)$ корень $x_1 = \frac{x^2}{1 + 2x} \cong -1 - (1 + x)^2 + O((1 + x)^2)$, близкий к минус единице. При $i = p = 3$ соответственно имеем $x_3 = \frac{4x^4}{(1 + 2x)^7} + O(x^5) \cong 4x^4 + O(x^5)$, $x_3 = -1 + \frac{4(x + 1)^4}{(1 + 2x)^7} + O((1 + x)^5) \cong -1 - 4(1 + x)^4 + O((1 + x)^5)$.

В скалярном случае (когда $n = 1$) из (7) и (8) получаем метод П.Л. Чебышева [1] вычисления корней уравнений. В частности, при $p = 4$, удерживая в (8) первые четыре члена, получим

$$x_{i+1} = x_i - N_1(x_i) - \frac{1}{2}[N_1(x_i)]^2\frac{F^{(2)}(x_i)}{F'(x_i)} -$$
$$- \frac{[N_1(x_i)]^3}{2F'(x_i)}\left[\frac{[F^{(2)}(x_i)]^2}{F'(x_i)} - \frac{1}{3}F^{(3)}(x_i)\right] -$$
$$- \frac{5}{8}\frac{[N_1(x_i)]^4}{F'(x_i)}\left[\frac{[F^{(2)}(x_i)]^3}{[F'(x_i)]^2} -\right.$$
$$\left. - \frac{2}{3}\frac{F^{(2)}(x_i)F^{(3)}(x_i)}{F'(x_i)} + F^{(4)}(x_i)\right].$$

Здесь погрешности $|F(x_4(x))|$ и $|x_{i+1} - x_*|$ являются величинами пятого порядка малости. Ряд по степеням $N_1(x_i)$ (ряд Чебышева) сходится согласно критерию Коши.

Аналогично тому, как это делалось в [5], можно показать, что формулы (7) и (8) дают возможность решить задачи нелинейного программирования с погрешностью p-го порядка малости.

Заметим, что формулу (8) можно получить из разложения Тейлора обратной функции $x = \Phi(y)$ в точке $y_* = F(x_*) = 0$, т.е.

$$x_p(x) = x - (\Phi(y))_y(y - y_*) + \ldots$$

$$\ldots + (-1)^p(\Phi(y))_y^{(p)}\frac{[y - y_*]^p}{p!} = x - (\Phi(y))_y(y) + \ldots$$

$$\ldots + (-1)^p(\Phi(y))_y^{(p)}\frac{[y]^p}{p!} = x - (\Phi(F(x)))_yF(x) + \ldots$$

$$\ldots + (-1)^p(\Phi(F(x)))_y^{(p)}\frac{[F(x)]^p}{p!}.$$

Этот подход использовался П.Л. Чебышевым в его студенческой работе [1].

Теорема 1 может быть обобщена на случай, когда отображение F: $\mathbb{R}^n \to \mathbb{R}^m$ и $m < n$. Предполагаем, что не пусто множество решений X следующей системы из m уравнений с n неизвестными:

$$F(x) = 0_m. \qquad (10)$$

В данном случае матрица Якоби F' – прямоугольная матрица размера $m \times n$ ранга m. Введем невырожденную матрицу Грамма $G = F'[F']^\top$. Для матрицы F' псевдообратной (правой обратной) будет $(n \times m)$-матрица $[F']^+ = [F']^\top G^{-1}$, так как

$$F'(x)[F'(x)]^+ = I_{mm}. \qquad (11)$$

Квадратная $(n \times n)$-матрица $K = [F']^+F'$ есть матрица проектирования на n-мерное пространство строк матрицы F', так как $K[F']^\top = [F']^\top G^{-1}F'[F']^\top = [F']^\top$.

Пусть $\pi(x)$ обозначает проекцию вектора x на множество X, т.е.

$$\pi(x) = \underset{y \in X}{\mathrm{Argmin}}\|y - x\|^2. \qquad (12)$$

Введем вектор множителей Лагранжа $\lambda \in \mathbb{R}^m$ и функцию Лагранжа

$$L(y) = \frac{1}{2}\|y - x\|^2 - \lambda^\mathrm{T}F(y).$$

Пусть минимум по y в задаче (12) достигается при $y = \pi(x)$. Необходимые условия минимума имеют вид

$$\begin{gathered} F(\pi(x)) = 0_m, \\ L_y(\pi(x), \lambda) = \pi(x) - x - [F'(\pi(x))]^\mathrm{T}\lambda = 0_n. \end{gathered} \qquad (13)$$

Умножим последнее выражение слева на матрицу $F'(\pi(x))$, получим

$$F'(\pi(x))(\pi(x) - x) = G(\pi(x))\lambda.$$

Считаем, что ранг матрицы $F'(x)$ равен m; тогда система (13) имеет единственное решение

$$\lambda = [G(\pi(x))]^{-1}F'(\pi(x))(\pi(x) - x).$$

Отсюда и из (13) следует, что

$$[I_{nn} - [F'(\pi(x))]^+F'(\pi(x))](x - \pi(x)) = 0_n, \qquad (14)$$

где I_{nn} – единичная $(n \times n)$-матрица.

Из (14) следует, что вектор $x - \pi(x)$ принадлежит пространству образов матрицы $[F'(\pi(x))]^\top$ и его проекция на нуль-пространство (ядро) равна нулю.

В системе (10), в отличие от (1), может существовать неединственное решение, поэтому в качестве невязки здесь и ниже используем вектор $h(x) = x - \pi(x)$, определяющий расстояние между точкой x и ее проекцией на множество решений системы (10). Считаем, что $\|h(x)\|$ – величина первого порядка малости, $\pi(x) \in X$. В случае $m < k$ анализ проведем по схеме, близкой к приведенной выше. Разложение в ряд Тейлора (2) в данном случае запишем в виде

$$F(\pi(x)) = 0_m = F(x) - F'(x)h + \frac{1}{2}F^{(2)}(x)[h]^2 + \ldots$$

$$\ldots + \frac{(-1)^p}{p!}F^{(p)}(x)[h]^p + O_m(\|h\|^{p+1}). \qquad (15)$$

Умножим равенство (15) слева на $[F'(x)]^+$, обозначим $\bar{N}_1(x) = [F'(x)]^+F(x)$, тогда

$$\bar{N}_1(x) = [F'(x)]^+F'(x)h - [F'(x)]^+ \times$$

$$\times\left[\frac{1}{2}F^{(2)}(x)[h]^2 + \ldots + \frac{(-1)^p}{p!}F^{(p)}(x)[h]^p\right] + O_m(\|h\|^{p+1}). \qquad (16)$$

Отбрасывая члены второго порядка малости, из (16) получим

$$\bar{N}_1(x) = [F'(x)]^+F'(x)h(x) + O_n(\|h\|^2) =$$

$$= [F'(\pi(x))]^+F'(\pi(x))h(x) + O_n(\|h(x)\|^2). \qquad (17)$$

Здесь использовалась теорема о малом возмущении псевдообратного оператора [7–9]. В первом приближении в качестве решения системы (10) возьмем $x_1(x) = x - \bar{N}_1(x)$. Тогда с учетом (11), (14) и (17) имеем

$$x_1(x) - \pi(x) = h(x) - \bar{N}_1(x) = h(x) -$$

$$- [F'(\pi(x))]^+F'(\pi(x))h(x) + O_n(\|h(x)\|^2) \approx$$

$$\approx O_n(\|h(x)\|^2),$$

$$F(x_1(x)) = F(x) - F'(x)\bar{N}_1(x) +$$

$$+ O_m(\|h(x)\|^2) \cong O_m(\|h(x)\|^2).$$

Аналогично во втором приближении определим

$$x_2(x) = x - \bar{N}_2(x),$$

$$\bar{N}_2(x) = \bar{N}_1(x) + \frac{1}{2}[F'(x)]^+F^{(2)}(x)[\bar{N}_1(x)]^2.$$

Из разложения в ряд Тейлора следует, что $F(x_2(x)) \approx O_m(\|x - \pi(x)\|^3)$. Аналогично для произвольного $p \geq 2$ получим

$$\bar{N}_p(x) = \bar{N}_1(x) + [F'(x)]^+ \times$$

$$\times \left\{ \sum_{k=2}^{p} \frac{(-1)^k}{k!} F^{(k)}(x)[\bar{N}_{p-1}(x)]^k \right\}, \quad (18)$$

$$x_p(x) = x - \bar{N}_p(x).$$

Несложной проверкой, используя теорему Тейлора, получаем, что

$$\|F(x_p(x))\| \approx O(\|x - \pi(x)\|^{p+1}). \quad (19)$$

Здесь не гарантируется, что $\|x_p(x) - \pi(x)\| \approx O(\|x - \pi(x)\|^{p+1})$.

Вместе с тем нетрудно показать, см. [8], что

$$\|x_p(x) - x_*\| \leq c\|x - x_*\|, \quad c > 0.$$

Обозначим $x_k = x_k(x)$, $k = 1, 2, ..., p$, причем из (19) имеем

$$\|F(x_p)\| \leq c_1\|x - \pi(x)\|^{p+1}.$$

Используя теорему Люстерника [6, 8], получим

$$\|x_p - \pi(x_p)\| = \rho(x_p, X) \leq \bar{c}\|F(x_p)\| \leq \tilde{c}\|x - \pi(x)\|^{p+1},$$

$$\tilde{c} = \bar{c}c_1 > 0,$$

где $\rho(x_p, X)$ – расстояние от точки x_p до множества X.

Имеет место следующая

Теорема 2. *Пусть $F(x) \in C^{p+1}(\mathbb{R}^n, \mathbb{R}^m)$, x_* – решение системы* (10), *ранг матрицы $F'(x_*)$ равен $m \leq n$ и $x \in U_\varepsilon(x_*)$, где $\varepsilon > 0$ достаточно малое. Тогда для векторов $x_p(x)$, вычисленных по формулам* (18), *имеет место оценка*

$$\|x_p - \pi(x_p)\| \leq C\|x - \pi(x)\|^{p+1},$$

$C > 0$ – независимая константа.

Отметим, что в случае, когда $p = 1$, $m = n$, метод (18) совпадает с методом Ньютона, так как тогда $[F'(x_i)]^+ = [F'(x_i)]^{-1}$, а при $m < n$ получаем $x_1 = x - [F'(x)]^+F(x)$, т.е. приходим к методу Ньютона–Гаусса.

Работа выполнена при финансовой поддержке РФФИ (грант 14–07–00805), Программы поддержки ведущих научных школ НШ-4640.2014.1 и программы Президиума РАН П-18.

СПИСОК ЛИТЕРАТУРЫ

1. *Чебышев П.Л.* Собрание сочинение. М.; Л.: Изд-во АН СССР, 1951. Т. 5. С. 7–25.
2. *Третьяков А.А.* // ЖВМиМФ. 1984. Т. 24. № 7. С. 986–992.
3. *Денисов А.А., Карманов В.Г., Третьяков А.А.* // ДАН. 1985. Т. 281. № 6. С. 1293–1297.
4. *Нечепуренко М.И.* // УМН. 1954. Т. 9. В. 2(60).
5. *Евтушенко Ю.Г., Третьяков А.А.* // ДАН. 2012. Т. 442. № 2. С. 156–161.
6. *Алексеев В.М., Тихомиров В.М., Фомин С.В.* Оптимальное управление. М.: Физматгиз, 1979.
7. *Stewart G.W., Sun J.* Matrix Perturbation Theory, Computer Science and Scientific Computing. N.Y.: Acad. Press, 1990.
8. *Измаилов А.Ф., Третьяков А.А.* Факторанализ нелинейных отображений. М.: Наука, 1994.
9. *Prusinska A., Tret'yakov A.A.* // Set-Valued Anal. 2008. V. 16. P. 93–104.