维普资讯 http://www.cqvip.com

2001 年 8 月
第 31 卷第 4 期
西北大学学报(自然科学版)
Journal of Northwest University(Natural Science Edition)
Aug. 2001
Vol. 31 No. 4

求解非线性方程组的信赖域方法

薛西峰，邢志栋，孟红云
(西北大学 数学系，陕西 西安 710069)

摘要：用信赖域方法求解非线性方程组问题，计算中仅使用目标函数及其一阶导数信息，通过 BFGS 校正方法构造 Hessian 阵的近似，最后给出了相应的数值结果。

关 键 词：信赖域方法；Samaskii 技巧；BFGS 校正方法；交替方向法
中图分类号：O241.7 **文献标识码**：A **文章编号**：1000-274X(2001)04-0289-03

考虑非线性方程组

$$F(x) = 0, x \in D \subseteq R^n。 \quad (1)$$

其中：$F(x) = (F_1(x), F_2(x), \cdots, F_n(x))^T$，$F_i: R^n \to R(i = 1, 2, \cdots, n)$ 是连续可微的实值函数。

求解式(1)的方法有多种[1]，熟知以 Newton 法为代表的迭代法及其变形。求解式(1)时，一般都是采用线性搜索技术，寻找搜索方向 d_k 和步长因子 α_k，使得目标函数值在每一步都有所下降，即满足 $\| F(x_k) + \alpha_k d_k \| < \| F(x_k) \|$。这种搜索技术，在实际计算时对初始点 x_0 要求比较严格，且每次都要计算 $F'(x)$，当导数值出现奇异状况或非常小时，使计算无法进行，且收敛性不能保证，因而使方法受到一定的限制。因此，在具有线性搜索的 Newton 方法中，修改 Hessian 阵，以保证搜索方向的下降性，寻求在解的附近使算法具有整体收敛性，是很有意义的工作。信赖域方法既具有 Newton 法的快速收敛性又有理想的总体收敛性，而且可以解决 Hessian 阵不正定和 x_k 为鞍点等困难[2]。有鉴于此，我们对应用信赖域方法求解非线性方程组进行了深入的讨论，并得出了一定的结论。

1 信赖域方法的算法模型

用无约束优化方法求解式(1)时，通常将其转化为求极小问题

$$\min f(x) = \frac{1}{2} F(x)^T F(x)。 \quad (2)$$

信赖域方法的基本思想是：设 x_k 是当前迭代点，在以 x_k 为中心，δ_k 为半径的闭球(或取某框形区域)邻域内求解一子问题，要求试探步 d_k 在信赖域内，即在每次迭代时有一正数 δ_k，且要求试探步满足 $\| d_k \| < \delta_k$，用一评价函数决定该 d_k 是否被接受，若被接受，则令 $x_{k+1} = x_k + d_k$；否则令 $x_{k+1} = x_k$。如果试探成功，信赖域的半径保持不变或扩大，否则，缩小半径。信赖域方法的模型问题是

$$\min q^{(k)}(s) = f(x_k) + g_k^T s + 1/2 s^T G_k s, \quad (3)$$
$$s.t. \ \| s \| \leqslant h_k。$$

而式(3)可以通过求解

$$(G_k + \mu_k I)s = - g_k \quad (4)$$

来表征，其中：二次模型 $q^{(k)}(s)$ 是在 x_k 处对目标函数 $f(x)$ 的一个合适模拟，g_k，G_k 是目标函数 $f(x)$ 在 x_k 处的梯度和 Hessian 阵 μ_k 是一非负数。令

$$R(x) = \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_n}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial x_n} & \cdots & \frac{\partial F_n}{\partial x_n} \end{bmatrix}, g(x) = R(x)' F(x)。 \quad (5)$$

本文将上述信赖域方法用于求解非线性方程组问题，在保持信赖域方法总体收敛的前提下和在具体求解过程中，用目标函数 f 及其一阶导数 g 的信息，构造 Hessian 阵 G_k 的近似序列 B_k，通过 BFGS 校正公式产生目标函数的曲率近似，不仅保持了矩阵序列的正定性，克服了数值上的奇异性，而且方法具有下降性及收敛速度快等优点，其中，g 和 f 的定义

收稿日期：2000-03-20
基金项目：陕西省教委专项基金资助项目(99JK096)；西北大学校内基金资助项目(00NW51)
作者简介：薛西峰(1961-)，男，陕西华县人，西北大学副教授，从事非线性问题的研究。

维普资讯 http://www.cqvip.com

见式(5),式(2),所用的BFGS校正方法如下

$$B_{k+1}^{BFGS} = B_k + \frac{g_k^T + g_k}{g_k d_k} + \frac{y_k^T y_k}{\alpha_k y_k^T y_k}。\quad (6)$$

其中 $s_k = x_{k+1} - x_k,\ y_k = g_{k+1} - g_k,$

$$g_k = g(x_k),g_{k+1} = g(x_{k+1})。\quad (7)$$

对于搜索步长因子 α_k 的确定,为克服在接近极小点时有可能出现迂回现象及步长因子较小的情形,应用交替方向法,记 $\alpha_k = (-g_k,s_k)/(\|-g_k\|\ \|s_k\|)$,若 $\alpha_k < \sqrt{2}/2$ 时,搜索方向取原拟牛顿方向 s_k,否则,搜索方向取最速下降方向 $-g_k$。同时,为了提高算法的效率,还可在每一迭代步中使用Samaskii技巧[1]。

2 收敛性分析

信赖域方法的一个突出优点是具有总体收敛性,下面给出方法的总体收敛性定理及其在较强条件下的二阶收敛结果。以下两个定理的证明可参照文献[2]中定理3.6.3和定理5.1.3。

定理1 (BFGS公式的正定性)当且仅当 $s_k^T y_k > 0$ 时,BFGS校正公式(6)保持正定性。

定理2 (信赖域方法的总体收敛性定理)

假定 $f(x)$ 是二阶强凸可微函数,即存在 $M \geqslant m \geqslant 0$,使对任意 $x,y \in R^n$,有 $m\|x\|^2 \leqslant f''(y)x,x) \leqslant M\|x\|^2$。则有:

1)该算法产生一个满足一阶和二阶必要条件的聚点 x^*;

2)若 $f(x)$ 的Hessian阵 G^∞ 正定,则对主序列 $r_k \to 1, x_k \to x^\infty$,$x_k$ 的最小下界大于零,约束 $\|s\| < h_k$;此外,该算法有二阶收敛速度。

3 具体算法及实现步骤

3.1 算法一

step0 给出初始点 $x_1 \in R^n$,其中 $\epsilon \geqslant 0, h_1 = \|g_1\|_2, B_1 = I \in R^{n\times n}$,其中 g 定义如前,置 $k:=0$(以下用 B_k 近似目标函数中的 G_k):

step1 分别按式(5,6)计算 g_k, B_k,若 $\|g_k\| < \epsilon$,则停止;

step2 分解 $B_k + \mu_k I$,若不正定,令 $\mu_k = 4\mu_k$,重复判断直到 $B_k + \mu_k I$ 正定;

step3 求解式(4),得 s_k,并计算 $r_k = \dfrac{f(x_k) - f(x_k + s_k)}{f(x_k) - q^{(k)}(s_k)}$,同时用改进的黄金分割法[2]计算步长因子 α_k,

若 $r_k > 0$,计算 $\alpha_k = (g_k,s_k)/(\|g_k\|\ \|s_k\|)$

如果 $\alpha_k < \sqrt{2}/2$,令 $d_k = \alpha_k s_k$,否则,令 $d_k = \alpha_k(-g_k)$;$x_{k+1} = x_k + d_k$,

否则,$x_{k+1} = x_k$,

step4 若 $r_k \geqslant 0.25$,$\mu_k = 4\mu_k, h_k = h_k$,转step5;$h_k = 0.5h_k$,转step6;

step5 若 $r_k \leqslant 0.75$,且 $\|h_k\| = \|s_k\|$,$\mu_k = \mu_k$,转step6;$h_k = 2h_k$;

step6 $h_k = h_k$,由式(6)计算 B_k,$k = k+1$,转step2。

算法的改进:

我们发现若在step3中,使用Samaskii加速技巧可以提高算法一的效率,即在从 x_k 到 x_{k+1} 的迭代过程中进行 m 次简化迭代。其具体算法如下:

3.2 算法二

step0 给出初始点 $x_1 \in R^n, \epsilon \leqslant 0, h_1 = \|g_1\|_2, B_1 = I \in R^{n\times n}$,其中 g 定义如前面所述,

置 $k:=0$(以下用 B_k 近似目标函数中的 G_k);

step1 分别按式(5,6),计算 g_k, B_k,若 $\|g_k\| < \epsilon$,则停止;

step2 分解 $B_k + \mu_k I$,若不正定,令 $\mu_k = 4\mu_k$,重复判断直到 $B_k + \mu_k I$ 正定;

step3 求解式(4),得 s_k,并计算 r_k,同时用改进的黄金分割法计算步长因子 α_k,

若 $r_k > 0$,计算 $\alpha_k = (-g_k,s_k)./(\|-g_k\|\ \|s_k\|)$,

if $\alpha_k < \sqrt{2}/2$,令 $d_k = \alpha_k s_k$,

else 令 $d_k = \alpha^k(-g_k)$,

$x_{k+1} = x_k + d_k$,

$x_{k,0} = x_k$,

$x_{k,i} = x_{k,i-1} + [F'(x_k)]^{-1}F(x_{k,i-1})$,

$i = 1,\cdots,m,\quad k = 0,1,\cdots$

$x_{k+1} = x_{k,m}$;

step4 若 $r_k \geqslant 0.25$,$\mu_k = 4\mu_k, h_k = h_k$,转step5,$h_k = 0.5h_k$,转step6;

step5 若 $r_k \leqslant 0.75$ 且 $\|h_k\| = \|s_k\|$,$\mu_k = \mu_k$,转step6;$h_k = 2h_k$;

step6 $h_k = h_k$,由式(2.5)计算 B_k,$k = k+1$,转step2。

否则, $x_{k+1} = x_k$。

维普资讯 http://www.cqvip.com

4 数值分析及数值实验结果

为了验证本文提出方法的可行性,对文献[1,3]给出的算例进行试算,在计算中采用 Visual C ++ 6.0 编写程序,结果如下:

例 1[1]

$$\begin{cases} f_1(x_1,x_2) = x_1x_1 - x_2 + 1 = 0, \\ f_2(x_1,x_2) = x_1 - \cos(0.5 \times 3.141\,592\,6 \times x_2) = 0。\end{cases}$$

精确解为 $x^* = (-1/\sqrt{2}, 1.5)$ 和 $x^* = (0, 1)^T$。

例 2[3]

$$\begin{cases} f_1(x_1,x_2) = x_1 \times x_1 - x_2 - 1 = 0, \\ f_2(x_1,x_2) = (x_1 - 2)^2 + (x_2 - 0.5)^2 - 1 = 0。\end{cases}$$

精确解为 $x^* = (1.546\,342, 1.391\,174)^T$ 和 $x^* = (1.067\,412, 0.139\,46)^T$。

表 1 对例 1,例 2 求解的数值结果

Tab. 1 Solving numerical result for example 1 and 2

例子	算法	初始点	次数	x_1^*	x_2^*	截止误差	$f(x_1^*, x_2^*)$
例 1	算法一	$(0.5,0.5)^T$	10	−0.707 108 1	1.500 000	0.000 003	0.000 002
		$(2.0,2.0)^T$	20	0.000 867 1	0.999 324	0.000 002	0.000 001
		$(1.0,1.0)^T$	15	−0.000 017 2	1.000 133	0.000 001	0.000 002
	算法二	$(0.5,0.5)^T$	6	−0.711 814	1.503 806	0.000 046	0.00 001
		$(2.0,2.0)^T$	12	0.000 871	0.999 441	0.000 031	0.00 003
		$(1.0,1.0)^T$	10	0.000 179	1.000 129	0.00 007	0.00 002
例 2	算法一	$(2.0,1.0)^T$	21	1.067 346	0.139 228	0.000 001	0.000 000
		$(1.0,1.0)^T$	18	1.067 300	0.138 384	0.000 001	0.000 002
		$(0.0,0.0)^T$	16	1.067 272	0.139 234	0.000 001	0.000 000
	算法二	$(2.0,1.0)^T$	9	1.546 342	1.391 174	0.000 002	0.000 000
		$(1.0,1.0)^T$	8	1.067 312	0.139 056	0.000 002	0.000 001
		$(0.0,0.0)^T$	8	1.067 349	0.139 233	0.000 013	0.000 000

参考文献:

[1] 李庆扬,莫孜中,祁力君,等.非线性方程组的数值解法[M].北京:科学出版社,1997.

[2] 袁亚湘,孙文瑜.最优化理论与方法[M].北京:科学出版社,1997.

[3] 孔 敏,沈祖和.解非线性方程组的极大熵法[J].高等学校计算数学学报,1999,(1):1-7.

(编 辑 曹大刚)

Trust region method for sloving nonlinear equations

XUE Xi-feng, XING Zhi-dong, MENG Hong-yun

(Department of Mathematics, Northwest University, Xi'an 710069, China)

Abstract: Trust region method was used to solve nonlinear equation problems, at the same time, objective function and its once derivation, were used in the computation only through BFGS justification matrix was costructed. Finally some numerical results were given.

Key words: trust region method; samaskii techniqe; BFGS justification; alternative direction method