

改进的布洛依登算法

李福祥, 黄佳玥

(哈尔滨理工大学 理学院, 黑龙江 哈尔滨 150080)

摘要: 提出了一种改进的布洛依登算法, 证明了新方法的收敛性, 通过进行数值实验验证了改进布洛依登算法的收敛阶数及有效性, 并通过与牛顿法以及拟牛顿法的比较体现了改进的布洛依登算法的优越性。

关键词: 非线性方程组; 拟牛顿迭代法; 改进拟牛顿迭代法

DOI: 10.15938/j.jhust.2017.06.024

中图分类号: O22 文献标志码: A 文章编号: 1007-2683(2017)06-0127-04

A Modified Broyden Algorithm

LI Fu-xiang, HUANG Jia-yue

(School of Sciences, Harbin University of Science and Technology, Harbin 150080, China)

Abstract: A Modified Broyden algorithm is presented to solve nonlinear equations in this paper. The convergence of the new algorithm is proved. The convergent order and effectiveness of improved Broyden method can be verified by numerical experiments. Through comparing with Newton's method and quasi-newton method, the advantages of the improvement Broyden method are showed.

Keywords: nonlinear equations; quasi Newton iteration method; improved quasi newton iteration method

0 引言

求解非线性方程组的牛顿法是一个最基本而且十分重要的方法, 目前很多有效的迭代法都是以牛顿法为基础改进的。求解非线性方程组的牛顿法是求解非线性方程的牛顿法的推广^[1-7]。

设非线性方程

$$f(x) = 0 \quad (1)$$

若 x^* 是式(1)的根, x^0 是 x^* 的近似值, 将函数 $f(x)$ 在点 x^0 处泰勒展开, 即

$$f(x) \approx f(x^0) + f'(x^0)(x - x^0)$$

于是式(1)可以表示为

$$f(x^0) + f'(x^0)(x - x^0) = 0$$

记其根为 x^1 , 则有

$$x^1 = x^0 - \frac{f(x^0)}{f'(x^0)}$$

进行推广迭代则有

$$x^{k+1} = x^k - \frac{f(x^k)}{f'(x^k)}$$

推广到 n 维空间中则有

$$x^{k+1} = x^k - [F'(x^k)]^{-1} F(x^k) \quad (2)$$

式(2)即为非线性方程组的牛顿迭代法, 此方法二阶收敛, 具体证明见文[8]。

收稿日期: 2017-09-02

基金项目: 黑龙江省教育厅科学技术研究项目(11521045); 黑龙江省自然科学基金(A200811)。

作者简介: 李福祥(1972—), 男, 教授, E-mail: lifx2013@163.com;

黄佳玥(1996—), 女, 硕士研究生。

1 拟牛顿迭代法

由于牛顿法中 $F(x^k)$ 的计算量非常大,人们希望通过较小的计算量构造矩阵 A_k 来代替 $F(x^k)$ 要求在拟牛顿算法中,并通过 A_k 构造的 A_{k+1} 满足拟牛顿条件或拟牛顿方程: $A_{k+1}s^k = y^k$, 其中 $A_{k+1} = A_k + \Delta A_k$, $s^k = x^{k+1} - x^k$, $y^k = F(x^{k+1}) - F(x^k)$, 拟牛顿法的迭代形式为:

$$\begin{cases} x^{k+1} = x^k - A_k^{-1} F(x^k), k=0, 1, L \\ A_{k+1}(x^{k+1} - x^k) = F(x^{k+1}) - F(x^k) \\ A_{k+1} = A_k + \Delta A_k, \text{rank}(\Delta A_k) = m \geq 1 \end{cases} \quad (3)$$

选取不同的 ΔA_k 即对应着不同的拟牛顿方法。例如,如果规定 A_{k+1} 与 A_k 在某向量 $c \in R^n$ 的正交补 $c^\perp$ 上两者无任何差别,也即 $A_{k+1}z = A_kz, \forall z \in c^\perp$, 这时可令 $\Delta A = A_{k+1} - A_k = uc^T$, 由

$$\begin{aligned} (A_{k+1} - A_k)s^k &= y^k - A_k s^k = \\ &= F(x^{k+1}) - F(x^k) - \\ &= [-F(x^k)] = \\ &= F(x^{k+1}) \end{aligned}$$

推得 $u = \frac{F(x^{k+1})}{c^T s^k}$ 。于是得到秩 1 修正公式^[8]

$$A_{k+1} = A_k + \frac{F(x^{k+1})c^T}{c^T s^k} \quad (4)$$

下面介绍一种比较重要的拟牛顿迭代法。

2 布洛依登算法

布洛依登首先在 1965 年对拟牛顿法提出一种修正 A_k 的计算方案(在式(4)中取 $c = s^k$):

$$\begin{cases} x^{k+1} = x^k - A_k^{-1} F(x^k) \\ A_{k+1} = A_k + \frac{(y^k - A_k s^k)(s^k)^T}{(s^k)^T s^k} \end{cases} \quad (5)$$

迭代法式(5)中的第 2 个公式叫做布洛依登秩 1 校正公式。它有一个重要性质,即在所有满足拟牛顿方程的矩阵中 A_{k+1} 是在 F -范数意义下最接近于 A_k 的矩阵^[8]。在第一次迭代时,其中的 A_0 通常用 $F'(x^0)$ 代替。

拟牛顿算法与牛顿法相比,拟牛顿法降低了导数值的运算量,收敛速度比牛顿法快,但每一步迭代都要计算新的矩阵的逆矩阵,在一定程度上也会增加计算的难度,改进拟牛顿算法就是基于这个缺点改进的,通过改进布洛依登秩 1 校正公式,降低逆矩

阵的计算量。

3 改进的布洛依登算法

由于布洛依登算法每次需要计算矩阵 A_k 的逆或解线性方程组

$$A_{k+1}(x^{k+1} - x^k) = F(x^{k+1}) - F(x^k)$$

这给布洛依登算法的应用带来不便。本文在公式

$$A_{k+1} = A_k + \frac{(y^k - A_k s^k)(s^k)^T}{(s^k)^T s^k}$$

中令

$$A_{k+1}^{-1} = A_k^{-1} + \frac{(s^k - A_k^{-1} y^k)(y^k)^T}{(y^k)^T y^k}$$

即得改进布洛依登算法的迭代格式为

$$\begin{cases} x^{k+1} = x^k - A_k^{-1} F(x^k) \\ A_{k+1}^{-1} = A_k^{-1} + \frac{(s^k - A_k^{-1} y^k)(y^k)^T}{(y^k)^T y^k} \end{cases} \quad (6)$$

用这种方式改进的布洛依登算法,在得到通过 x^0 迭代出的 x^1 后,进行再迭代,计算 x^2 时,不需像拟牛顿法一样再计算 A_{k+1} 的逆。即每一步迭代只需计算一个矩阵的逆。

4 收敛性分析

引理 1^[9] (Sherman-Morrison) 设矩阵 $A \in L(R^n)$ 非奇异, $u, v \in R^n$, 当 $\vartheta = 1 + v^T A^{-1} u \neq 0$ 时, 则 $A + uv^T$ 非奇异, 且有

$$(A + uv^T)^{-1} = A^{-1} - \frac{1}{\vartheta} (A^{-1} uv^T A^{-1}) \quad (7)$$

引理 2^[9] 若 $x^* \in D_0$ 是 $F(x) = 0$ 的解。 $F: D \subset R^n \rightarrow R^n$ 在开凸集 $D_0 \subset D$ 上连续可导, 且 $F'(x^*)$ 非奇异, 条件

$$\|F'(x) - F'(x^*)\| \leq L \|x - x^*\| \quad (8)$$

成立, $U(x, A)$ 是定义在 $D_0 \times D_M$ 上 F 的修正函数, $\bar{A} \in U(x, A)$ 。若存在 $\alpha_1 > 0$ 及 $\alpha_2 > 0$ 使

$$\begin{aligned} &\|\bar{A} - F'(x^*)\| \\ &\leq [1 + \alpha_1 \sigma(x, \bar{x})] \|A - F'(x^*)\| + \alpha_2 \sigma(x, \bar{x}) \end{aligned} \quad (9)$$

其 $\bar{x} = x - A^{-1} F(x)$, $\sigma(x, \bar{x}) = \max\{\|x - x^*\|_2, \|\bar{x} - x^*\|_2\}$, $\bar{A}$ 是由 A 得到的新近似, 则对每个 $\gamma \in (0, 1)$, $\exists \varepsilon(\gamma) > 0, \delta(\gamma) > 0$, 使当 $x^0 \in D_0, A_0 \in D_M$ 满足

$$\|x^0 - x^*\| \leq \varepsilon(\gamma), \|A_0 - F'(x^*)\| \leq \delta(\gamma)$$

时,拟牛顿法(4)生成的序列 $\{x^k\}$ 是适定的且收敛于 x^* ,并有

$$\|x^{k+1} - x^*\|_2 \leq \gamma \|x^k - x^*\|_2, k=0,1,\dots,$$

序列 $\{\|A_k\|\}$ 及 $\{\|A_k^{-1}\|\}$ 一致有界。

定理1 设 $F: D \subset R^n \rightarrow R^n$ 在 x^* 凸领域 $D_0 \subset D$ 内连续可导, $F(x^*)=0$,且 $F'(x^*)$ 非奇异,条件(8)成立,则改进拟牛顿方法(6)生成的序列 $\{x^k\}$ 是适定且超线性收敛 x^* 。

证明:先证明(6)生成的序列 $\{x^k\}$ 适定且收敛于 x^* ,这只需验证条件(9)成立即可,由引理1可得:

$$\begin{aligned} A_{k+1} - F'(x^*) &= \\ \left[A_k^{-1} + \frac{(s^k - A_k^{-1}y^k)(y^k)^T}{(y^k)^T y^k} \right]^{-1} - F'(x^*) &= \\ [A_k - F'(x^*)] - \frac{1}{\vartheta} \left(A_k \frac{(s^k - A_k^{-1}y^k)(y^k)^T}{(y^k)^T y^k} A_k \right) &= \\ [A_k - F'(x^*)] - \frac{1}{\vartheta} \left(\frac{(A_k s^k - y^k)(y^k)^T A_k}{(y^k)^T y^k} \right) \end{aligned}$$

设 $(y^k)^T y^k = a$,上式两端取模,并取 $\frac{1}{\vartheta}$ 的值为 M ,整理上式为:

$$\begin{aligned} \|A_{k+1} - F'(x^*)\|_F &\leq \\ \|A_k - F'(x^*)\|_F + \\ |Ma^{-1}| \|(A_k s^k - y^k)(y^k)^T A_k\|_2 &\leq \\ \|A_k - F'(x^*)\|_F + |Ma^{-1/2}| \|A_k s^k - y^k\|_2 \\ \|A_k\|_2 &\leq \\ \|A_k - F'(x^*)\|_F + |Ma^{-1/2}| \|A_k\|_2 \\ \|A_k s^k - y^k + F'(x^*)s^k - F'(x^*)s^k\|_2 &\leq \\ \|A_k - F'(x^*)\|_F + B \|(A_k - F'(x^*))s^k\|_2 + \\ B \|y^k - F'(x^*)s^k\|_2 &\leq \\ \|A_k - F'(x^*)\|_F (1 + B\sigma) + BL\sigma \end{aligned}$$

其中 $B = |M((y^k)^T y^k)^{-1/2}| \|A_k\|_2$,这表明条件(9): $\alpha_1 = B, \alpha_2 = BL$ 时成立,于是引理2条件全部满足。故(6)生成的序列 $\{x^k\}$ 收敛于 x^* 。

5 数值实验

本节共选择了三个非线性方程组进行数值实验,其中一个为二元方程组,一个三元方程组;一个四元方程组。

例1 解非线性方程组

$$\begin{cases} x^3 - 3x^2 + 8x + 3y - 36 = 0 \\ x^2 + y - 4 = 0 \end{cases}$$

真解为 $X^* = [6, -40]^T$ 。

例2 解非线性方程组

$$\begin{cases} x^2 + y^2 + z^2 - 14 = 0 \\ xyz - 6 = 0 \\ x + y - z^2 = 0 \end{cases}$$

真解为 $X^* = [3, 1, 2]^T$ 。

例3 解非线性方程组

$$\begin{cases} x^3 + y^2 - 3xz^2 + 4t - 18 = 0 \\ xy - yt + 2xt + z = 0 \\ 6xyz - 3zt - 6 = 0 \\ x^2y + z - t = 0 \end{cases}$$

真解为 $X^* = [1, 3, 2, 5]^T$ 。

利用 Mathematica 软件进行编程计算,分别用牛顿法、拟牛顿法和改进拟牛顿法对上面的六个方程组进行迭代求解,以真解 X^* 与最后一次迭代得到的 $\bar{X}$ 之间的差 $\varepsilon = X^* - \bar{X}$ 为指标进行比较。

由于计算机不能显示出过小的数,因此分别选取迭代次数为2、3、4。为了使表格简洁明了,在不影响结果真实性、准确性的前提下,表格中只记录了 ε 的指数。

计算例1可得比较结果如表1;例2的结果比较见表2;例3的结果比较见表3。

表1 例1数值结果

迭代次数	2	3	4
牛顿迭代法	-26	-51	-103
拟牛顿迭代法	-58	-175	-523
改拟牛顿迭代法	-59	-176	-528

表2 例2数值结果

迭代次数	2	3	4
牛顿迭代法	-11	-21	-42
拟牛顿迭代法	-22	-66	-199
改拟牛顿迭代法	-25	-78	-222

表3 例3数值结果

迭代次数	2	3	4
牛顿迭代法	-6	-11	-22
拟牛顿迭代法	-12	-35	-104
改拟牛顿迭代法	-12	-35	-104

由表1、表2、表3可以看出,在收敛阶数方面:牛顿迭代法为2阶收敛;拟牛顿迭代法与改拟牛顿迭代法均为3阶收敛。从效率指数方面看:牛顿迭

代法的效率指数为 $2^{\frac{1}{2n}}$, 拟牛顿迭代法的效率指数为 $3^{\frac{1}{3n}}$, 改拟牛顿迭代法的效率指数均为 $3^{\frac{1}{3n}}$, 即拟牛顿迭代法与改拟牛顿迭代法的计算效率高于牛顿迭代法, 而改拟牛顿迭代法的计算效率与拟牛顿迭代法的计算效率是一样的。但是, 改进拟牛顿算法比拟牛顿算法少算了一个矩阵的逆, 因此对拟牛顿迭代法的修改是有一定意义的。

6 结 论

本文主要介绍了改进拟牛顿迭代法的形式, 证明了该方法的收敛性, 利用数值实验证明了改进拟牛顿迭代法的收敛阶数为 3 阶, 进行了三种迭代法的比较, 说明了改进拟牛顿迭代法的有效性。

参 考 文 献:

- [1] HALLEY E. A New. Exact and Easy Method for Finding the Roots of Equations Generally and without Any Previous Reduction [J]. Philos. Trans. R. Soc. Lond., 1694(18): 136–148.

- [2] KOU J, LI Y. Modified Chebyshev's Method Free from Second Derivative for Nonlinear Equations [J]. J. Appl. Math. Comput., 2007, 187(2): 1027–1032.
- [3] GUTIERREZ J M, HERNANDEZ M A. An Acceleration of Newton's Method: Super Halley Method [J]. J. Appl. Math. Comput., 2001, 117(2): 223–239.
- [4] KING R F. A Family of Fourth Order Methods for Nonlinear Equations [J]. SI AMJ. Numer. Anal., 1973(10): 876–879.
- [5] LIU Z, ZHENG Q, ZHAO P. A Variant of Ste Ensens Method of Fourth-order Convergence and Its Applications [J]. Applied Mathematics and Computation, 2010, 216(7): 1978–1983.
- [6] OSTROWSKI A M. Solutions of Equations and Systems of Equations [M]. New York: Academic Press, 1966.
- [7] KUNG H T, TRAUB J F. Optimal Order of One-point and Multi-point Iteration [J]. J. Assoc. Comput. Mach., 1974(21): 643–651.
- [8] 黄象鼎, 曾钟钢, 马亚南. 非线性数值分析的理论与方法 [M]. 武汉: 武汉大学出版社, 2004.
- [9] 李庆杨, 莫孜中, 祁力群. 非线性方程组的数值解法 [M]. 北京: 科学出版社, 1999.

(编辑: 关 毅)

(上接第 126 页)

- [10] KIRK W A, SIMS B. Handbook of Metric Fixed Point Theory [M]. Dordrecht: Kluwer Academic Publishers, 2001.
- [11] 左占飞, 崔云安. 广义凸性模在不动点中的应用 [J]. 黑龙江大学自然科学学报, 2009, 26(2): 206–210.
- [12] 王俊明, 陈丽丽, 崔云安. 平均非扩张映射的不动点性质 [J]. 黑龙江大学自然科学学报, 2006, 23(3): 298–301.
- [13] CUI Y A, HUDZIK H, WISLA M. M-Constant, Dominguez-Benavides Coficiet, and Weak Fixed Point Property in Orlicz Sequence Spaces Equipped with Thep-Amemiya Norm [J]. Fixed Point Theory and Applications, 2016, 89: 1–14.
- [14] 左占飞, 王良伟, 刘学飞, 等. Banach 空间中弱收敛序列系数的估计 [J]. 数学学报(中文版), 2016, 59(2): 145–150.
- [15] GOEBEL K, KIRK W A. Topics in Metric Fixed Point Theory, in: Cambridge Studies in Advanced Mathematics 28 [M]. New

York: Cambridge University Press, 1990.

- [16] SIMS B. A Class of Spaces with Weak Normal Structure [J]. Bull. Austral. Ma th. Soc., 1994(50): 523–528.
- [17] CHEN S T. Geometry of Orlicz Spaces [M]. Warszawa: Dissertations Math., 1996.
- [18] RAO M M, REN Z D. Theory of Orlicz Spaces [M]. New York: Marcel Dekker, 1991.
- [19] KRASNOSELSKII M A, RUTICKII Y B. Convex Function and Orlicz Spaces [M]. Groningen: P. Noordhoff, 1961.
- [20] MALLGRANDA L. Orlicz Spaces and Interpolation, in: Seminars in Mathematics 5 [M]. Sao Paulo: Campinas Univ., 1989.

(编辑: 温泽宇)