用排除-Newton 型混合方法解非线性代数方程组*

祁 力 群

(清 华 大 学)

SOLVING SYSTEMS OF NONLINEAR ALGEBRAIC EQUATIONS WITH EXCLUSION-NEWTON METHOD

Qi Li-qun

(Qing Hua University)

Abstract

An Exclusion-Newton procedure based on some new easily verifiable criterions for existence and convergence is given for solving systems of nonlinear algebraic equations. The method for finding the other solutions after one solution has been obtained is also presented.

一、引

非线性代数方程组有着广泛的实际背景. 有的方程组的未知量个数很多，从几百到几千，出现在非线性代数型微分方程（组）的差分化和非线性代数型有限元问题中. 比如流体力学中的基本方程——Navier-Stokes 方程:

$$\frac{\partial \boldsymbol{v}}{\partial t}+(\boldsymbol{v}\cdot \operatorname{grad})\boldsymbol{v}=-\operatorname{grad}\left(\Omega+\frac{1}{\rho}p\right)+\nu\Delta\boldsymbol{v},$$

离散化以后，就得到一个二次代数方程组，这在连续介质力学中是常见的. 有的未知量个数从 10 多个到百余，例如在电力网络问题中遇到过这样的复二次方程组[10]:

$$\left[\sum_{j=1}^{n} a_{ij}z_j\right] z_i = w_i,\quad i=1,2,\cdots,n.$$

分离实部和虚部后，就得到一个实二次方程组. 在电子网络问题中，还可能有高次方程组，未知量个数只有几个到 10 余个，例如最优化问题中出现的方程组，再如 [9] 中实际求解了 13 个来自双曲扁壳非线性稳定问题的三元三次方程组.

近年来，在一般形式的非线性方程组数值解方面文献虽然很多，但对非线性代数方程组的特殊解法讨论较少[7]，本文试图对非线性代数方程组给出一套简易可行的混合方法；用排除法寻找初值，方法虽然原已有之，但无解的判别原则却是新的，并且简单易行；收敛

* 1979 年 11 月 9 日收到.

判别原则利用了高次方程组可二次化的特点，对二次方程组尤其简便；还给出了求出一解后再求其它解的方法. 文末附计算例题.

二、无解判别原则

一个非线性代数方程组可以记为

$$f_i(\boldsymbol{x})=0,\quad i=1,2,\cdots,n, \tag{1}$$

这里变量 $\boldsymbol{x}\in R^N$, f_i 为 n 元多项式，或简记为

$$F(\boldsymbol{x})=\mathbf{0}. \tag{2}$$

$F:R^n\to R^n$ 为多项式组. 在一个 n 维长方体 X:

$$X=\langle\boldsymbol{u},\boldsymbol{v}\rangle=\{\boldsymbol{x}\,|\,\boldsymbol{u}\leqslant\boldsymbol{x}\leqslant\boldsymbol{v}\} \tag{3}$$

内怎样判定方程组无解？这里不等号均对偏序 F 有意义. [5]中提出的一个充分原则是，若对某一个 i,

$$0\bar{\in}[\min_{x\in X}f_i(\boldsymbol{x}),\ \max_{x\in X}f_i(\boldsymbol{x})], \tag{4}$$

则(1)在 X 中无解. 道理是显然的. 但在一般情况下，min 和 max 只能估算，不很简单[3].

现在我们提出一个简易的原则代替(4)，只限于 $\mathbf{0}\leqslant\boldsymbol{u}<\boldsymbol{v}$ 的情况. 对于一般问题，总可以在一开始就分在 2^n 个卦限内，把别的卦限用改变正负号的方法转换到这种情况. 又改写每个 f_i 为两个正系数多项式 p_i 与 q_i 之差：

$$f_i(\boldsymbol{x})=p_i(\boldsymbol{x})-q_i(\boldsymbol{x}),\ i=1,2,\cdots,n, \tag{5}$$

则有

无解判别原则 1. 如有一个 i，使得

$$0\bar{\in}[p_i(\boldsymbol{u})-q_i(\boldsymbol{v}),p_i(\boldsymbol{v})-q_i(\boldsymbol{u})], \tag{6}$$

则(1)在 X 中无解.

由 p_i 和 q_i 在 $\langle\boldsymbol{u},\boldsymbol{v}\rangle$ 为保序，有

$$[\min_{x\in X}f_i(\boldsymbol{x}),\ \max_{x\in X}f_i(\boldsymbol{x})]\subset[p_i(\boldsymbol{u})-q_i(\boldsymbol{v}),\ p_i(\boldsymbol{v})-q_i(\boldsymbol{u})].$$

从而(6)比(4)更为充分，亦即判别能力弱一点. 但是，

1) 当每个变量在 f_i 中最多出现一次时，

$$[\min_{x\in X}f_i(\boldsymbol{x}),\ \max_{x\in X}f_i(\boldsymbol{x})]=[p_i(\boldsymbol{u})-q_i(\boldsymbol{v}),\ p_i(\boldsymbol{v})-q_i(\boldsymbol{u})],$$

(6)和(4)一致.

2) 如果

$$\langle\boldsymbol{u}_0,\boldsymbol{v}_0\rangle\supset\langle\boldsymbol{u}_1,\boldsymbol{v}_1\rangle\supset\langle\boldsymbol{u}_2,\boldsymbol{v}_2\rangle\supset\cdots\supset\langle\boldsymbol{u}_k,\boldsymbol{v}_k\rangle\supset\cdots$$

都不符合无解判别原则 1，且 $\|\boldsymbol{u}_k-\boldsymbol{v}_k\|\xrightarrow[k\to\infty]{}0$，则有唯一点 $\boldsymbol{x}^*=\bigcap\limits_{k=1}^{\infty}\langle\boldsymbol{u}_k,\boldsymbol{v}_k\rangle$ 为(1)的解. 这由 p_i 和 q_i 的连续性即得.

3) 方法简单，便于计算.

由于这三点，(6)在排除法中比(4)更实用.

三、收敛判别原则

常用的精确化方法是 Newton 法：

$$x^{k+1}=x^k-F'(x^k)^{-1}F(x^k),\ k=0,1,2,\cdots \tag{7}$$

或简化 Newton 法:

$$x^{k+1}=x^k-F'(x^0)^{-1}F(x^k),\ k=0,1,2,\cdots. \tag{8}$$

常用的收敛定理是 Newton-Канторович 定理. 但这个定理要求估计 $\|F''(x)\|$ 在一个域上的值,不太好用,然而对于二次方程组,却较方便. 一个二次方程组可以记为

$$F(x)\equiv \mathrm{A}(x,x)+Bx+c=\mathbf{0}, \tag{9}$$

这里 c 为 n 维向量, B 为 n 维矩阵, A 表示 $R^N\times R^N\to R^N$ 的双重线性算子,则有

收敛判别原则. 如对 $x^0=w$, 有

$$\|F'(w)^{-1}F(w)\|\|F'(w)^{-1}\mathrm{A}\|\leqslant\frac{1}{4}, \tag{10}$$

则(7)和(8)均收敛于(9)在 $S(w,r)$ 内的唯一解,这里

$$r=[1+\sqrt{1-4\|F'(w)^{-1}\mathrm{A}\|\|F'(w)^{-1}F(w)\|}]/2\|F'(w)^{-1}\mathrm{A}\|, \tag{11}$$

$S(w,r)$ 表示 R^n 中以 w 为心、r 为半径的开球. 如果有

$$\|F'(w)^{-1}F(w)\|\|F'(w)^{-1}\|\|\mathrm{A}\|\leqslant\frac{1}{4}, \tag{10a}$$

则结论对下面的 r 成立,这里

$$r=[1+\sqrt{1-4\|F'(w)^{-1}F(w)\|\|F'(w)^{-1}\|\|\mathrm{A}\|}]/2\|F'(w)^{-1}\|\|\mathrm{A}\|. \tag{11a}$$

证明. 当 (10) 成立时, 由 [1] 的定理 2 和 (9) 中的 $F''(x)\equiv 2\mathrm{A}$, 即得结论. 由 $\|F'(w)^{-1}\mathrm{A}\|\leqslant\|F'(w)^{-1}\|\|\mathrm{A}\|$ 即得后半个结论. 证毕.

$\|\cdot\|$ 用何种范数均可,但用 l_∞ 范数较方便,下同.

对于高次方程组,可以增设变量,化为二次方程组,当然这就增加了维数.

[5]中所用的精确化方法是 Krawczyk 法,即区间 Newton 法,用区间算法. 关于这个方法及其收敛判别原则,可见[2, 4, 6]. 但是,区间算法终究比普通算法烦琐.

四、其它无解判别原则

对于二次方程组(9),还有

无解判别原则 2. 记 $w=\frac{1}{2}(u+v)$, 如有

$$\|F(w)\|>\frac{1}{2}\|F'(w)\|\|u-v\|+\frac{1}{4}\|\mathrm{A}\|\|u-v\|^2 \tag{12}$$

或当 $F'(w)^{-1}$ 存在时有

$$\|F'(w)^{-1}F(w)\|>\frac{1}{2}\|u-v\|+\frac{1}{4}\|F'(w)^{-1}\mathrm{A}\|\|u-v\|^2 \tag{13}$$

或

$$\|F'(w)^{-1}F(w)\|>\frac{1}{2}\|u-v\|+\frac{1}{4}\|F'(w)^{-1}\|\|\mathrm{A}\|\|u-v\|^2, \tag{13a}$$

则 $X=\langle u,v\rangle$ 是(9)的无解域.

证明　若 X 内有解 x^*, 则 $\|w-x^*\|\leqslant\frac{1}{2}\|u-v\|$, 且由(9)有

$$F(\boldsymbol{x}^*) = \mathrm{A}(\boldsymbol{x}^* - \boldsymbol{w})^2 + F'(\boldsymbol{w})(\boldsymbol{x}^* - \boldsymbol{w}) + F(\boldsymbol{w}) = 0,$$

$$\begin{aligned}\|F(\boldsymbol{w})\| &\leqslant \|\mathrm{A}\|\|\boldsymbol{x}^* - \boldsymbol{w}\|^2 + \|F'(\boldsymbol{w})\|\|\boldsymbol{x}^* - \boldsymbol{w}\| \\ &\leqslant \frac{1}{4}\|\mathrm{A}\|\|\boldsymbol{u} - \boldsymbol{v}\|^2 + \frac{1}{2}\|F'(\boldsymbol{w})\|\|\boldsymbol{u} - \boldsymbol{v}\|.\end{aligned}$$

若 $F'(\boldsymbol{w})^{-1}$ 存在,则

$$F'(\boldsymbol{w})^{-1}F(\boldsymbol{x}^*) = F'(\boldsymbol{w})^{-1}\mathrm{A}(\boldsymbol{x}^* - \boldsymbol{w})^2 + (\boldsymbol{x}^* - \boldsymbol{w}) + F'(\boldsymbol{w})^{-1}F(\boldsymbol{w}) = \mathbf{0},$$

即

$$\begin{aligned}\|F'(\boldsymbol{w})^{-1}F(\boldsymbol{w})\| &\leqslant \|F'(\boldsymbol{w})^{-1}\mathrm{A}\|\|\boldsymbol{x}^* - \boldsymbol{w}\|^2 + \|\boldsymbol{x}^* - \boldsymbol{w}\| \\ &\leqslant \frac{1}{4}\|F'(\boldsymbol{w})^{-1}\mathrm{A}\|\|\boldsymbol{u} - \boldsymbol{v}\|^2 + \frac{1}{2}\|\boldsymbol{u} - \boldsymbol{v}\| \\ &\leqslant \frac{1}{4}\|F'(\boldsymbol{w})^{-1}\|\|\mathrm{A}\|\|\boldsymbol{u} - \boldsymbol{v}\|^2 + \frac{1}{2}\|\boldsymbol{u} - \boldsymbol{v}\|.\end{aligned}$$

从而当(12)或(13)或(13a)成立时,X 为(9)的无解域. 证毕.

无其它解域定理. 如 $\boldsymbol{x}^*$ 为(9)的解,$F'(\boldsymbol{x}^*)$ 非奇异,则在 $S(\boldsymbol{x}^*, r)$ 内无其它解,这里

$$r = 1/\|F'(\boldsymbol{x}^*)^{-1}\mathrm{A}\| \geqslant 1/\|F'(\boldsymbol{x}^*)^{-1}\|\|\mathrm{A}\|. \tag{14}$$

证明. 在(11)和(11a)中取 $\boldsymbol{w} = \boldsymbol{x}^*$ 即得,证毕.

[8]中已有与上述类似的结果. 无解判别原则 2 比无解判别原则 1 麻烦,但可以作为用收敛判别原则的一个副结果得到.

五、实际方案

对于二次方程组(9)在给定的 n 维长方体 $X_0 = \langle \boldsymbol{u}_0, \boldsymbol{v}_0\rangle$ 上的求解问题(这里 $\mathbf{0} \leqslant \boldsymbol{u}_0 < \boldsymbol{v}_0$,一般情况分块处理),可以预先算出 $\|\mathrm{A}\|$,并记 $X = \langle \boldsymbol{u}, \boldsymbol{v}\rangle$ 为机器正在判别的子域,$\boldsymbol{w} = \frac{1}{2}(\boldsymbol{u} + \boldsymbol{v})$ 为中心点,T 为所有待判别的子域表. 对每个子域,除记录其二个端点向量外,还增记一个在 1 到 n 之间的对分指标,而 X 的对分指标记在 i 中,i 表示一个计数器. 又记预先给定的正数 h_1 为计算机容量所决定的最小可对分子域的直径,h_2 为预先给定的进行收敛判别子域的直径,$h_2 > 2h_1$. 用 h_2 的目的是避免一开始就用烦琐的收敛判别原则,于是有如下对分进行的方案:

排除-Newton 方案

步 1. 清 T, $X_0 \to X$, $1 \to$.

步 2. 用无解原则 1 判别 X 是否无解域. 若是,转入步 5.

步 3. 若 $\|\boldsymbol{u} - \boldsymbol{v}\| \leqslant h_1$, 转入步 9; 若 $h_1 < \|\boldsymbol{u} - \boldsymbol{v}\| \leqslant h_2$, 转入步 6.

步 4. 按第 i 个分量对分 X, 选择或任取一半为 X, 另一半以对分指标 $i + 1$ (当 $i = n$ 取 1) 从后面依次记入表 T, 转入步 2.

步 5. 若表 T 是空的,停止. 否则取出其中最后一个为 X, 其对分指标 $\to i$, 转入步 2.

步 6. 于 $\boldsymbol{w}$ 用收敛判别原则判别,如收敛转入步 8.

步 7. 用无解判别原则 2 判别 X, 如无解,转入步5,否则转入步 4.

步 8. 从 $\boldsymbol{x}^0=\boldsymbol{w}$ 用 Newton 法或简化 Newton 法计算. 求出解后,如仍需求其它解,则据(11)或(11a)或(14)求出无其它解域,从 T 内各子域除去,转入步 5, 否则停止.

步 9. 令 $\boldsymbol{x}^0=\boldsymbol{w}$, 用(8)求出 $\boldsymbol{x}^1$. 如在 X 内,按可能有解域打印出 X; 如不在 X 内,按可能无解域打印出 X; 之后均转入步 5.

由于 T 内只是记录了一些边长递减的子域,并不需要很大容量,用计算机解几十个未知量的二次方程组,是可以容纳得了的. 对于小型问题,用计算器即可解决. 在步 9 给出了不能判别无解和收敛的小子域，这种情况只可能在重根处发生. 事实上，不满足（12）即

$$\begin{aligned}\|F(\boldsymbol{w})\| &\leqslant \frac{1}{2}\|F'(\boldsymbol{w})\|\|\boldsymbol{u}-\boldsymbol{v}\|+\frac{1}{4}\|\mathrm{A}\|\|\boldsymbol{u}-\boldsymbol{v}\|^2\\ &\leqslant \frac{1}{2}\|F'(\boldsymbol{w})\|h_1+\frac{1}{4}\|\mathrm{A}\|h_1^2.\end{aligned}$$

由于 $\|F'(\boldsymbol{x})\|$ 在 X_0 上有界,当 h_1 很小,就表示 $\|F(\boldsymbol{w})\|$ 很小,不满足（10a）和（13a)即

$$\|F'(\boldsymbol{w})^{-1}F(\boldsymbol{w})\|>\frac{1}{4}(\|F'(\boldsymbol{w})^{-1}\|\|\mathrm{A}\|)^{-1},$$

$$\begin{aligned}\|F'(\boldsymbol{w})^{-1}F(\boldsymbol{w})\| &\leqslant \frac{1}{2}\|\boldsymbol{u}-\boldsymbol{v}\|+\frac{1}{4}\|F'(\boldsymbol{w})^{-1}\|\|\mathrm{A}\|\|\boldsymbol{u}-\boldsymbol{v}\|^2\\ &\leqslant \frac{1}{2}h_1+\frac{1}{4}\|F'(\boldsymbol{w})^{-1}\|\|\mathrm{A}\|h_1^2,\end{aligned}$$

即

$$(\|F'(\boldsymbol{w})^{-1}\|\|\mathrm{A}\|h_1)^2+2(\|F'(\boldsymbol{w})^{-1}\|\|\mathrm{A}\|h_1)-1>0,$$

$$\|F'(\boldsymbol{w})^{-1}\|>(\sqrt{2}-1)/\|\mathrm{A}\|h_1.$$

$\|F'(\boldsymbol{w})^{-1}\|$ 很大,或者 $F'(\boldsymbol{w})^{-1}$ 根本不存在,从而 $\boldsymbol{w}$ 只能在重根附近. 所以，这个方案能够求出二次方程组所有近似实解. 对于高次方程组,在用后面几个式子时需二次化,如维数增加太多,就不适用.

六、计算简例

下面的例是个简单的三元二次方程组,借此可以说明一些问题.

例.

$$\begin{cases} x_1^2+x_2=3, & \text{①}\\ 2x_1+x_2^2-x_3=1, & \text{②}\\ x_2+x_3^2=5. & \text{③}\end{cases}$$

解. 容易判定解有界 $\|\boldsymbol{x}\|_\infty\leqslant 6$[11]. 下面按卦限分别讨论. 在 $x_1\leqslant 0$, $x_2\geqslant 0$, $x_3\geqslant 0$ 卦限,令 $u_1=-x_1$, 原方程组变为

$$\begin{cases} u_1^2+x_2=3, & p_1=u_1^2+x_2, & q_1=3, & \text{①}\\ -2u_1+x_2^2-x_3=1, & p_2=x_2^2, & q_2=2u_1+x_3+1, & \text{②}\\ x_2+x_3^2=5, & p_3=x_2+x_3^2, & q_3=5. & \text{③}\end{cases}$$

在 $<0,0,0;6,6,6>$ 求解,下面就是排除表格

1. $\langle 3,0,0;6,6,6\rangle$①; $\langle 0,0,0;3,6,6\rangle$2.
2. $\langle 0,3,0;3,6,6\rangle$①; $\langle 0,0,0;3,3,6\rangle$3.
3. $\langle 0,0,3;3,3,6\rangle$③; $\langle 0,0,0;3,3,3\rangle$4.
4. $\langle 2,0,0;3,3,3\rangle$①; $\langle 0,0,0;2,3,3\rangle$5.
5. $\langle 0,2,0;2,3,3\rangle$6; $0,0,0;2,2,3\rangle$
6. $\langle 0,2,2;2,3,3\rangle$③; $\langle 0,2,0;2,3,2\rangle$7.
7. $\langle 1,2,0;2,3,2\rangle$①; $\langle 0,2,0;1,3,2\rangle$8.
8. $\langle 0,2,1;1,3,2\rangle$9; $\langle 0,2,0;1,3,1\rangle$
9. $\langle 0.5,2,1;1,3,2\rangle$10; $\langle 0,2,1;0.5,3,2\rangle$
10. $\langle 0.5,2.5,1;1,3,2\rangle$②; $\langle 0.5,2,1;1,2.5,2\rangle$11
11. $\langle 0.5,2,1.5;1,2.5,2\rangle$12; $\langle 0.5,2,1;1,2.5,1.5\rangle$
12. $\langle 0.7,2,1.5;1,2.5,2\rangle$13; $\langle 0.5,2,1.5;0.7,2.5,2\rangle$
13. $\langle 0.7,2.2,1.5;1,2.5,2\rangle$14; $\langle 0.7,2,1.5;1,2.2,2\rangle$15
14. $\langle 0.7,2.2,1.7;1,2.5,2\rangle$③; $\langle 0.7,2.2,1.5;1,2.5,1.7\rangle$②
15. $\langle 0.7,2,1.7;1,2.2,2\rangle$16; $\langle 0.7,2,1.5,1,2.2,1.7\rangle$
16. $\langle 0.8,2,1.7;1,2.2,2\rangle$17; $\langle 0.7,2,1.7;0.8,2.2,2\rangle$
17. $\langle 0.8,2,1.8;1,2.2,2\rangle$③; $\langle 0.8,2,1.7;1,2.2,1.8\rangle^*$

为了便于计算，没有完全用对分，而是取整数值分点. 每个子域的右面标有 ①，②或③的，表示被这个方程用无解判别原则 1 排除. 我们总是把数据大的放在左面，先行判别. 子域右面标有 1,2… 数字的，表示进入第几步判别对分，如在 14 两块均被排除，则退回，找到最小且尚未对分的一块到15判别对分. 最后得怀疑块$\langle 0.8,2,1.7;1,2.2,1.8\rangle^*$，取中心点 $\boldsymbol{w}=(0.9,2.1,1.75)^T$，用收敛判别原则判别收敛，$r=1.687$. 用简化 Newton 法得

$$\boldsymbol{x}^1=(0.9307856,2.1345859,1.6936897)^T,$$
$$\boldsymbol{x}^2=(0.9305722,2.1340222,1.6929449)^T.$$

结果已经很好了，即原方程组的解是

$$(-0.9305722,2.1340222,1.6929449)^T.$$

利用 $S(\boldsymbol{w},r)$ 和无解判别原则1，可知这一卦限内不再有解，用类似的办法可得原方程组其余三个怀疑块 $\langle -1.35,1.25,-1.95;-1.25,1.35,-1.90\rangle$；$\langle -2.2,-1.7,-2.6;-2.1,-1.6,-2.55\rangle$；$\langle -2.45,-3,2.8;-2.4,-2.9,2.85\rangle$ 都可以用 Newton 法或简化 Newton 法继续求解.

从以上分析和这个简例可以看出，这个方法的特点是能大范围求出代数方程组特别是二次方程组的所有解，不象连续法等在导数奇异处会产生通不过的困难. 还应当指出的是，无解判别原则(5)可以推广到一般的非线性函数 F，只要它可以在 $\langle \boldsymbol{u},\boldsymbol{v}\rangle$ 分成二个保序算子之差.

最后，谨向赵访熊教授和李庆扬副教授深致谢意，在我工作的过程中，赵访熊教授自始至终给以指导；李庆扬副教授在撰写本文时，给以大力帮助.

参 考 文 献

[1] P. Deuflhard, G. Heindl, *SIAM J. Numer. Anal.*, 16, (1979), 1—10.
[2] R. Krawczyk, *Computing*, 4, (1969). 181—201.
[3] R. E. Moore, *Computing*, 16, (1976), 1—15.
[4] R. E. Moore, *SIAM J. Numer. Anal.*, 14, (1977), 611—615.
[5] R. E. Moore, S. T. Jones, *SIAM J. Numer. Anal.*, 14, (1977), 1051—1066.
[6] R. E. Moore, *SIAM J. Numer. Anal.*, 15, (1978), 1194—1196.
[7] W. C. Rheinboldt, Methods for Solving systems of nonlinear equations, SIAM Book Series, Philadelphia, (1974), 79—94.
[8] 赵访熊,解非线性方程组的牛顿法,清华大学工程力学数学系,总 64077 力 013, (1964).
[9] 江善标,非线性代数方程组的解法,清华大学数力系计算方法专业研究生毕业论文, (1965).
[10] 潮流计算小组,用牛顿-块超松弛法解电网潮流问题,电子技术与自动控制,清华大学电子工程系编,1,(1972)。
[11] 祁力群，关于代数方程组的数值解法，参加全国计算数学年会学术论文，清华大学应用数学专业，(1979), 1—16。