中国农业大学学报 1997,2(2):29～32
Journal of China Agricultural University

混合不精确牛顿法

陈 静① 李正锋
(中国农业大学工程基础科学部)

摘 要 林正华提出的一个求解非线性方程组的混合牛顿与二阶拟牛顿迭代法是有效的,但有缺点。本文中提出了一个改进的算法:在每一迭代步,只需求解线性方程组的近似解。在合理的假设下证明了本算法具有与林算法相同的收敛性质。

关键词 非线性方程组;不精确牛顿法;二阶拟牛顿法;混合迭代法
中图分类号 O221.2

Mixed Inexact Newton Method

Chen Jing Li Zhengfeng
(College of Applied Engineering Sciences, CAU)

Abstract Lin Zhenghua presented an algorithm-mixed Newton and second-order quasi-Newton method for solving the system of nonlinear equations. That algorithm is efficient, but there are some drawbacks. An improved algorithm is developed which requires an approximate solution to the system of linear equations at each iteration. Based on the rational assumption for the residuals, it is shown that all such methods share the same convergence properties as Lin's algorithm.

Key words nonlinear equations; inexact Newton method; second-order quasi-Newton method; hybrid iterative method

考虑求解非线性方程组

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

这里 $F:\mathbf{R}^n\to\mathbf{R}^n$ 是充分光滑的向量函数。牛顿法和拟牛顿法是求解这一问题的两类最重要的方法:前者每次迭代需计算雅可比矩阵,但具有二阶收敛速率;后者每次迭代不必计算雅可比矩阵,但仅具有超线性收敛速率。林正华在文[1]中提出了一个交替使用牛顿迭代和拟牛顿迭代的算法,并对一维情形证明了它具有二步四阶收敛速率,还指出该结论可望推广到高维。这表明虽然插入了计算量较少的拟牛顿迭代,但大体上仍具有牛顿法的收敛速率。看来这是一个值得进一步研究的问题。他的算法可叙述如下。

算法 1(混合牛顿与二阶拟牛顿迭代法)

步 0:给定 $x_0\in\mathbf{R}^n$,$k:=0$。

步 1:修正 $\boldsymbol{B}_k$。

若 $k=2m$ 时,$\boldsymbol{B}_k=F'(x_k)$,

收稿日期:1996-08-24
①陈 静,北京清华东路 17 号中国农业大学(东校区)71 信箱,100083

若 $k=2m+1$ 时，$B_k=B_{k-1}+\frac{(y_{k-1}-B_{k-1}s_{k-1})s_{k-1}^{\mathrm{T}}}{s_{k-1}^{\mathrm{T}}s_{k-1}}$。

这里 $s_{k-1}=x_k-x_{k-1}$，$y_{k-1}=2F(x_k)-F(x_{k-1})$。

步 2：$x_{k+1}=x_k-(B_k)^{-1}F(x_k)$。

步 3：$k:=k+1$ 转步 1 直到某个终止准则满足为止。

但是，在每一迭代点，都要求解一个线性方程组。当方程组的变量个数很大时，这个计算量是很大的，并且也是不必要的，尤其是在当前迭代点远离解点时。为了克服这一缺点，笔者提出一种改进的算法：在每一迭代步，只需求解线性方程组的一个近似解。这一思想可见于文[2]。更进一步，笔者还证明了改进算法具有与林算法相同的收敛性质。

笔者改进的算法(以下简称算法 2)类似于算法 1，它们的区别只在于步 2。

算法 2 的步 2：选取一参数 η_k，求出近似解 s_k 满足

$$B_k s_k=-F(x_k)+r_k$$

其中 r_k 满足

$$\frac{\|r_k\|}{\|F(x_k)\|}\leqslant\eta_k \tag{2}$$

这里 η_k 可能依赖于 x_k；值得指出的是如果令 $\eta_k\equiv0$，则算法 2 退化为林算法。

引理 1[1] $\forall s\in\mathbf{R}^n$，$s\neq0$，那么

$$\left\|I-\frac{2ss^{\mathrm{T}}}{s^{\mathrm{T}}s}\right\|=1$$

引理 2[3] 假设函数 $F(x)$ 是某个定义域 $D\subset\mathbf{R}^n$ 上的连续函数，并且在其上 $F'(x)$ 是李普希兹连续的，其李氏常数为 γ，则对 $\forall x+p\in D$ 有

$$\|F(x+p)-F(x)-F'(x_*)p\|\leqslant\frac{\gamma}{2}\|p\|^2$$

引理 3[2] 假设 x_* 为式(1)的一个解且 $F'(x_*)$ 非奇异，则存在 x_* 的一个邻域，使得

$$c^{-1}\|x-x_*\|\leqslant\|F(x)\|\leqslant c\|x-x_*\|$$

这里 c 为某常数。

引理 4 假设 $F(x)$ 满足引理 2 和引理 3 的条件，并设序列 $\{x_k\}$ 和 $\{B_k\}$ 是由算法 2 生成的，若存在常数 c' 使得

$$\|\eta_k\|\leqslant c'\|s_k\|^2 \tag{3}$$

成立，则存在 x_* 的一个邻域使得下面结论成立：

1)若 $k=2m$ 时，有 $\|B_k-F'(x_*)\|\leqslant\gamma\|x_k-x_*\|$；

2)若 $k=2m+1$ 时，有

$$\|B_k-F'(x_*)\|\leqslant\gamma\|x_k-x_*\|+(\gamma+1+cc')\|x_{k-1}-x_*\|。$$

证明 1)的证明平凡。下面证明 2)。

为了书写简单，现略去下标 k，记新迭代点 x_k 为 x_+，记 $F'(x_*)$ 为 J^*，于是由 B_k 的定义有

$$B_+-J^*=B-J^*+\frac{(y-Bs)s^{\mathrm{T}}}{s^{\mathrm{T}}s}=B-J^*+\frac{[2F(x_+)-F(x)-Bs]s^{\mathrm{T}}}{s^{\mathrm{T}}s}$$

利用 $Bs=-F(x)+r$，有

$$B_+-J^*=B-J^*+\frac{2[F(x_+)-F(x)]+r-2Bss^{\mathrm{T}}}{s^{\mathrm{T}}s}=$$

$$B-J^*+\frac{2[F(x_+)-F(x)-J^*s]s^{\mathrm{T}}-2(B-J^*)ss^{\mathrm{T}}+r}{s^{\mathrm{T}}s}=$$

$$(B-J^*)\left(I-\frac{2ss^{\mathrm{T}}}{s^{\mathrm{T}}s}\right)+\frac{2[F(x_+)-F(x)-J^*s]s^{\mathrm{T}}}{s^{\mathrm{T}}s}+\frac{r}{s^{\mathrm{T}}s}$$

由引理 1,引理 2 和式(2)得

$$\|B_+-J^*\|\leqslant\|B-J^*\|+\gamma(\|x_+-x\|)+\eta_k\|F(x_k)\|/\|s\|^2$$

再由引理 3 及条件式(3)有

$$\|B_+-J^*\|\leqslant\gamma\|x_+-x_*\|+(\gamma+1+cc')\|x-x_*\|$$

定理 1　假设引理 4 的条件成立,则存在 $\delta_1>0$ 和 $\delta_2>0$ 使得当

$$\|x_0-x_*\|\leqslant\delta_1,\|B_0-F'(x_*)\|\leqslant\delta_2$$

时,由算法 2 生成的 $\{x_k\}$Q-线性收敛到 x_*。

证明　由引理 4 和文[3]中的定理 8.2.2,可知该定理成立。

引理 5　若 $F:\mathbf{R}^1\to\mathbf{R}^1$ 是二次连续可微函数,$F''(x)$是李普希兹连续函数,那么若式(3)成立,则存在 $\delta>0$,当 $\|x_0-x_*\|<\delta$ 时,由算法 2 生成的 $\{B_k\}$满足

1)当 k 为偶数时,$B_k=F'(x_k)$;

2)当 k 为奇数时,$\|B_k-F'(x_k)\|\leqslant\bar{c}s_{k-1}^2$。　　(4)

其中 $s_{k-1}=x_k-x_{k-1}$,$\bar{c}$为某一常数。

证明　仅证 k 为奇数的情形。

利用泰勒展式有

$$F(x_{k-1})=F(x_k)-F'(x_k)s_{k-1}+\frac{1}{2}F''(t_1)s_{k-1}^2 \tag{5}$$

和

$$F(x_k)=F(x_{k-1})+F'(x_{k-1})s_{k-1}+\frac{1}{2}F''(t_2)s_{k-1}^2 \tag{6}$$

其中 t_1 和 t_2 均位于 x_{k-1}和 x_k 之间的线段上。式(6)减去(5)有

$$2F(x_k)-2F(x_{k-1})=[F'(x_{k-1})+F'(x_k)]s_{k-1}+\frac{1}{2}[F''(t_2)-F''(t_1)]s_{k-1}^2$$

从而

$$2F(x_k)-F(x_{k-1})-F'(x_k)s_{k-1}=F(x_{k-1})+F'(x_{k-1})+\frac{1}{2}[F''(t_2)-F''(t_1)]s_{k-1}^2$$

利用 $B_ks_{k-1}=2F(x_k)-F(x_{k-1})$有

$$[B_k-F'(x_k)]s_{k-1}=[F(x_{k-1})+F'(x_{k-1})s_{k-1}]+\frac{1}{2}[F''(t_2)-F''(t_1)]s_{k-1}^2$$

即

$$B_k-F'(x_k)=\frac{r_{k-1}}{s_{k-1}}+\frac{1}{2}[F''(t_2)-F''(t_1)]s_{k-1}$$

因为 $F''(x)$是李普希兹连续,进而有

$$\|B_k-F'(x_k)\|\leqslant\frac{\eta_{k-1}\|F(x_{k-1})\|}{\|s_{k-1}\|}+\frac{\gamma_0}{2}s_{k-1}^2$$

其中 γ_0 为 $F''(x)$的李氏常数。注意到条件式(3)和引理 3,则有

$$\|B_k-F'(x_k)\|\leqslant\left(cc'+\frac{\gamma_0}{2}\right)s_{k-1}^2$$

令 $\bar{c}=cc'+\frac{\gamma_0}{2}$，则知式(4)成立。

由引理5知，当 $n=1$ 时，即得到了比引理4更强的一个结果。

定理2 假设引理5的条件成立，则当 x_0 充分接近 x_* 时，由算法2生成的点列 $\{x_k\}$ 收敛到 x_*，并且是二步Q-4次收敛的。

证明 因偶数迭代步是牛顿步，知(见文[3])存在 $c_2>0$ 使得

$$\|x_k-x_*\|\leqslant c_2\|x_{k-1}-x_*\|^2 \tag{7}$$

当 k 为奇数时，由引理5知

$$\|\boldsymbol{B}_k-F'(x_k)\|\leqslant\frac{\gamma_0}{2}\|s_{k-1}\|^2$$

由于 $\{x_k\}$ 局部线性收敛到 x_*(见定理1)，因此

$$\|\boldsymbol{B}_k-F'(x_k)\|\leqslant\gamma_0\|x_{k-1}-x_*\|^2 \tag{8}$$

由于 $\boldsymbol{B}_k s_k=-F(x_k)+r_k$，于是有

$$\begin{aligned}\boldsymbol{B}_k(x_{k+1}-x_*)&=\boldsymbol{B}_k(x_{k+1}-x_k)+\boldsymbol{B}_k(x_k-x_*)=\\&\boldsymbol{B}_k(x_k-x_*)-F(x_k)+r_k=\\&[\boldsymbol{B}_k-F'(x_k)](x_k-x_*)-F(x_k)+F'(x_k)(x_k-x_*)+r_k\end{aligned}$$

从而

$$\|x_{k+1}-x_*\|\leqslant\|\boldsymbol{B}_k^{-1}\|[\|\boldsymbol{B}_k-F'(x_k)\|\|x_k-x_*\|+\|F(x_k)+F'(x_k)(x_*-x_k)\|+\|r_k\|] \tag{9}$$

由泰勒展式知

$$F(x_*)=0=F(x_k)+F'(x_k)(x_*-x_k)+O(\|x_k-x_*\|^2)$$

进而

$$\|F(x_k)+F'(x_k)(x_*-x_k)\|\leqslant c_1\|x_k-x_*\|^2 \tag{10}$$

其中 c_1 为某一常数。

再由引理3及把式(3)，(7)，(8)和(10)代入(9)知存在常数 $c_3>0,c_4>0$ 使得

$$\|x_{k+1}-x_*\|\leqslant c_3\|x_{k-1}-x_*\|^4+c_4\|x_{k-1}-x_*\|^6$$

于是存在常数 $c_5>0$ 使得

$$\|x_{k+1}-x_*\|\leqslant c_5\|x_{k-1}-x_*\|^4。$$

参考文献

1 林正华.牛顿与二阶拟牛顿混合迭代法.高等学校计算数学学报，1994(3)：217～224

2 Dembo R，Eisenstat S C，Steihaug T. Inexact Newton methods. SIAM J Numer Anal，1982，19(2)：400～408

3 Dennis J E Jr，Schnabel R B. Numerical Methods for Unconstrained Optimization and Nonlinear Equations. New Jersey：Prentice-Hall Inc，1983. 70～82，168～190