应用数学和力学，第16卷 第7期（1995年7月）
Applied Mathematics and Mechanics

应用数学和力学编委会编
重庆出版社出版

一类求解非线性代数方程组的并行多分裂AOR算法

白 中 治

（复旦大学数学研究所，上海 200433）

（张汝清推荐，1994年5月27日收到）

摘 要

本文提出了求解大型非线性代数方程组$A\varphi(x)+B\psi(x)=b$的并行多分裂AOR(Accelerated Overrelaxation)算法。在一定的条件下，证明了非线性代数方程组解的存在唯一性，并建立了新算法的全局收敛性理论。

关键词 非线性代数方程组 并行算法 松弛 H-矩阵

一、引 言

Stefan 问题和许多弱非线性椭圆型偏微分方程经过有限元和差分离散后，均可产生大型非线性代数方程组

$$A\varphi(\boldsymbol{x})+B\psi(\boldsymbol{x})=\boldsymbol{b},\quad A,B\in L(R^n),\quad \boldsymbol{x},\boldsymbol{b}\in R^n, \tag{1.1}$$

这里，$\varphi(\boldsymbol{x})=(\varphi_i(x_i))$，$\psi(x)=(\psi_i(x_i))\in R^n$为连续函数，$\boldsymbol{x}$为未知向量而$\boldsymbol{b}$为常向量。

利用矩阵多分裂的思想[1]，White[2] 于1986年为这类具有重要的实际应用价值的特殊问题的求解设计了一种并行非线性Gauss-Seidel 算法。该算法在具体实施中取得了良好的数值效果。

本文进一步在文[2]的算法中引进松弛参数，建立了求解大型非线性方程组(1.1)的一类并行多分裂 AOR(Accelerated Overrelaxation) 算法。由于其中有两个参数可供任意选择，从而使得该算法不仅灵活实用，还可获得较快的收敛速度。相应于松弛参数的特殊选取，新算法不仅以文[2]中的并行非线性 Gauss-Seidel 算法为特例，而且还可产生并行多分裂外插Gauss-Seidel，并行多分裂 SOR(Successive Overrelaxation) 等求解非线性方程组(1.1)的许多实用而有效的算法。在适当的条件下，证明了非线性方程组 (1.1) 在R^n上解的存在唯一性，并建立了新算法的全局收敛性理论。

二、并行算法的建立

记 $A=(a_{ij})$，$B=(b_{ij})$，$D=\mathrm{diag}(A)$，$E=\mathrm{diag}(B)$。给定正整数 $\alpha(\alpha\leqslant n)$，对 $k=$

$1,2,\cdots,\alpha$，设$L_k=(l_{ij}^{(k)})$，$M_k=(m_{ij}^{(k)})\in L(R^n)$分别为严格下三角矩阵，$U_k=(u_{ij}^{(k)})$，$V_k=(V_{ij}^{(k)})\in L(R^n)$分别为对角元是零的矩阵，而 $E_k=\mathrm{diag}(e_1^{(k)},e_2^{(k)},\cdots,e_n^{(k)})\in L(R^n)$ 为非负矩阵。如果

(i) $A=D-L_k-U_k \quad (k=1,2,\cdots,\alpha)$

(ii) $B=E-M_k-V_k \quad (k=1,2,\cdots,\alpha)$

(iii) $\det(D)\neq 0$，$\det(E)\neq 0$；

(iv) $\sum\limits_k E_k=I \quad (I\in L(R^n)$为单位矩阵)

则称$(D-L_k,U_k;\ E-M_k,\ V_k;\ E_k)(k=1,2,\cdots,\alpha)$，为矩阵对$(A;\ B)$的一个多分裂。

由此，构造求解大型非线性代数方程组(1.1)的并行多分裂 AOR(Accelerated Overrelaxation)算法如下：

算法：给定初始向量$x^0\in R^n$对$p=0,1,2,\cdots$，计算

$$x_i^{p+1}=\sum_k e_i^{(k)}x_i^{p,k} \qquad (i=1(1)n) \tag{2.1}$$

其中

$$\left.\begin{aligned} x_i^{p,k}&=\frac{\omega}{r}\ \bar{x}_i^{p,k}+\left(1-\frac{\omega}{r}\right)x_i^p \\ \bar{x}_i^{p,k}&=r\hat{x}_i^{p,k}+(1-r)x_i^p \end{aligned} \quad (i=1(1)n;\ k=1,2,\cdots,\alpha)\right\} \tag{2.2}$$

而 $\hat{x}_i^{p,k}(i=1(1)n,k=1,2,\cdots,\alpha)$逐次地由方程组

$$\begin{aligned} a_{ii}\varphi_i(\hat{x}_i^{p,k})+b_{ii}\psi_i(\hat{x}_i^{p,k})-\sum_{j<i}[\,l_{ij}^{(k)}\varphi_j(\bar{x}_j^{p,k})+m_{ij}^{(k)}\psi_j(\bar{x}_j^{p,k})\,] \\ -\sum_{j\neq i}[u_{ij}^{(k)}\varphi_j(x_j^p)+v_{ij}^{(k)}\psi_j(x_j^p)]=b_i \\ (i=1(1)n;\ k=1,2,\cdots,\alpha) \end{aligned} \tag{2.3}$$

确定。这里，r为松弛因子，ω为加速因子。

显然，(2.2)可等价地表示为

$$\left.\begin{aligned} x_i^{p,k}&=\omega\hat{x}_i^{p,k}+(1-\omega)x_i^p \\ \bar{x}_i^{p,k}&=r\hat{x}_i^{p,k}+(1-r)x_i^p \end{aligned} \quad (i=1(1)n,k=1,2,\cdots,\alpha)\right\} \tag{2.4}$$

在由(2.1)，(2.3)～(2.4)所定义的并行算法中，相应于参数(r,ω)的特殊选取$(0,1)$，$(0,\omega)$，$(1,1)$，$(1,\omega)$和(ω,ω)，可得到求解大型非线性方程组(1.1)的并行多分裂Jacobi法，外插型Jacobi法，Gauss-Seidel法[2]，外插型Gauss-Seidel法[2]以及SOR算法。特别是，当$\varphi(x)=x$，$B=I$时，新算法即退化为文[3]中所建立的求解大型非线性代数方程组$Ax+\psi(x)=b$的并行多分裂AOR算法。

三、预 备 知 识

在下面的讨论中，将沿用[1]～[2]中的符号和概念；用$\langle\cdot\rangle$，$\rho(\cdot)$分别表示相应矩阵的比较矩阵和谱半径。另外，将用到非线性方程组(1.1)的如下基本假设：

(A_1) $A\in L(R^n)$为H-矩阵；

(A_2) φ,ψ：$R^n\to R^n$均为连续的对角映射，且对任意的$x,y\in R^n$，有

$$\begin{cases}|\varphi(x)-\varphi(y)|\geqslant\varepsilon|x-y|, & \varepsilon=\operatorname{diag}(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\geqslant0\\ |\psi(x)-\psi(y)|\geqslant\eta|x-y|, & \eta=\operatorname{diag}(\eta_1,\eta_2,\cdots,\eta_n)\geqslant0\end{cases},$$

(A_3) $P:=|D|\varepsilon+|E|\eta$为正对角矩阵；

(A_4) $\operatorname{sgn}(a_{ii}b_{ii})(\varphi_i(s)-\varphi_i(t))(\psi_i(s)-\psi_i(t))\geqslant0.(\forall s,t\in R^1;\ i=1(1)n)$

(A_5) $\langle A\rangle|D|^{-1}|E|\leqslant\langle B\rangle$；

(A_6) 对任意的$x,y,z\in R^n$，$t\in R^1$，有

$$\begin{cases}|\varphi(tx+(1-t)y)-\varphi(z)|\leqslant|t||\varphi(x)-\varphi(z)|+|1-t||\varphi(y)-\varphi(z)|\\ |\psi(tx+(1-t)y)-\psi(z)|\leqslant|t||\psi(x)-\psi(z)|+|1-t||\psi(y)-\psi(z)|\end{cases}$$

为得到非线性方程组(1.1)的解的存在唯一性，及建立新算法的全局收敛定理，需要如下几个引理。

引理1 定义$g:R^n\to R^n$为

$$g(x)=D\varphi(x)+E\psi(x) \tag{3.1}$$

则有：

(1) 当基本假设(A_2)～(A_4)满足时，g是$R^n\to R^n$的同胚映像，并且有

$$\begin{aligned}|g(x)-g(y)|&=|D||\varphi(x)-\varphi(y)|+|E||\psi(x)-\psi(y)|\\&\geqslant P|x-y|,\quad\forall x,y\in R^n\end{aligned} \tag{3.2}$$

(2) 当基本假设(A_2)与(A_6)满足时，对任意给定的正整数N和任何x，$x^{(i)}\in R^n$，$t_i\in R^1$，$i=1(1)N$，只要$\sum_{i=1}^N t_i=1$，就有

$$\left|g\left(\sum_{i=1}^N t_ix^{(i)}\right)-g(x)\right|\leqslant\sum_{i=1}^N|t_i||g(x^{(i)})-g(x)| \tag{3.3}$$

证明 直接推演即知结论(1)成立。现利用归纳法证明(2)。

$N=1$时，(3.3)是显然的。假设当$N=m$时，(3.3)成立，则当$N=m+1$时，记

$$T^{(m)}=\sum_{i=1}^m t_i,\quad X^{(m)}=\sum_{i=1}^m\frac{t_i}{T(m)}x^{(i)},$$

由于

$$\sum_{i=1}^{m+1}t_ix^{(i)}=T^{(m)}X^{(m)}+t_{m+1}x^{(m+1)},$$

$$T^{(m)}+t_{m+1}=1,\quad\sum_{i=1}^m\frac{t_i}{T^{(m)}}=1,$$

利用归纳假设可得

$$\begin{aligned}\left|g\left(\sum_{i=1}^m t_ix^{(i)}\right)-g(x)\right|&=|g(T^{(m)}X^{(m)}+t_{m+1}x^{(m+1)})-g(x)|\\&\leqslant|T^{(m)}||g(X^{(m)})-g(x)|+|t_{m+1}||g(x^{(m+1)})-g(x)|\end{aligned}$$

$$\leqslant |T^{(m)}| \sum_{i=1}^{m} \frac{|t_i|}{|T^{(m)}|} |g(\boldsymbol{x}^{(i)}) - g(\boldsymbol{x})| + |t_{m+1}| |g(\boldsymbol{x}^{(m+1)}) - g(\boldsymbol{x})|$$

$$= \sum_{i=1}^{m+1} |t_i| |g(\boldsymbol{x}^{(i)}) - g(\boldsymbol{x})|,$$

这说明结论(3.3)对于$N=m+1$亦成立。根据归纳法，(3.3)得证。

引理2[1]. 若 $A\in L(R^n)$ 为M-矩阵，$A=B_k-C_k(k=1,2,\cdots,a)$ 均为A的弱正则分裂，则 $\rho\left(\sum_k E_k B_k^{-1} C_k\right)<1$.

四、并行算法的全局收敛性分析

首先，证明非线性方程组(1.1)在R^n中解的存在唯一性。

定理1 如果基本假设$(A_1)\sim(A_5)$满足．则非线性方程组(1.1)对于任何右端向量$\boldsymbol{b}\in R^n$，在R^n中存在唯一解$x^*\in R^n$.

证明 记

$$G=D-A,\quad H=E-B.$$

由于A是 H-矩阵，则 $\rho(|D|^{-1}|G|)=\rho(|G||D|^{-1})<1$。根据非负矩阵理论的 Perron-Frobinius定理和谱半径的连续性知，对充分小的$\delta>0$，有

$$\rho_\delta=\rho(|G||D|^{-1}+\delta \boldsymbol{e}\boldsymbol{e}^T)<1,\quad \boldsymbol{e}=(1,1,\cdots,1)^T\in R^n \tag{4.1}$$

且存在正向量$\boldsymbol{x}_\delta\in R^n$，使得

$$\rho(|G||D|^{-1}+\delta \boldsymbol{e}\boldsymbol{e}^T)\boldsymbol{x}_\delta=\rho_\delta \boldsymbol{x}_\delta. \tag{4.2}$$

设$g: R^n\to R^n$由(3.1)定义．现对于给定的初值$\boldsymbol{x}^0\in R^n$，构造迭代序列$\{\boldsymbol{x}^p\}$为

$$g(\boldsymbol{x}^{p+1})=\boldsymbol{b}+G\varphi(\boldsymbol{x}^p)+H\psi(\boldsymbol{x}^p)\quad (p=0,1,2,\cdots) \tag{4.3}$$

由引理1(1)易知，$\{x^p\}$在R^n上唯一确定，且存在$\sigma>0$，使得

$$|g(x^1)-g(x^0)|\leqslant\sigma \boldsymbol{x}_\delta \tag{4.4}$$

注意到(A_5)亦等价于

$$|H|\leqslant|G||D|^{-1}|E| \tag{4.5}$$

再利用(3.2)和(4.2)可得

$$\begin{aligned}|g(\boldsymbol{x}^{p+1})-g(\boldsymbol{x}^p)|&\leqslant|G||\varphi(\boldsymbol{x}^p)-\varphi(\boldsymbol{x}^{p-1})|+|H||\psi(\boldsymbol{x}^p)-\psi(\boldsymbol{x}^{p-1})|\\&=|G||D|^{-1}|g(\boldsymbol{x}^p)-g(\boldsymbol{x}^{p-1})|+(|H|-|G||D|^{-1}|E|)|\psi(\boldsymbol{x}^p)-\psi(\boldsymbol{x}^{p-1})|\\&\leqslant|G||D|^{-1}|g(\boldsymbol{x}^p)-g(\boldsymbol{x}^{p-1})|\\&\leqslant(|G||D|^{-1}+\delta\boldsymbol{e}\boldsymbol{e}^T)^p|g(\boldsymbol{x}^1)-g(\boldsymbol{x}^0)|\\&\leqslant\rho_\delta^p\rho\boldsymbol{x}_\delta,\end{aligned}$$

从而，对任何正整数q，成立

$$|g(\boldsymbol{x}^{p+q+1})-g(\boldsymbol{x}^p)|\leqslant\frac{\sigma}{1-\rho_\delta}\rho_\delta^p\boldsymbol{x}_\delta \tag{4.6}$$

(4.6)表明$\{g(\boldsymbol{x}^p)\}$为R^n中的Cauchy序列。又据引理1(1)知，g是 $R^n\to R^n$ 的同胚映像，故$\{x^p\}$亦为R^n中的Cauchy序列。因此，$\lim\limits_{p\to\infty}\boldsymbol{x}^p$存在。记$\lim\limits_{p\to\infty}\boldsymbol{x}^p=x^*$.

现在(4.3)两端取极限，即知x^*为方程组(1.1)的解。

设$y^*\in R^n$为方程组(1.1)的另一解。利用(3.2)和(4.5)并经简单计算，可得

$$|g(x^*)-g(y^*)|\leqslant|G||D|^{-1}|g(x^*)-g(y^*)|,$$

或

$$(I-|G||D|^{-1})|g(x^*)-g(y)|\leqslant 0 \tag{4.7}$$

考虑到$\rho(|G||D|^{-1})<1$，则$(I-|G||D|^{-1})^{-1}\geqslant 0$。结合(4.7)有

$$g(x^*)=g(y^*)$$

再根据$g: R^n\to R^n$的同胚性质，即得

$$y^*=x^*$$

此即表明$x^*\in R^n$为非线性方程组(1.1)在R^n中的唯一解。

现在，讨论新算法的全局收敛性。

定理2 如果基本假设$(A_1)\sim(A_4)$和(A_6)满足，$(D-L_k,U_k;\ E-M_k,\ V_k;\ E_k)(k=1,2,\cdots,\alpha)$为矩阵对$(A;\ B)$的一个多分裂，且

$$\langle A\rangle=|D|-|L_k|-|U_k|=|D|-|G| \qquad (k=1,2,\cdots,\alpha) \tag{4.8}$$

$$|M_k|\leqslant|L_k||D|^{-1}|E|,\quad |V_k|\leqslant|U_k||D|^{-1}|E| \qquad (k=1,2,\cdots,\alpha) \tag{4.9}$$

则当

$$0\leqslant r\leqslant\omega,\quad 0<\omega<2/(1+\rho(|D|^{-1}|G|)) \tag{4.10}$$

时，由(2.1)～(2.3)定义的并行多分裂AOR算法从任何初值$x^0\in R^n$出发所产生的序列$\{x^p\}$均收敛到非线性方程组(1.1)在R^n中的唯一解$x^*\in R^n$。

证明 (4.9)显然蕴含有假设(A_5)。根据定理1，非线性方程组(1.1)在R^n中存在唯一解$x^*\in R^n$。

注意到(2.3)，显然x^*满足

$$a_{ii}\varphi_i(x_i^*)+b_{ii}\psi_i(x_i^*)-\sum_{j<i}[l_{ij}^{(k)}\varphi_j(x_j^*)+m_{ij}^{(k)}\psi_j(x_j^*)]$$

$$-\sum_{j\neq i}[u_{ij}^{(k)}\varphi_j(x_j^*)+v_{ij}^{(k)}\psi_j(x_j^*)]=b_i$$

$$(i=1(1)n;\quad k=1,2,\cdots,\alpha) \tag{4.11}$$

(2.3)与(4.11)相减并利用(3.1)知，

$$\begin{aligned}|g(\hat{x}^{p,k})-g(x^*)|\leqslant&|L_k||\varphi(\bar{x}^{p,k})-\varphi(x^*)|+|M_k||\psi(\bar{x}^{p,k})-\psi(x^*)|\\&+|U_k||\varphi(x^p)-\varphi(x^*)|+|V_k||\psi(x^p)-\psi(x^*)|\end{aligned} \tag{4.12}$$

又由(3.2)，我们有

$$\left.\begin{aligned}&|\varphi(\bar{x}^{p,k})-\varphi(x^*)|=|D|^{-1}(|g(\bar{x}^{p,k})-g(x^*)|-|E||\psi(\bar{x}^{p,k})-\psi(x^*)|)\\&|\varphi(x^p)-\varphi(x^*)|=|D|^{-1}(|g|(x^p)-g(x^*)|-|E||\psi(x^p)-\psi(x^*)|)\end{aligned}\right\} \tag{4.13}$$

将(4.13)代入(4.12)并整理得

$$\begin{aligned}|g(\hat{x}^{p,k})-g(x^*)|\leqslant&|L_k||D|^{-1}|g(\bar{x}^{p,k})-g(x^*)|+|U_k||D|^{-1}|g(x^p)-g(x^*)|\\&+(|M_k|-|L_k||D|^{-1}|E|)|\psi(\bar{x}^{p,k})-\psi(x^*)|\\&+(|V_k|-|U_k||D|^{-1}|E|)|\psi(x^p)-\psi(x^*)|.\end{aligned}$$

现利用(4.9)，即可得

$$|g(\hat{x}^{p,k})-g(x^*)|\leqslant|L_k||D|^{-1}|g(\bar{x}^{p,k})-g(x^*)|+|U_k||D|^{-1}|g(x^p)-g(x^*)| \tag{4.14}$$

由(2.2)知，

$$\bar{x}^{p,k}=\frac{r}{\omega}x^{p,k}\left(1-\frac{r}{\omega}\right)x^p \tag{4.15}$$

再考虑到(2.4)并利用引理1(2)，我们有

$$\left.\begin{aligned}&|g(x^{p,k})-g(x^*)|\leqslant\omega|g(\hat{x}^{p,k})-g(x^*)|+|1-\omega||g(x^p)-g(x^*)|\\&|g(\bar{x}^{p,k})-g(x^*)|\leqslant\frac{r}{\omega}|g(x^{p,k})-g(x^*)|+\left(1-\frac{r}{\omega}\right)|g(x^p)-g(x^*)|\end{aligned}\right\} \tag{4.16}$$

结合(4.14)与(4.16)可得

$$\begin{aligned}|g(x^{p,k})-g(x^*)|\leqslant&(|D|-r|L_k|)^{-1}[|1-\omega||D|+(\omega-r)|L_k|+\omega|U_k|]\\&|D|^{-1}|g(x^p)-g(x^*)|\end{aligned} \tag{4.17}$$

记

$$\mathscr{L}(r,\omega)=\sum_k E_k(|D|-r|L_k|)^{-1}[|1-\omega||D|+(\omega-r)|L_k|+\omega|U_k|],$$

由(4.17)得

$$\sum_k E_k|g(x^{p,k})-g(x^*)|\leqslant|D|\mathscr{L}(r,\omega)|D|^{-1}|g(x^p)-g(x^*)| \tag{4.18}$$

现对每个$i\in\{1,2,\cdots,n\}$，根据引理1(2)，有

$$\begin{aligned}|g_i(x_i^{p+1})-g_i(x_i^*)|&=\left|g_i\left(\sum_k e_i^{(k)}x_i^{p,k}\right)-g_i(x_i^*)\right|\\&\leqslant\sum_k e_i^{(k)}|g_i(x_i^{p,k})-g_i(x_i^*)|,\end{aligned}$$

利用(4.18)，最终可得

$$|g(x^{p+1})-g(x)|\leqslant|D|\mathscr{L}(r,\omega)|D|^{-1}|g(x^p)-g(x^*)| \tag{4.19}$$

令

$$A(\omega)=\frac{1}{\omega}(1-|1-\omega|)|D|-|G|,$$

$$B_k(r,\omega)=\frac{1}{\omega}(|D|-r|L_k|),$$

$$C_k(r,\omega)=\frac{1}{\omega}[|1-\omega||D|+(\omega-r)|L_k|+\omega|U_k|],$$

$$(k=1,2,\cdots,\alpha)$$

显然，

$$A(\omega)=B_k(r,\omega)-C_k(r,\omega)\qquad(k=1,2,\cdots,\alpha)$$

均为$A(\omega)$的弱正则分裂，并且当r,ω在由(4.10)所确定的范围内时，$A(\omega)$为M-矩阵。依引理2，成立

$$\rho(|D|\mathscr{L}(r,\omega)|D|^{-1})=\rho(\mathscr{L}(r,\omega))<1.$$

因此，

$$\lim_{p\to\infty}g(x^p)=g(x^*).$$

$g:R^n\to R^n$的同胚性质即保证

$$\lim_{p\to\infty} x^p = x^*.$$

至此，定理证毕.

致谢 审稿人提出的宝贵意见，使得本文增色不少. 在此，表示衷心的感谢!

参 考 文 献

[1] D. P. O'Leay and R. E. White, Multisplittings of matrices and parallel solution of linear systems, *SIAM J. Alg. Disc. Meth.*, 6(1985), 630—640.

[2] R. E. White, A nonlinear parallel algorithm with application to the Stelan problem, *SIAM J. Numer. Anal.*, 23(1986),639—652.

[3] Z. Z. Bai, *Parallel nonlinear AOR method and its convergence*, to appear (1994).

[4] R. E. White, Parallel algorithms for nonlinear problems, *SIAM J. Alg.Disc. Mesh.*, 7(1986),137—149.

[5] R. E. White, An enthalpy formulation of the Stefan problem, *SIAM J. Numer. Anal.*, 19(1982), 1129—1157.

[6] R. E. White, The binary alloy solidification problem: existence, uniquess and numerical approximation, *SIAM J. Numer. Anal.*, 22(1985), 205—244.

[7] *J. M.* Ortega and W. C. Rheinboldt, *Iterattve Solution of Nonlinear Equations in Several Variables*, Academic New York Press (1970).

Parallel Multisplitting AOR Method for Solving a Class of System of Nonlinear Algebraic Equations

Bai Zhongzhi

(*Institute of Mathematics*, *Fudan University*, *Shanghai* 200433)

Abstract

In this paper, we propose a class of parallel multisplitting AOR method for solving large-scale system of nonlinear equations $A\varphi(x)+B\psi(x)=b$. Under certain conditions, we prove the existence and uniqueness of the solution of this system of nonlinear equations and set up the global convergence theory of our new method.

Key words system of nonlinear algebraic equations, parallel method, relaxation, H-matrix