第15卷 第4期 北京工业大学学报 Vol.15, No.4
1989年12月 JOURNAL OF BEIJING POLYTECHNIC UNIVERSITY Dec., 1989

非线性ABS算法的推广及其收敛性质

诸梅芳

（应用数学系）

【摘要】 将求解非线性方程组的ABS算法加以推广，并证明了推广了的算法具有局部收敛性和二阶收敛速率.

关键词：非线性ABS算法，二阶收敛速率

本文考虑求解非线性方程组

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

这里$f(x)=(f_1(x),\cdots,f_n(x))^T$是$R^n$到$R^n$的连续映射，其Jacobi矩阵记为

$$A(x)=\begin{pmatrix} a_1^T(x) \\ \cdots\cdots \\ a_n^T(x)\end{pmatrix}=\left(\frac{\partial f_i}{\partial x_j}\right) \tag{2}$$

Abaffy，Broyden和Spedicato在文献[1]中提出了一类求解线性方程组的算法[2]，后来这类方法又被推广到求解非线性方程组[2,3]. 这些算法都具有两种形式，即逐个求解的形式和分组分块求解的形式. 对线性方程组来说，逐个方程求解形式产生的近似点列包含了分块求解形式所有可能产生的近似点列. 然而文献[2]或文献[3]中对非线性方程组的算法并不具有上述性质. 本文进一步推广了非线性ABS算法，使其具有与线性ABS算法完全相对应的结果，从而使理论系统更加完善.

为简单起见，本文仅叙述逐个方程求解形式的算法及其收敛性质. 对于一般分块求解形式的算法及其相应结论，可以用与本文类似的方法得到.

1 非线性ABS算法的推广

本节首先建立推广的非线性ABS算法，然后做一些必要的说明.

1.1 推广的非线性ABS算法

适当选取初始点x_1. 下面列出从x_i计算x_{i+1}的内部迭代格式：

1.1.1 置$y_1=x_i$，H_1为单位阵，$k=1$.

1.1.2 构造y_1，…，y_k的适当的凸组合u_k和s_k：

$$u_k=\sum_{j=1}^{k}\mu_j y_j,\qquad \mu_j\geqslant 0,\quad \sum_{j=1}^{k}\mu_j=1$$

$$s_k=\sum_{j=1}^{k}\lambda_j y_j,\qquad \lambda_j\geqslant 0,\quad \sum_{j=1}^{k}\lambda_j=1$$

1.1.3 计算搜索方向$p_k=H_k^T z_k$，其中z_k是满足$P_k^T[A(u_k)]^T v_k\neq 0$的任意向量，这里要

本文于1989年8月30日收到.

求v_k与v_1、…，v_{k-1}线性无关.

1.1.4 计算步长

$$d_k=[v_k^T A(u_k)p_k]^{-1}v_k^T \overline{f}_k$$

其中

$$\overline{f}_k=f(s_k)-A(u_k)\sum_{j=1}^{k-1}\overline{\lambda}_j p_j d_j$$

$$\overline{\lambda}_j=\sum_{i=1}^{j}\lambda_i$$

1.1.5 计算新迭代点

$$y_{k+1}=y_k-p_k d_k$$

若$k=n$，则置$x_{i+1}=y_{n+1}$并转出该内部迭代.

1.1.6 修正迭代矩阵

$$H_{k+1}=H_k-H_k[A(u_k)]^T v_k w_k^T H_k$$

其中w_k是满足$w_k^T H_k[A(u_k)]^T v_k=1$的任意向量.

1.1.7 置$k=k+1$，转1.1.2.

1.2 说明

1.2.1 若在上述算法的步骤1.1.2中，限定恒取

$$u_k=0y_1+\cdots+0y_{k-1}+1y_k=y_k$$

$$s_k=0y_1+\cdots+0y_{k-1}+1y_k=y_k$$

则得到文献[2]中的算法. 若仅限定取$s_k=y_k$(u_k仍可取为y_1, …, y_k的任意凸组合)，则得到文献[3]中的算法. 可见上述算法是它们的推广. 另外，若在上述算法中限定$u_k=y_1=x_i$，$s_k=y_1=x_i(k=1, \cdots, n)$，则得到Newton法. 显然，当仅考虑逐个方程求解形式的算法时，文献[2]和文献[3]中的算法都不包括Newton法，而上述算法则以Newton法为其特殊情形.

1.2.2 上述逐个方程求解形式的算法所产生的近似点列，包含了相应分块求解形式产生的近似点列，而文献[2]和文献[3]中的算法都不具有这一性质.

2 算法的收敛性质

本节除特别申明外，范数$\|\cdot\|$均指l_2-范数

2.1 定理

设

(1) $f(x^*)=0$，且$f(x)$在x^*的Jacobi矩阵$A(x^*)$非奇异;

(2) 存在着正数δ_0和τ_0，使对任意的$x\in U(x^*, \delta_0)$和$y\in U(x^*, \delta_0)$均有

$$\|f(x)-f(y)\|\leqslant\tau_0\|x-y\|$$

其中$U(x^*, \delta_0)$是以x^*为中心、δ_0为半径的闭球;

(3) 存在着正数τ_1，使对任意的$x\in U(x^*, \delta_0)$和$y\in U(x^*, \delta_0)$均有

$$\|A(x)-A(y)\|_F\leqslant\tau_1\|x-y\|$$

其中$\|\cdot\|_F$是矩阵的Frobenius范数;

(4) 存在着正数τ_2和τ_3，使对所有算法中使用的p_k，v_k和u_k一致地有

$$\|p_k[v_k^T A(u_k)p_k]^{-1}v_k^T\|_F \leqslant \tau_2$$

$$\mathrm{Cond}(V) \leqslant \tau_3$$

其中 $\mathrm{Cond}(V)=\|V\|\cdot\|V^{-1}\|$ 是矩阵 $V=(v_1, \cdots, v_n)$ 的条件数．则存在着正数 δ，使当 $x_1\in U(x^*, \delta)$ 时，由上述推广的非线性ABS算法产生的点列 $\{x_i\}$ 收敛到 x^*，且收敛速率是二阶的．

为证明这一定理，需要下面的两个引理：

引理 1　若定理的条件成立，则存在着正数 δ_1 和 τ_4，使当 $x\in U(x^*, \delta_1)$ 时，有

$$\|x-x^*\| \leqslant \tau_4\|f(x)\|$$

证明　由文献[4]中的一个结果可以得到引理 1 的结论．

引理2　若定理的条件成立，则存在着正数 δ_2 和 τ_5，使得当 $y_1\in U(x^*, \delta_2)$，$k=2, \cdots, n$ 时，$y_1, \cdots, y_k$ 的任意凸组合 $\sum_{j=1}^{k}\eta_j y_j$ 满足

$$\left\|\sum_{j=1}^{k}\eta_j y_j - x^*\right\| \leqslant \tau_5\|y_1-x^*\| \tag{3}$$

$$\sum_{j=1}^{k}\eta_j y_j \in U(x^*, \delta_0)\cap U(x^*, \delta_1) \tag{4}$$

其中 δ_1 由引理 1 所示．

证明　令

$$\tau_5 = (1+\tau_0\tau_2+M\tau_2)^n \tag{5}$$

$$\delta_2 = \min\{\delta_0/\tau_5, \ \delta_1/\tau_5\} \tag{6}$$

其中

$$M = \max\{\|A(x)\|_F: \ x\in U(x^*, \delta_0)\} \tag{7}$$

我们即可证明引理 2 的结论成立．事实上，只需证明当 $k=2, \cdots, n$ 时，下列各式成立：

$$\|p_{k-1}d_{k-1}\| \leqslant \tau_0\tau_2(1+\tau_0\tau_2+M\tau_2)^{k-2}\|y_1-x^*\| \tag{8}$$

$$\sum_{j=1}^{k}\eta_j y_j \in U(x^*, \delta_0)\cap U(x^*, \delta_1) \tag{9}$$

其中 η_j 为任意满足 $\sum_{j=1}^{k}\eta_j=1$ 的非负数．

我们采用数学归纳法证明(8)式和(9)式成立．

当 $k=2$ 时，由算法步骤（1.1.4）和定理的条件(1)、(2)和(4)可知

$$\begin{aligned}\|p_1d_1\| &\leqslant \|p_1[v_1^T A(u_1)p_1]^{-1}v_1^T\|_F\cdot\|f(y_1)\| \\ &\leqslant \tau_2\|f(y_1)-f(x^*)\| \\ &\leqslant \tau_0\tau_2\|y_1-x^*\|\end{aligned}$$

进而可知，y_1 与 y_2 的任意凸组合 $\eta_1y_1+\eta_2y_2$ 满足

$$\begin{aligned}\|\eta_1y_1+\eta_2y_2-x^*\| &\leqslant \eta_1\|y_1-x^*\|+\eta_2\|y_1-x^*\|+\eta_2\|p_1d_1\| \\ &\leqslant \|y_1-x^*\|+\|p_1d_1\| \\ &\leqslant (1+\tau_0\tau_2)\|y_1-x^*\| \\ &\leqslant \tau_5\|y_1-x^*\|\end{aligned}$$

再由(6)式便知

$$\eta_1 y_1+\eta_2 y_2 \in U(x^*,\ \delta_0)\cap U(x^*,\ \delta_1)$$

因此，当$k=2$时，(8)式和(9)式成立.

假定当$2\leqslant k\leqslant i$ ($i\leqslant n-1$) 时(8)式和(9)式成立. 于是，当$k=i+1$时，由算法步骤(1.1.4) 知

$$p_i d_i=p_i[v_i^{\mathrm{T}}A(u_i)p_i]^{-1}v_i^{\mathrm{T}}[f(s_i)-A(u_i)\sum_{j=1}^{i-1}\overline{\lambda}_j p_j d_j] \tag{10}$$

由归纳假定知

$$u_i\in U(x^*,\ \delta_0)\cap U(x^*,\ \delta_1),\ s_i\in\subset U(x^*,\ \delta_0)\cap U(x^*,\ \delta_1) \tag{11}$$

注意到

$$\begin{aligned} s_i&=\sum_{j=1}^{i}\lambda_j y_j=\sum_{j=1}^{i}\lambda_j(y_1+\sum_{t=1}^{j-1}p_t d_t)\\ &=y_1+\sum_{j=1}^{i}\lambda_j(\sum_{t=1}^{j-1}p_t d_t)\end{aligned}$$

进而由定理的条件(1)和(2)可知

$$\begin{aligned}\|f(s_i)\|&=\|f(s_i)-f(x^*)\|\leqslant\tau_0\|s_i-x^*\|\\ &\leqslant\tau_0[\|y_1-x^*\|+\sum_{t=1}^{i-1}\|p_t d_t\|]\end{aligned} \tag{12}$$

由定理的条件(4)和(10)～(12)式可得

$$\begin{aligned}\|p_i d_i\|&\leqslant\tau_2[\|f(s_i)\|+M\sum_{j=1}^{i-1}\|p_j d_j\|]\\ &\leqslant\tau_0\tau_2\|y_1-x^*\|+(\tau_0\tau_2+M\tau_2)\sum_{j=1}^{i-1}\|p_j d_j\|\end{aligned}$$

据此由归纳假定便知

$$\begin{aligned}\|p_i d_i\|&\leqslant\tau_0\tau_2\|y_1-x^*\|+(\tau_0\tau_2+M\tau_2)\tau_0\tau_2\|y_1-x^*\|\sum_{j=1}^{i-1}(1+\tau_0\tau_2+M\tau_2)^{j-1}\\ &=\tau_0\tau_2(1+\tau_0\tau_2+M\tau_2)^{i-1}\|y_1-x^*\|\end{aligned}$$

并且对y_1, …, y_{i+1}的任意凸组合$\sum_{j=1}^{i+1}\eta_j y_j$均有

$$\begin{aligned}\|\sum_{j=1}^{i+1}\eta_j y_j-x^*\|&=\|y_1-x^*+\sum_{j=1}^{i+1}\eta_j(\sum_{t=1}^{j-1}p_t d_t)\|\\ &\leqslant\|y_1-x^*\|+\sum_{t=1}^{i}\|p_t d_t\|\\ &\leqslant(1+\tau_0\tau_2+M\tau_2)^i\|y_1-x^*\|\\ &\leqslant\tau_5\|y_1-x^*\|\end{aligned}$$

再由(6)式便知

$$\sum_{j=1}^{i+1}\eta_j y_j\in U(x^*,\ \delta_0)\cap U(x^*,\ \delta_1)$$

至此，引理2的结论得证.

2.2 定理的证明

首先根据算法的构造，证明$v_k^{\mathrm{T}}f(x_{i+1})=v_k^{\mathrm{T}}f(y_{n+1})$满足

$$v_k^{\mathrm{T}}f(y_{n+1})=v_k^{\mathrm{T}}[\overline{A}_k-A(u_k)](y_{n+1}-s_k),\ k=1,\ \cdots,\ n \tag{13}$$

其中

$$\overline{A}_k=\begin{pmatrix} a_1^{\mathrm T}\ (s_k+\theta_1(y_{n+1}-s_k)) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_n^{\mathrm T}\ (s_k+\theta_n(y_{n+1}-s_k)) \end{pmatrix},\quad 0<\theta_i<1,\ i=1,\ \cdots,\ n \tag{14}$$

为此考察 $f(y_{n+1})$ 和$f(s_k)$ 的差．由中值定理知

$$f(y_{n+1})-f(s_k)=\overline{A}_k(y_{n+1}-s_k)$$

其中$\overline{A}_k$如式(14)所示．

另外，由于算法的步骤（1.1.4）相当于求解下列以d_k为未知量的方程

$$v_k^{\mathrm T}A(u_k)p_kd_k=v_k^{\mathrm T}\overline{f}_k=v_k^{\mathrm T}[f(s_k)-A(u_k)\sum_{j=1}^{k-1}\overline{\lambda}p_jd_j]$$

所以有

$$v_k^{\mathrm T}f(s_k)-v_k^{\mathrm T}A(u_k)\sum_{j=1}^{k}\overline{\lambda}p_jd_j=0$$

因此

$$\begin{aligned} v_k^{\mathrm T}f(y_{n+1})&=v_k^{\mathrm T}[f(s_k)+\overline{A}_k(y_{n+1}-s_k)] \\ &\quad+v_k^{\mathrm T}A(u_k)\sum_{j=1}^{k}\overline{\lambda}_jp_jd_j-v_k^{\mathrm T}A(u_k)\sum_{j=1}^{k}\overline{\lambda}_jp_jd_j \\ &=v_k^{\mathrm T}\overline{A}_k(y_{n+1}-s_k)-v_k^{\mathrm T}A(u_k)\sum_{j=1}^{k}\overline{\lambda}_jp_jd_j \end{aligned} \tag{15}$$

注意到当$j>k$时，搜索方向p_j满足

$$v_k^{\mathrm T}A(u_k)p_j=0$$

可知

$$\begin{aligned} v_k^{\mathrm T}A(u_k)\sum_{j=1}^{k}\overline{\lambda}_jp_jd_j&=v_k^{\mathrm T}A(u_k)(\sum_{j=1}^{k}\overline{\lambda}_j\ p_jd_j+\sum_{j=k+1}^{n}p_jd_j) \\ &=-v_k^{\mathrm T}A(u_k)(y_{n+1}-s_k) \end{aligned} \tag{16}$$

将(16)式代入(15)式便知(13)式成立．

令

$$\delta=\min\{\delta_2,\delta_3\}$$

其中δ_2如引理 2 所示，$\delta_3=[16\tau_1\tau_5^2\tau_4n^{3/2}\mathrm{Cond}(V)]^{-1}$

试估计$|v_k^{\mathrm T}f(y_{n+1})|$ $(k=1,\cdots,n)$．先考虑矩阵 $\overline{A}_k-A(u_k)$ 的第j行．由定理条件（3）知

$$\begin{aligned} \|a_j^{\mathrm T}\ (s_k+\theta_j(y_{n+1}-s_k))-a_j^{\mathrm T}(u_k)\|&\leqslant\tau_1\|s_k-u_k+\theta_j(y_{n+1}-s_k)\| \\ &\leqslant4\tau_1\tau_5\|y_1-x^*\| \end{aligned}$$

上边最后一个不等式是根据引理 2 得到的．同时由引理 2 可知

$$\|y_{n+1}-s_k\|\leqslant2\tau_5\|y_1-x^*\|$$

把以上两式代入(13)式，可得

$$\begin{aligned} |v_k^{\mathrm T}f(y_{n+1})|&\leqslant\|v_k\|\cdot\sqrt{n}\cdot8\tau_1\tau_5^2\|y_1-x^*\|^2 \\ &\leqslant\|V\|_{\mathrm F}\cdot8\tau_1\tau_5^2\sqrt{n}\ \|y_1-x^*\|^2 \end{aligned}$$

因此

$$\|V^T f(y_{n+1})\| \leqslant \|V\|_F \cdot 8\tau_1\tau_5^2 n \|y_1 - x^*\|^2$$

利用 $f(y_{n+1}) = V^{-T}V^T f(y_{n+1})$ 及 $\|V\|_F \leqslant \sqrt{n}\|V\|$ 有

$$\begin{aligned}\|f(y_{n+1})\| &\leqslant \|V^{-T}\| \cdot \|V^T f(y_{n+1})\| \\ &\leqslant \|V^{-1}\| \cdot \|V\| \cdot 8\tau_1\tau_5^2 n^{3/2} \|y_1 - x^*\|^2 \\ &= \mathrm{Cond}(V) \cdot 8\tau_1\tau_5^2 n^{3/2} \|y_1 - x^*\|^2\end{aligned}$$

于是根据引理1便知

$$\begin{aligned}\|y_{n+1} - x^*\| &\leqslant \tau_4 \|f(y_{n+1})\| \\ &\leqslant \mathrm{Cond}(V) \cdot 8n^{3/2}\tau_1\tau_4\tau_5^2 \|y_1 - x^*\|^2\end{aligned}$$

可见当 $y_1 \in U(x^*, \delta)$ 时有

$$\|y_{n+1} - x^*\| \leqslant \frac{1}{2}\|y_1 - x^*\|$$

即知当 $x_1 \in U(x^*, \delta)$ 时，序列 $\{x_k\}$ 收敛到 x^*．同时易见它收敛的阶不低于二阶．定理证毕．

参考文献

[1] Abaffy J, Broyden C G and Spedicato E. A Class of Direct Methods for Linear Equations. Numer. Math, 1984; 45: 361～376

[2] Abaffy J and Spedicato E. ABS Projection Algorifhms (mathematical techniques for linear and nonlinear equations), Halsted Press a division of John Wiley & Sons, 1989

[3] Abaffy J, Galantai A and Spedicato E. The Local Convergence of ABS Methods for Nonlinear Algebraic Systems, Numer. Math, 1987; 51: 429～439

[4] Gay D M. Brown's Method and Some Generalizations with Applications to Minimization Problems, Report No. TR75-225, Cornell University, 1975

Generalization of Nonlinear ABS Algorithm and Its Convergence Properties

Zhu Meifang

(Department of Applied Mathematics)

[Abstract] The ABS algorithm for nonlinear algebraic systems is generalized. The local convergence and quadratic convergence of the generalized algorithm have also been proved.

Key Words: Nonlinear ABS algorithm, Quadratic convergence