中国农业大学学报　1998,3(6):24～27
Journal of China Agricultural University

关于用 Broyden 方法解线性方程组终止性的讨论①

钟　萍②

(中国农业大学工程基础科学部)

摘　要　D. P. O'Leary 在一定条件下证明了 Broyden 方法在求解 n 维线性方程组时至多 $2n$ 步终止,又指出当该条件不成立时终止将提前发生。本文中对后者证明了既存在着提前终止的情况,也存在着恰好 $2n$ 步终止的情况。

关键词　线性方程组; Broyden 方法; 终止性

中图分类号　O 221.2

A Note on Termination of Broyden Method for Linear System of Equations

Zhong Ping

(College of Applied Engineering Sciences, CAU)

Abstract　D. P. O'Leary has proved that Broyden method must terminate in at most $2n$ steps on linear system of equations with n variables under certain conditions. He said that the termination would actually occur earlier when the conditions do not hold. It is proved that the Broyden method can be terminated earlier and/or terminated in $2n$ steps exactly when the conditions do not hold.

Key words　linear equations; Broyden's method; termination

Broyden 在 1965 年引进了一种解非线性方程组 $f(x)=0, x\in\mathbf{R}^n$(其中 $f(x):\mathbf{R}^n\mapsto\mathbf{R}^n$ 为可微函数)的迭代方法[1]。其迭代格式为 $x_{k+1}=x_k-\boldsymbol{H}_k f(x_k)$,此式中的 $\boldsymbol{H}_k$ 由 $\boldsymbol{H}_{k-1}$通过下式修正得来:$\boldsymbol{H}_k=\boldsymbol{H}_{k-1}+(s_{k-1}-\boldsymbol{H}_{k-1}y_{k-1})v_{k-1}^{\mathrm{T}}$,这里 $\boldsymbol{H}_0$ 是非奇异初始矩阵,$v_{k-1}\in\mathbf{R}^n$ 是参数,满足 $v_{k-1}^{\mathrm{T}}y_{k-1}=1, s_{k-1}=x_k-x_{k-1}, y_{k-1}=f_k-f_{k-1}$。

现考虑用 Broyden 方法求解 n 维线性方程组

$$g(x)=\boldsymbol{A}x-b \tag{1}$$

其中 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$非奇异,$x\in\mathbf{R}^n$。Gay 在 1979 年证明了此时至多 $2n$ 步终止[2]。O'Leary 在 1995 年重新以更简明的形式证明了这一结论[3],他提出以下 2 个条件成立:条件 1. $v_j^{\mathrm{T}}y_{j-1}\neq 0, j=1,2,\cdots,k$,且 $y_k\neq 0$;条件 2. v_0 属于 $\boldsymbol{F}_0^{\mathrm{T}}$ 的值域。

O'Leary 首先在这 2 个条件下证明了用 Broyden 方法解 n 维线性方程组至多 $2n$ 步终止,然后指出当这 2 个条件不完全成立时,终止将提前发生,但对此未给出证明。笔者在本文中证明当这 2 个条件不完全成立时确实有终止提前发生的情况,但也存在着恰好 $2n$ 步终止的情

收稿日期:1998-07-15

①国家自然科学基金和北京市自然科学基金资助项目

②钟　萍,北京清华东路 17 号 中国农业大学(东校区)71 信箱,100083

况。下面分 3 种情况进行讨论:情况 Ⅰ. 对 $j\geqslant 1$,有 $y_j\neq 0, v_j^{\mathrm{T}}y_{j-1}=0, v_j^{\mathrm{T}}y_{j-2}\neq 0$。情况 Ⅱ. 对 $j\geqslant 1$,有 $y_j\neq 0, v_j^{\mathrm{T}}y_{j-1}=0$ 和 $v_j^{\mathrm{T}}y_{j-1}\neq 0$ 都存在。情况 Ⅲ. 对 $j\geqslant 1$,有 $y_j\neq 0, v_j^{\mathrm{T}}y_{j-1}=0$;另外 $v_j^{\mathrm{T}}y_{j-2}\neq 0$ 和 $v_j^{\mathrm{T}}y_{j-2}=0$ 都存在。

讨论情况 Ⅰ。

为了讨论方便,先以引理的形式给出文献[3]中的 2 个结论。

引理 1　考虑用 Broyden 方法解 n 维线性方程组(1),可以得到:

1)若令 $\boldsymbol{F}_k=\boldsymbol{I}-\boldsymbol{A}\boldsymbol{H}_k$($\boldsymbol{H}_k$ 为 Broyden 方法中的迭代矩阵),则有 $\boldsymbol{F}_k=\boldsymbol{F}_{k-1}(\boldsymbol{I}-y_{k-1}v_{k-1}^{\mathrm{T}})$。

2)若令 $\boldsymbol{P}_k=\boldsymbol{F}_k\boldsymbol{F}_{k-1}\cdots\boldsymbol{F}_0=\boldsymbol{F}_k\boldsymbol{P}_{k-1}$,则有 $g_{k+1}=\boldsymbol{P}_k g_0$,其中 $g_i=g(x_i)$。

引理 2　设情况 Ⅰ 成立且 $\mathrm{rank}(\boldsymbol{F}_1)=n-1$,那么对 $i\geqslant 2$,有 $\mathrm{rank}(\boldsymbol{F}_i)=n-2$,$\mathrm{Null}(\boldsymbol{F}_i)=\mathrm{span}\{y_{i-1},y_{i-2}\}$。

证明　利用归纳法证明。当 $i=2$ 时有 $\boldsymbol{F}_2y_1=\boldsymbol{F}_1(\boldsymbol{I}-y_1v_1^{\mathrm{T}})y_1=0$,$\boldsymbol{F}_2y_0=\boldsymbol{F}_1(\boldsymbol{I}-y_1v_1^{\mathrm{T}})y_0=0$。令 $ay_0+by_1=0$,则由 $\boldsymbol{F}_1(ay_0+by_1)=0$,$\boldsymbol{F}_1y_0=0$ 和 $\boldsymbol{F}_1y_1\neq 0$ 知 $b=0$,从而知 $a=0$,所以 y_0,y_1 线性无关。如果还存在着 $y\neq 0$ 使 $\boldsymbol{F}_2y=0$ 成立,那么有 $y=(v_1^{\mathrm{T}}y)y_1+c_0y_0$,其中 $c_0\in\mathbf{R}$,所以 $\mathrm{Null}(\boldsymbol{F}_2)=\mathrm{span}\{y_0,y_1\}$,$\mathrm{rank}(\boldsymbol{F}_2)=n-2$。

假设 $i=l$ 时该引理成立,下面证明当 $i=l+1$ 时该引理也成立。容易证明 $\boldsymbol{F}_{l+1}y_l=0$,$\boldsymbol{F}_{l+1}y_{l-1}=0$,$\boldsymbol{F}_{l+1}y_{l-2}\neq 0$,且 y_{l-1},y_l 线性无关。如果还存在着 $y\neq 0$ 使 $\boldsymbol{F}_{l+1}y=0$ 成立,那么 $y=(v_l^{\mathrm{T}}y)y_l+c_{l-1}y_{l-1}+c_{l-2}y_{l-2}$,其中 $c_{l-1},c_{l-2}\in\mathbf{R}$;又由 $\boldsymbol{F}_{l+1}y=c_{l-2}\boldsymbol{F}_{l+1}y_{l-2}=0$,得到 $c_{l-2}=0$;所以 $\mathrm{Null}(\boldsymbol{F}_{l+1})=\mathrm{span}\{y_{l-1},y_l\}$,$\mathrm{rank}(\boldsymbol{F}_{l+1})=n-2$。证毕。

引理 3　设 $k=3m+1$(m 为非负整数),情况 Ⅰ 成立且 $\mathrm{rank}(\boldsymbol{F}_1)=n-1$,那么存在一组向量 $\beta_1,\alpha_2,\cdots,\beta_{3m+1},\alpha_{3m+2},\cdots,\beta_k,\alpha_{k+1}$满足:

1)当 $i=1$ 时,有 $\beta_1^{\mathrm{T}}\boldsymbol{F}_j=0$,$j=1,2,\cdots,k+1$;当 $i=3m+1(m=1,2,\cdots)$时,有 $\beta_i^{\mathrm{T}}\boldsymbol{F}_j=\beta_{i-3}^{\mathrm{T}}$,$j=i,i+1,\cdots,k+1$;当 $i=3m+1(m=0,1,\cdots)$时,有 $\beta_i^{\mathrm{T}}\boldsymbol{P}_j=0$,$j=i,i+1,\cdots,k+1$;当 $i=3m+1(m=0,1,\cdots)$时,有 $\beta_i^{\mathrm{T}}g_j=0$,$j=i+1,i+2,\cdots,k+2$;当 $i=3m+1(m=0,1,\cdots)$时,有 $\beta_i^{\mathrm{T}}y_j=0$,$j=i+1,i+2,\cdots,k+1$。

2)当 $i=2$ 时,有 $\alpha_2^{\mathrm{T}}\boldsymbol{F}_j=0$,$j=2,3,\cdots,k+1$;当 $i=3m+2(m=1,2,\cdots)$时,有 $\alpha_i^{\mathrm{T}}\boldsymbol{F}_j=\alpha_{i-3}^{\mathrm{T}}$,$j=i,i+1,\cdots,k+1$;当 $i=3m+2(m=0,1,\cdots)$时,有 $\alpha_i^{\mathrm{T}}\boldsymbol{P}_j=0$,$j=i,i+1,\cdots,k+1$;当 $i=3m+2(m=0,1,\cdots)$时,有 $\alpha_i^{\mathrm{T}}g_j=0$,$j=i+1,i+2,\cdots,k+2$;当 $i=3m+2(m=0,1,\cdots)$时,有 $\alpha_i^{\mathrm{T}}y_j=0$,$j=i+1,i+2,\cdots,k+1$。

3)$\beta_1,\alpha_2,\beta_4,\alpha_5,\cdots,\beta_{3m+1},\alpha_{3m+2},\cdots,\beta_k,\alpha_{k+1}$是一组线性无关的向量。

证明　利用归纳法证明。当 $i=1$ 时,由 $\mathrm{rank}(\boldsymbol{F}_1)=n-1$ 可知存在着 $\beta_1\neq 0$ 使 $\beta_1^{\mathrm{T}}\boldsymbol{F}_1=0$,并且有

$$\beta_1^{\mathrm{T}}\boldsymbol{F}_j=\beta_1^{\mathrm{T}}\boldsymbol{F}_{j-1}(I-y_{j-1}v_{j-1}^{\mathrm{T}})=\beta_1^{\mathrm{T}}\boldsymbol{F}_1(I-y_1v_1^{\mathrm{T}})\cdots(I-y_{j-1}v_{j-1}^{\mathrm{T}})=0,j=2,3,\cdots,k+1;$$

$$\beta_1^{\mathrm{T}}\boldsymbol{P}_j=\beta_1^{\mathrm{T}}\boldsymbol{F}_j\boldsymbol{P}_{j-1}=0,j=1,2,\cdots,k+1;$$

$$\beta_1^{\mathrm{T}}g_j=\beta_1^{\mathrm{T}}\boldsymbol{P}_{j-1}g_0=0,j=2,3,\cdots,k+2;$$

$$\beta_1^{\mathrm{T}}y_j=\beta_1^{\mathrm{T}}g_{j+1}-\beta_1^{\mathrm{T}}g_j=0,j=2,3,\cdots,k+1。$$

当 $i=2$ 时,由 $\mathrm{rank}(\boldsymbol{F}_2)=n-2$ 可知存在着与 β_1 线性无关的向量 α_2 使 $\alpha_2^{\mathrm{T}}\boldsymbol{F}_2=0$,并且容易证明 $\alpha_2^{\mathrm{T}}\boldsymbol{F}_j=0$,$j=2,3,\cdots,k+1$;$\alpha_2^{\mathrm{T}}\boldsymbol{P}_j=0$,$j=2,3,\cdots,k+1$;$\alpha_2^{\mathrm{T}}g_j=0$,$j=3,4,\cdots,k+2$;$\alpha_2^{\mathrm{T}}y_j=0$,

$j=3,4,\cdots,k+1$。

假设存在着一组线性无关的向量 $\beta_1,\alpha_2,\beta_4,\alpha_5,\cdots,\beta_{3m+1},\alpha_{3m+2},\cdots,\beta_{k-3},\alpha_{k-2}$ 使该引理成立，下面证明存在着一组线性无关的向量 $\beta_1,\alpha_2,\beta_4,\alpha_5,\cdots,\beta_{3m+1},\alpha_{3m+2},\cdots,\beta_k,\alpha_{k+1}$ 使该引理成立。考虑 $x^{\mathrm{T}}\boldsymbol{F}_k=\beta_{k-3}^{\mathrm{T}}$，设 $\boldsymbol{B}_k=(\boldsymbol{F}_k^{\mathrm{T}},\beta_{k-3})$，因为有 $\boldsymbol{B}_k^{\mathrm{T}}y_{k-2}=0,\boldsymbol{B}_k^{\mathrm{T}}y_{k-1}=0$，所以 $x^{\mathrm{T}}\boldsymbol{F}_k=\beta_{k-3}^{\mathrm{T}}$ 有非零解，设为 β_k，即有 $\beta_k^{\mathrm{T}}\boldsymbol{F}_k=\beta_{k-3}^{\mathrm{T}}$。同理，存在着 $\alpha_{k+1}\neq 0$ 使 $\alpha_{k+1}^{\mathrm{T}}\boldsymbol{F}_{k+1}=\alpha_{k-2}^{\mathrm{T}}$，并由归纳假设可知：$\beta_k^{\mathrm{T}}\boldsymbol{F}_{k+1}=\beta_k^{\mathrm{T}}\boldsymbol{F}_k(\boldsymbol{I}-y_kv_k^{\mathrm{T}})=\beta_{k-3}^{\mathrm{T}}$；$\beta_k^{\mathrm{T}}\boldsymbol{P}_j=0,j=k,k+1$；$\beta_k^{\mathrm{T}}g_{k+1}=0$；$\beta_k^{\mathrm{T}}y_{k+1}=0$；$\alpha_{k+1}^{\mathrm{T}}\boldsymbol{F}_{k+2}=\alpha_{k+1}^{\mathrm{T}}\boldsymbol{F}_{k+1}(\boldsymbol{I}-y_{k+1}v_{k+1}^{\mathrm{T}})=\alpha_{k-2}^{\mathrm{T}}$；$\alpha_{k+1}^{\mathrm{T}}\boldsymbol{P}_j=0,j=k+1,k+2$；$\alpha_{k+1}^{\mathrm{T}}g_{k+2}=0$；$\alpha_{k+1}^{\mathrm{T}}y_{k+2}=0$。

下面证明 $\beta_1,\alpha_2,\beta_4,\alpha_5,\cdots,\beta_{3m+1},\alpha_{3m+2},\cdots,\beta_k,\alpha_{k+1}$ 是一组线性无关的向量。设 $c_1\beta_1+c_2\alpha_2+c_4\beta_4+c_5\alpha_5+\cdots+c_k\beta_k+c_{k+1}\alpha_{k+1}=0$，其中 $c_i\in\mathbf{R},i=1,2,\cdots,k+1$，因为有 $(c_1\beta_1^{\mathrm{T}}+c_2\alpha_2^{\mathrm{T}}+\cdots+c_k\beta_k^{\mathrm{T}}+c_{k+1}\alpha_{k+1}^{\mathrm{T}})\boldsymbol{F}_{k+1}=0$，所以 $c_4\beta_1^{\mathrm{T}}+c_5\alpha_2^{\mathrm{T}}+\cdots+c_k\beta_{k-3}^{\mathrm{T}}+c_{k+1}\alpha_{k-2}^{\mathrm{T}}=0$。由归纳假设知 $c_4,c_5,\cdots,c_k,c_{k+1}$ 均为零，又由 β_1,α_2 线性无关知 $c_1=c_2=0$。证毕。

由引理 3 可以得到下面的定理。

定理 1 设情况 Ⅰ 成立且 $\mathrm{rank}(\boldsymbol{F}_1)=n-1$，那么用 Broyden 方法解 n 维线性方程组(1)经过 $k+2$ 步得到的 g_{k+2} 垂直于一组线性无关的向量 $\beta_1,\alpha_2,\cdots,\beta_{3m+1},\alpha_{3m+2},\cdots,\beta_k,\alpha_{k+1}$，因此用 Broyden 方法解 n 维线性方程组至多 $2n-[n/2]$ 步终止，其中 $[n/2]$ 为不超过 $n/2$ 的最大整数。

显然有下面的推论。

推论 1 设情况 Ⅰ 成立，$\mathrm{rank}(\boldsymbol{F}_1)=q$($q$ 为不超过 $n-2$ 的整数)，那么用 Broyden 方法解 n 维线性方程组至多 $2n-[n/2]$ 步终止。

讨论情况 Ⅱ 。

设 $\mathrm{rank}(\boldsymbol{F}_1)=n-1$，在该条件下，现给出一个确定下标 k^* 并构造 n 个线性无关且都与 g_{k^*} 垂直的向量的方法，称之为“向量构造方法”：1)当 $i=1$ 时，取 $z_1^1\neq 0$ 使 $(z_1^1)^{\mathrm{T}}\boldsymbol{F}_1=0$；2)当 $\{z_l^p\}$ 的元素个数为 n 时停止；3)令 $i=i+1,k^*=i+1$；4)若 $\mathrm{rank}(\boldsymbol{F}_{i-1})-\mathrm{rank}(\boldsymbol{F}_i)=m_i>0$，则取 m_i 个使 $x^{\mathrm{T}}\boldsymbol{F}_i=0$ 成立的向量 $z_i^1,z_i^2,\cdots,z_i^{m_i}$，并且使所有满足 $x^{\mathrm{T}}\boldsymbol{F}_i=0$ 的向量 $z_l^p(l=1,2,\cdots,i,p\geqslant 1)$ 线性无关，转 2)；5)若 $\mathrm{rank}(\boldsymbol{F}_{i-1})=\mathrm{rank}(\boldsymbol{F}_i)$，考察 $x^{\mathrm{T}}\boldsymbol{F}_i=(z_l^p)^{\mathrm{T}}(l=1,2,\cdots,i-1,p\geqslant 1)$，若有非零解，则取 $z_i^p\neq 0$ 使 $(z_i^p)^{\mathrm{T}}\boldsymbol{F}_i=(z_l^p)^{\mathrm{T}}$，转 2)，否则转 3)。

由文献[3]中的证明可以看出当 $\mathrm{rank}(\boldsymbol{F}_k)=n-1(k=1,2,\cdots)$ 时，用 Broyden 方法解 n 维线性方程组(1)至多经过 $2n$ 步即可以由向量构造方法找到 n 个线性无关且都与 g_{2n} 垂直的向量。当情况 Ⅱ 成立时，有 $\mathrm{rank}(\boldsymbol{F}_k)\leqslant\mathrm{rank}(\boldsymbol{F}_{k-1})\cdots\leqslant\mathrm{rank}(\boldsymbol{F}_1)=n-1$，所以此时由向量构造方法至多经过 $2n$ 步即可找到 n 个线性无关且都与 g_{2n} 垂直的向量。

定理 2 设情况 Ⅱ 成立且 $\mathrm{rank}(\boldsymbol{F}_1)=n-1$，那么存在着 $k^*\geqslant 1$，使得由向量构造方法得到 n 个线性无关且都与 g_{k^*} 垂直的向量 $z_l^p(l=1,2,\cdots,k,p\geqslant 1)$，因此用 Broyden 方法解 n 维线性方程组(1)至多 $2n$ 步终止。

推论 2 设情况 Ⅱ 成立，$\mathrm{rank}(\boldsymbol{F}_1)=q$($q$ 为不超过 $n-2$ 的整数)，那么用 Broyden 方法解 n 维线性方程组至多 $2n$ 步终止。

在情况 Ⅱ 中，用 Broyden 方法解 n 维线性方程组存在着恰好 $2n$ 步终止的情况，这可以用下面的例子予以说明。

例 $g(x)=\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}\begin{pmatrix}x^{(1)}\\ x^{(2)}\end{pmatrix}-\begin{pmatrix}1\\ 1\end{pmatrix}$。

解　取 $\boldsymbol{H}_0=\boldsymbol{I}, x_0=0$,则 $g_0=\begin{pmatrix}-1\\-1\end{pmatrix}$。

第 1 步:$x_1=x_0-\boldsymbol{H}_0 g_0=\begin{pmatrix}1\\1\end{pmatrix}, g_1=\begin{pmatrix}2\\6\end{pmatrix}, s_0=\begin{pmatrix}1\\1\end{pmatrix}, y_0=\begin{pmatrix}3\\7\end{pmatrix}$;取 $v_0=\begin{pmatrix}-2\\1\end{pmatrix}$,则 $\boldsymbol{H}_1=\boldsymbol{H}_0+(s_0-\boldsymbol{H}_0 y_0)v_0^{\mathrm{T}}=\begin{pmatrix}5 & -2\\12 & -5\end{pmatrix}$。

第 2 步:$x_2=\begin{pmatrix}3\\7\end{pmatrix}, g_2=\begin{pmatrix}16\\36\end{pmatrix}, s_1=\begin{pmatrix}2\\6\end{pmatrix}, y_1=\begin{pmatrix}14\\30\end{pmatrix}$;取 $v_1=\begin{pmatrix}-\frac{29}{14}\\1\end{pmatrix}$,则 $v_1^{\mathrm{T}}y_1=1, v_1^{\mathrm{T}}y_0=\frac{11}{14}\neq 0, \boldsymbol{H}_2=\begin{pmatrix}\frac{151}{7} & -10\\\frac{258}{7} & -17\end{pmatrix}$。

第 3 步:$x_3=\begin{pmatrix}\frac{125}{7}\\\frac{205}{7}\end{pmatrix}, g_3=\begin{pmatrix}\frac{528}{7}\\\frac{1188}{7}\end{pmatrix}, s_2=\begin{pmatrix}\frac{104}{7}\\\frac{156}{7}\end{pmatrix}, y_2=\begin{pmatrix}\frac{416}{7}\\\frac{936}{7}\end{pmatrix}$;取 $v_2=\begin{pmatrix}-\frac{35}{104}\\\frac{49}{312}\end{pmatrix}$,则 $v_2^{\mathrm{T}}y_2=1, v_2^{\mathrm{T}}y_1=0, \boldsymbol{H}_3=\begin{pmatrix}-2 & 1\\\frac{3}{2} & -\frac{1}{2}\end{pmatrix}$。

第 4 步:$x_4=\begin{pmatrix}-1\\1\end{pmatrix}, g_4=0$,终止。

讨论情况Ⅲ。

定理 3　设情况Ⅲ成立且 $\mathrm{rank}(\boldsymbol{F}_1)=n-1$,那么用 Broyden 方法解 n 维线性方程组至多 $2n-[n/2]$步终止。

由向量构造方法和定理 1 容易知道此定理成立。

推论 3　设情况Ⅲ成立,$\mathrm{rank}(\boldsymbol{F}_1)=q$($q$ 为不超过 $n-2$ 的整数),那么用 Broyden 方法解 n 维线性方程组至多 $2n-[n/2]$步终止。

本研究得到邓乃扬教授的指导,谨致谢意。

参 考 文 献

1　Broyden C G. A class of methods for solving nonlinear simultaneous equations. Math Comp, 1965, 19: 577～593

2　Gay D M. Some convergence properties of Broyden's method. SIAM: J Numer Anal, 1979, 16: 623～630

3　O'Leary D P. Why Broyden's nonsymmetric method terminates on linear equations. SIAM: J Optim, 1995, 6: 231～235