第19卷,第1期
2012年3月
中国传媒大学学报自然科学版
JOURNAL OF COMMUNICATION UNIVERSITY OF CHINA (SCIENCE AND TECHNOLOGY)
Vol.19,No.1
Mar,2012

调整步长牛顿法

刘停战,刘伟,何颖
(中国传媒大学 理学院,北京 100024)

摘要:本文研究了求解非线性方程组的迭代解法,提出了一种调整步长牛顿法。证明了该算法在不同条件下的二阶收敛性和大范围收敛性。

关键词:非线性方程组;牛顿法;调整步长牛顿法
中图分类号:O241.7 文献标识码:A 文章编号:1673-4793(2012)01-0008-03

Step-adjusting Newton Method

LIU Ting-zhan,LIU Wei,HE Ying
(School of Science,Communication University of China,Beijing 100024,China)

Abstract: In this paper,we studied iterative method for solving nonlinear equations and obtained step-adjusting Newton method. Second-order convergence and global convergence are also proved in different conditions.

Keywords: nonlinear equations; Newton method; step-adjusting Newton method

1 引言

设 F 是实的或复的高维 Banach 空间上的某个凸子集 Ω 到同型空间 S 上的非线性算子,考虑求方程组

$$F(x)=0,x\in\Omega \tag{1}$$

的解,其中 $F(x)=(f_1(x),\cdots,f_n(x))$。我们知道在迭代法中,牛顿法和牛顿下山法最具代表性,牛顿法有二阶收敛性,牛顿下山法有大范围收敛性。牛顿法和牛顿下山法的迭代格式分别为:

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

其中 $k=0,1,\cdots$。

2 调整步长牛顿法

我们构造方程组(1)的等价方程组

$$F_\alpha(x)=(f_1(x)^{\alpha_1},\cdots,f_n(x)^{\alpha_n})^T=0 \tag{2}$$

其中 $\alpha_i>0(i=1,2,\cdots,n)$。

对(2)式使用牛顿法,得到牛顿迭代格式:

$$x^{(k+1)}=x^{(k)}-[F_\alpha'(x^{(k)})]^{-1}F_\alpha(x^{(k)}),k=0,1,\cdots \tag{3}$$

由

$$F_\alpha(x)=(f_1(x)^{\alpha_1},\cdots,f_n(x)^{\alpha_n})^T$$
$$=\mathrm{diag}(f_1(x)^{\alpha_1-1},\cdots,f_n(x)^{\alpha_n-1})F(x),$$
$$F_\alpha'(x)=\mathrm{diag}(\alpha_1,\alpha_2,\cdots,\alpha_n)\,\mathrm{diag}(f_1(x)^{\alpha_1-1},f_2(x)^{\alpha_2-1},\cdots,f_n(x)^{\alpha_n-1})F'(x),$$
$$[F_\alpha'(x)]^{-1}=[F'(x)]^{-1}\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)[D_\alpha(x)]^{-1},$$

把(3)式化简为:

$$x^{(k+1)}=x^{(k)}-[F'(x^{(k)})]^{-1}\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)F(x^{(k)}),k=0,1,2,\cdots \tag{4}$$

收稿日期:2011-07-12
作者简介:刘停战(1954-),男(汉族),吉林长春人,中国传媒大学理学院教授. E-mail:tzliu@cuc.edu.cn.

其中 $\lambda_i = \dfrac{1}{\alpha_i}(i = 1,2,\cdots,n)$，对角矩阵 $\mathrm{diag}(\alpha_1,\alpha_2,\cdots,\alpha_n)$ 称作为步长矩阵，所对应的算法称为调整步长牛顿算法，称(4)式为调整步长牛顿法的迭代格式。

注　该算法是牛顿下山法的推广。当 $0 < \lambda_1 = \lambda_2 = \cdots = \lambda_n \le 1$ 时，调整步长牛顿法就简化为牛顿下山法。当 $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 1$ 时，调整步长牛顿法即为牛顿法。

3　调整步长牛顿法的收敛性

关于调整步长牛顿法的收敛性及收敛阶，我们有:

定理 1　设 $F_\alpha: \Omega \subset R^n \to R^n$ 及初始迭代点 $x^{(0)}$ 满足下列条件:

1) $[F'(x^{(0)})]^{-1}[D_\alpha(x^{(0)})]^{-1}$ 存在，且

$$\|[F'(x^{(0)})]^{-1}[D_\alpha(x^{(0)})]^{-1}\| \le \beta,$$

$$\|F'(x^{(0)})^{-1}F(x^{(0)})\| \le \eta,$$

2) 于 $x^{(0)}$ 的邻域 $S(x^{(0)},\delta) \subset \Omega$ 内，$D_\alpha(x)F'(x)$ 存在且满足 Lipschitz 条件: $\|D_\alpha(x)F'(x) - D_\alpha(y)F'(y)\| \le \kappa\|x-y\|$，$\forall x,y \in S(x^{(0)},\delta)$，并且

$$\rho = \kappa\beta\eta \le \frac{1}{2},$$

$$\delta \ge \frac{1-\sqrt{1-2\rho}}{\rho}\eta,$$

则方程组(1)于 $S(x^{(0)},\delta)$ 有解 x^* 存在，且(4)式收敛于 x^*。当 $\rho < \dfrac{1}{2}$ 时，调整步长牛顿法二阶收敛。

证明　只要说明本定理满足文献[2]的 Kantorovich 定理即可。由 $[F'(x^{(0)})]^{-1}[D_\alpha(x^{(0)})]^{-1}$ 存在及关系式

$$[F'_\alpha(x)]^{-1} = [F'(x)]^{-1}\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)[D_\alpha(x)]^{-1},$$

知 $[F'_\alpha(x^{(0)})]^{-1}$ 存在。再由 $\|[F'(x^{(0)})]^{-1}[D_\alpha(x^{(0)})]^{-1}\| \le \beta$，可以得到

$$\|[F'_\alpha(x^{(0)})]^{-1}\| = \|[F'(x^{(0)})]^{-1}\mathrm{diag}(\lambda_1,\lambda_2,\cdots\lambda_n)[D_\alpha(x^{(0)})]^{-1}\| \le \|\mathrm{diag}(\lambda_1,\lambda_2,\cdots\lambda_n)\|\,\|[F'(x^{(0)})]^{-1}[D_\alpha(x^{(0)})]^{-1}\| \le \beta,$$

上述的推导过程需要用到 $0 < \lambda_i \le 1(i = 1,\cdots,n)$，且取 $\|\mathrm{diag}(\lambda_1,\lambda_2,\cdots\lambda_n)\| = \|\mathrm{diag}(\lambda_1,\lambda_2,\cdots\lambda_n)\|_1$。同样可以得到 $\|[F'_\alpha(x^{(0)})]^{-1}F'_\alpha(x^{(0)})\| \le \eta$。由 $F'_\alpha(x) = \mathrm{diag}(\alpha_1,\alpha_2,\cdots,\alpha_n)D_\alpha(x)F'(x)$ 和 $\|D_\alpha(x)F'(x) - D_\alpha(y)F'(y)\| \le \kappa\|x-y\|$ 可以得到，

$$\|F'_\alpha(x) - F'_\alpha(y)\| \le \kappa\|x-y\|,\ \forall x,y \in S(x^{(0)},\delta),$$

由以上可知，满足 Kantorovich 定理的条件，所以结论成立。

定理 1　给出了调整步长牛顿法的半局部收敛性，下面讨论调整步长牛顿法的大范围收敛性。

定理 2　设 $F: \Omega \subset R^n \to R^n$ 满足下列条件:

1) $\|F(x^{(0)})\| \le \eta$，

2) 于区域 $\Omega_0 = \{x \mid \|x - x^{(0)}\| \le \gamma\beta\eta\} \subset \Omega$，$F'$ 有逆存在，且

$$\|F'(x)^{-1}\| \le \beta,\ \forall x \in \Omega_0, \tag{5}$$

$$\|F'(x) - F'(y)\| \le \kappa\|x-y\|,\ \forall x,y \in \Omega_0, \tag{6}$$

则当 $\gamma > \lambda_{\max}^{(0)}/(1/2\rho(\lambda_{\max}^{(0)})^2 + 1 - \lambda_{\max}^{(0)})$，$\lambda_{\min}^{(k)}/(\lambda_{\max}^{(k)})^2 > 1/2\rho > 0$，$0 < \lambda^{(k)} \le 1$，$k = 0,1,\cdots$，方程组 $F(x) = 0$ 于 Ω_0 有解 x^* 存在，其中 $\rho = \kappa\beta^2\eta$，由

$$x^{(k+1)} = x^{(k)} - [F'(x^{(k)})]^{-1}\mathrm{diag}(\lambda_1^{(k)},\lambda_2^{(k)},\cdots,\lambda_n^{(k)})F(x^{(k)}) \tag{7}$$

产生的迭代序列 $\{x^{(k)}\}$ 收敛于 x^*。

证明

$$\|F(x^{(k+1)})\| = \|F(x^{(k+1)}) - \mathrm{diag}(\lambda_1^{(k)},\lambda_2^{(k)},\cdots,\lambda_n^{(k)})F(x^{(k)}) - F'(x^{(k)})(x^{(k+1)} - x^{(k)})\| \le (1/2\|\mathrm{diag}(\lambda_1^{(k)},\lambda_2^{(k)},\cdots,\lambda_n^{(k)})\|^2\beta^2\eta\kappa + \|I - \mathrm{diag}(\lambda_1^{(k)},\lambda_2^{(k)},\cdots,\lambda_n^{(k)})\|)\|F(x^{(k)})\|,$$

令

$\lambda_{\max}^{(k)} = \max(\lambda_1^{(k)},\lambda_2^{(k)},\cdots,\lambda_n^{(k)})$，$\lambda_{\min}^{(k)} = \min(\lambda_1^{(k)},\lambda_2^{(k)},\cdots,\lambda_n^{(k)})$，

$$(1/2\|\mathrm{diag}(\lambda_1^{(k)},\lambda_2^{(k)},\cdots,\lambda_n^{(k)})\|^2\beta^2\eta\kappa + \|I - \mathrm{diag}(\lambda_1^{(k)},\lambda_2^{(k)},\cdots,\lambda_n^{(k)})\|)\|F(x^{(k)})\| = (1/2\rho(\lambda_{\max}^{(k)})^2 + 1 - \lambda_{\min}^{(k)})\|F(x^{(k)})\|,$$

由于

$$0 < 1/2\rho(\lambda_{\max}^{(k)})^2 + 1 - \lambda_{\min}^{(k)} < 1,$$

令 $1/2\rho(\lambda_{\max}^{(k)})^2 + 1 - \lambda_{\min}^{(k)} = \omega_k$。

即可得

$$\|F(x^{(k+1)})\| \le \omega_k\|F(x^{(k)})\|,$$

$x^{(0)}, x^{(1)} \in \Omega_0$ 显然，故当 $k = 0$ 时成立，我们有 $\|F(x^{(1)})\| \le \omega_0\|F(x^{(0)})\|$，由此特别有 $\|F(x^{(1)})\| \le \|F(x^{(0)})\|$，并且利用(7)式

$$\|x^{(2)} - x^{(1)}\| \le \lambda_{\max}^{(1)}\|[F'(x^{(1)})]^{-1}\|\,\|F$$

$(x^{(1)})\| \le \lambda_{\max}^{(1)}\beta\eta\omega_0$

$$\|x^{(2)}-x^{(0)}\| \le \|x^{(2)}-x^{(1)}\| + \|x^{(1)}-x^{(0)}\| \le \beta\eta(\lambda_{\max}^{(0)}+\lambda_{\max}^{(1)}\omega_0) \le \gamma\beta\eta \quad (8)$$

既可以得出 $x^{(2)} \in \Omega_0$。

利用数学归纳法，对 $k=0,1,2,\cdots$，均有 $x^{(k)} \in \Omega_0$，且

$$\|F(x^{(k)})\| \le (1/2\rho(\lambda_{\max}^{(k-1)})^2 + 1 - \lambda_{\max}^{(k-1)})\|F(x^{(k-1)})\| = \omega_{k-1}\|F(x^{k-1})\|,$$

$$\le \prod_{i=0}^{k-1}\omega_i\eta$$

可得

$$\|x^{(k+1)}-x^{(k)}\| \le \prod_{i=0}^{k}\omega_i\eta,$$

于是利用上式立即导出 $x^{(k)}$ 有极限 $x^* \in \Omega_0$ 存在，并注意 $\|[F'(x^{(k)})]^{-1}\| \le \beta$ 以及 λ^k 的有界性。对(7)式令 $k\to\infty$ 导出 $F(x^*)=0$。这样就证明了调整步长牛顿法的大范围收敛性。

4 数值实验

本节将考虑使用上述调整步长牛顿法与牛顿法来计算一个例子，迭代终止条件为 $\|x_k - x_{(k-1)}\| < 10^{-6}$。

例 1

$$\begin{cases} f_1(x) = e^{-0.2x_1} - x_2 = 0 \\ f_2(x) = e^{-x_1} - x_2 + 0.5 = 0 \end{cases},$$

方程组有解 $x^* = (1.3127, 0.7691)^T$ 和 $x^* = (2.9837, 0.5506)^T$。取初值 $x^{(0)} = (202, 300)^T$。其中 $(\lambda_1^{(0)}, \lambda_2^{(0)}) = (0.7, 0.6)^T$，若 $\|F(x^{(k+1)})\| > \|F(x^{(k)})\|$，则取 $\lambda_i^{(k+1)} = \lambda_i^{(k)}/2$，$(i=1,2)$。

表 1

算法名称	迭代步数	$x^{(k)}$	$\|F(x^{(k)})\|$
牛顿法	发散	发散	发散
迭代格式(7)	99	$x^* = (1.3127, 0.7691)^T$	9.4022×10^{-7}

通过表1的计算结果可以看出当初始迭代点 $x^{(0)}$ 距离解较远时，牛顿法发散，调整步长牛顿法却收敛，这就说明了迭代格式(7)具有大范围收敛性。

参考文献

[1] 刘兴龙. 解非线性方程组的一种带参数的 Newton 方法[J]. 哈尔滨工业大学学报，1979(2):97-104.

[2] 冯果忱. 非线性方程组迭代解法[M]. 上海: 上海科学技术出版社，1989.

[3] 卢兴江. 关于解非线性方程组的 Newton 型迭代法的若干研究[J]. 浙江丝绸工学院学报，1998，15(2):141-144.

[4] Ortega J M，Rheinboldt W C. 多元非线性方程组迭代解法[M]. 北京: 科学出版社，1983.

（责任编辑: 宋金宝）