上海科学技术大学

博士学位论文

SHANGHAI UNIVERSITY
OF SCIENCE AND TECHNOLOGY
DOCTORAL THESIS

研究生处

大型代数方程组的并行迭代算法

（摘要）

姓　名：白中治

专　业：计算数学

导　师：郭本瑜、王德人（教授）

入学年月：一九九〇年九月

一九九三年三月

本文由两部分组成，第一部分对于大型线性代数方程组

$$Ax=b,\quad A\in L(R^n)\ 非奇异,\quad x,b\in R^n.$$

运用矩阵多重分裂的概念，采用矩阵多重分裂的松弛加速技术，我们建立了并行矩阵多分裂松弛算法的一般框架；基于滞后信息充分利用的原则，我们设计了一类能够充分发挥MIMD系统并行计算之功能的异步并行矩阵多分裂松弛加速型算法；当系数矩阵$A\in L(R^n)$为块状对称正定矩阵时，利用分块多层迭代的思想，我们构造了一类并行多层迭代算法。对于这些新算法的收敛性理论，我们进行了深入细致的研究，同时，对于收敛速度，亦做了详细的估计。

第二部分对于大型非线性代数方程组

$$F(x)=0,\quad F:\mathscr{D}\subset R^n\to R^n.$$

在非线性多分裂的意义下，利用非线性多重分裂的松弛加速技术，我们建立了一类并行非线性多分裂松弛加速型算法，其中包含了实用而有效的多分裂AOR-Newton法，多分裂AOR-Chord以及多分裂AOR-Steffensen方法等。在适当的条件下，我们研究了新算法的局部收敛性、单调收敛性和全局收敛性，得到了比较深刻的理论结果。大量的数值实验表明，我们的算法是可行、有效的，所有的数值结果均与

理论完全相符；基于滞后信息充分利用的准则，我们设计了一类灵活而又高效的、适用于MIMD系统的异步并行非线性多分裂松弛加速型迭代法，其中亦包含了具有很高应用价值的异步并行非线性多分裂AOR-Newton法，AOR-Chord法以及AOR-Steffensen方法等。关于新的异步非线性多分裂算法模型的收敛性理论，我们进行了深入细致的讨论，同时，也估计了它们的收敛速度。

关键词：线性代数方程组，矩阵多分裂，L-矩阵，M-矩阵，H-矩阵，对称正定矩阵，多层迭代，预处理矩阵，条件数，非线性代数方程组，非线性多分裂，同步并行，异步并行，松弛加速，局部收敛性，单调收敛性，全局收敛性，收敛速度。

第一章 绪论

§1.引言

随着现代科学技术的迅速发展，并行计算方法的设计与理论分析，已成为当前国际计算数学界所面临的一项艰巨任务，同时，也是计算数学研究的热点。依照“以大化小”、“分而治之”的基本准则，O'Leary和White[1]提出了矩阵多重分裂的概念，即

定义1.1 给定矩阵$A\in L(R^n)$，如果

i) $A=M_i-N_i$, $\det(M_i)\neq 0$, $i=1,2,\cdots,\alpha$;

ii) $\sum_{i=1}^{\alpha}E_i=I$ ($E_i\in L(R^n)$ $(i=1,2,\cdots,\alpha)$ 为非负对角矩阵，$I\in L(R^n)$为单位矩阵)

则称三元组集(M_i,N_i,E_i), $i=1,2,\cdots,\alpha$是矩阵A的一个多分裂。这里，$\alpha(\alpha\leq n)$为给定的正整数，通常等效于处理机的台数，而E_i $(i=1,2,\cdots\alpha)$称作权矩阵。

由此概念，通过对大型线性代数方程组

(1.1) $Ax=b$, $A\in L(R^n)$非奇异, $x,b\in R^n$

进行系统分解，O'Leary和White设计了求解(1.1)的结构新颖、计算简便的并行矩阵多分裂迭代算法：

算法1.1：给定初始向量$x^0\in R^n$，对$p=0,1,2,\cdots$，计算

$$\begin{cases} x^{i,p}=M_i^{-1}(N_i x^p+b),\ i=1,2,\cdots,\alpha \\ x^{p+1}=\sum_{i=1}^{\alpha}E_i x^{i,p} \end{cases}$$

如果引进矩阵

$$H=\sum_{i=1}^{\alpha}E_iM_i^{-1}N_i,\quad G=\sum_{i=1}^{\alpha}E_iM_i^{-1}, \tag{1.2}$$

则该算法可等价地表述为

$$x^{p+1}=Hx^p+Gb. \tag{1.3}$$

显然，在算法1.1中，对于固定的p，$x^{i,p}$对于各个不同的i可以相互独立地计算，从而可同步并行地在SIMD系统上实现；通过权矩阵的合理选取，可以协调未知量之间的重复与不重复计算，平衡各处理机之间的任务分配；相应于E_i的零对角元的$x^{i,p}$的元素可不必计算，从而可节省工作量，减少冗余时间，提高算法的并行效率；适当选取α个分裂，可使得每个$x^{i,p}$由求解相应的低维线性方程组而得到。在实际应用中，通过权矩阵与α个分裂的选取，可以改善算法的收敛性态，提高算法的并行计算效率。

借鉴单分裂迭代法的构造技巧，利用较为特殊的矩阵多重分裂，Frommer和Mayer[2]建立了求解(1.1)的并行矩阵多分裂逐次超松弛(SOR)迭代算法，Wang Deren[3]提出了多分裂松弛加速技术，进一步推广并改进了该算法。

关于上述算法的理论研究和数值分析，O'Leary、White[1]，

White([4]-[5])，Neumann、Plemmons[6]，Frommer、Mayer[2]，Elsner[7]以及Wang Deren[3]等都已做了许多深入细致的工作。但所有这些分析均仅囿于系数矩阵$A\in L(R^n)$是H-矩阵类的情形，这在一定程度上限制了算法的适用范围。对于对称正定矩阵类，目前仅有的结果也只是White([4]-[5])对于具特殊结构的问题，采用特殊的权矩阵和多重分裂所得到的，因而，尚有待于进一步地推广和完善。对于更广泛的矩阵类(如L-矩阵)，有关并行矩阵多分裂迭代算法的收敛性问题，还需要进行深入研究。

Frommer[8]将矩阵多分裂的思想成功地推广到非线性映射，为设计求解大型非线性代数方程组

$$F(x)=0,\quad F:\mathcal{D}\subset R^n\to R^n \tag{1.4}$$

在多分裂意义下的并行迭代算法提供了条件。

除问题的非线性性质外，Frommer所建立的非线性多分裂迭代算法与算法1.1有完全相似的特点。鉴于非线性问题所固有的复杂性，在实际应用中，权矩阵与非线性多重分裂的合理选取将更加困难。特别是算法中涉及到的α个低维非线性方程组的求解，计算工作量仍十分可观。

所有这些研究工作，均是基于SIMD系统，计算是在同步环境下进行的。在矩阵多分裂的意义下，针对MIMD系

统的异步并行算法，目前讨论还很少。

Bru．Elsner，Neumann[9]将矩阵多分裂的概念与Chazan．Miranker[10]混乱迭代的思想有机地结合为一体，所提出的求解大型线性代数方程组(1.1)的两种并行混乱迭代模型。由于限制较多，信息交换不很灵活自由，致使多处理系统的实际效能并未能得到充分发挥。

苏仰锋[11]等人的进一步研究结果是引人注目的。作者在分析Bru等人异步模型的基础上，建立了以尽可能减少多处理机之间的同步等待时间，提高算法的并行计算效率为宗旨的异步多分裂算法模型。

针对上述问题，在本文的第二章，通过引进参数矩阵，我们建立了求解大型线性代数方程组(1.1)的并行矩阵多分裂松弛算法的一般框架。相应于参数矩阵的不同选取，它概括了几乎所有的并行矩阵多分裂松弛算法，从而形成了一组广泛的并行矩阵多分裂松弛算法序列，它既方便于应用，又可改善算法的收敛性质，并且在对参数矩阵和多重分裂的适当限制下，当系数矩阵为H-矩阵时，我们建立了这类算法收敛与不收敛的充分条件。进而，当系数矩阵为L-矩阵时，我们还得到了保证算法收敛的充分必要条件。从而，使算法的收敛理论更趋于完善。

在本文的第三章，根据MIMD系统的具体特性，采用滞后信息充分利用的思想，基于矩阵多重分裂的概念，我们设计了一类求解线性代数方程组(1.1)的新的异步并行多分裂松弛算法。该算法是Bru，苏仰锋等人异步迭代模型的推广与改进，具有计算方便，通讯充分自由灵活等特点，因而，更适用于MIMD系统。关于新算法的收敛理论，我们在与同步多分裂松弛算法相同的条件下，证明了它的收敛性；同时，亦给出了收敛速度的估计。

本文第四章，对于由二阶自共轭椭圆型边值问题，经有限元离散所产生的具对称正定系数矩阵的大型线性代数方程组，我们运用分块多层迭代的思想，构造了具良好条件数的关于系数矩阵的预处理矩阵，进而，建立了求解这类线性方程组的并行多层迭代算法。对于该算法的收敛性态，即预处理矩阵的条件数，我们做了详细的理论分析，同时，亦细致地讨论了新算法的具体实施方案，并估计了计算工作量。结果表明，新算法的收敛速度不仅与剖分网格的大小及网格的层数无关，并且其计算工作量以网格步长的线性函数为界。

在非线性多分裂的意义下，利用非线性多分裂的松弛加速技术，研究求解非线性方程组(1.4)的非线性多分裂松

弛加速型算法，是本文第五章的主要内容。我们建立了一类广泛的并行非线性多分裂松弛加速型算法，其中包含了多分裂AOR-Newton法，多分裂AOR-弦法以及多分裂AOR-Steffensen方法等颇为实用的非线性多分裂松弛算法。这些算法亦是Schechter[12]和冯果忱[13]结果在多重分裂意义下的推广与改进，且适用于超级多处理机系统，为求解大型非线性方程组提供了新型的计算方案。关于新算法的可行性于收敛性问题，我们建立了具一定概括性的一般的局部收敛性定理，特别，对于多分裂AOR-Newton法的收敛性，进行了细致的分析，同时，还得到了算法的敛速估计。最后，通过一系列数值例子，说明我们的算法是可行的，所得数值结果亦是满意的。

一个非线性迭代算法的全局收敛性是引人注目的，但是，要对一般的光滑映象去研究迭代算法的全局收敛问题，是十分困难的。在本文的第六章，对于一些具有特殊性质的函数类，我们进一步研究了第五章所建立的并行非线性多分裂松弛算法的单调收敛性和全局收敛性。我们证明了这类算法具有双侧收敛于所论方程的最大解和最小解的性质，以及松弛参数的不同选取导致算法不同的收敛速度。作为上述理论的特例，对于线性代数方程组(1.1)，我们研

究了矩阵多分裂迭代算法的单调收敛性理论，同时，也建立了单调性的比较定理。最后，我们列举了一类非线性边值问题，并对此进行数值实验，结果与所述理完全相符。

在本文的第七章，同样运用非线性多分裂的概念，采用非线性多重分裂的松弛加速技术和滞后信息充分利用的思想，我们提出了一类求解非线性方程组(1.4)的异步并行松弛加速迭代算法模型。该算法模型同时兼顾非线性多重分裂的优点和多处理机系统的具体特性，具有计算简便、通讯充分灵活自由等良好性质，因此，能够更加充分地发挥MIMD系统的实际计算效能。随着松弛参数的不同选取，不仅可改善新算法模型的收敛性态，而且还可得到几乎所有已知的和许多新的异步非线性多分裂松弛算法。同时，我们还讨论了具有很高应用价值的异步非线性多分裂AOR-Newton、-Chord、-Steffensen等算法，从而使得这类算法在具体实施中更为方便、有效。在一定条件下，我们建立了新算法模型的收敛性理论，同时，也估计了其收敛速度。

§2.预备知识

本节系统地列出了以后各章节中所要用到的符号、基本概念和已知结论。

第一部分　大型线性代数方程组

在矩阵多分裂的意义下，我们设计了具很强概括性的求解大型线性代数方程组(1.1.1)的同步与异步松弛迭代算法，同时，当(1.1.1)的系数矩阵为块状对称正定矩阵时，我们构造了一类并行多层迭代算法。对于这些新算法的收敛性理论，我们进行了深入细致的研究；同时，对于它们的收敛速度，亦做了详细的估计.

第二章　矩阵多重分裂与并行松弛算法

§1. 算法的建立

首先，我们将数集$\{1,2,\cdots,n\}$分为α个非空子集合$J_i(i=1,2,\cdots,\alpha)$，即$J_i\subset\{1,2,\cdots,n\}$，$i=1,2,\cdots,\alpha$，且满足

(1.1) $$\bigcup_{i=1}^{\alpha} J_i=\{1,2,\cdots,n\}。$$

考虑非奇异矩阵$A=(a_{mj})\in L(R^n)$。定义矩阵

$$\begin{cases} D(A)=diag(A), \quad det(D(A))\neq 0 \\ L_i(A)=(\mathcal{L}_{mj}^{(i)}(A)), \quad \mathcal{L}_{mj}^{(i)}(A)=\begin{cases} l_{mj}^{(i)}(A), & 若\ j<m\ 且\ (m,j)\in J_i \\ 0, & 其它 \end{cases} \\ U_i(A)=(\mathcal{U}_{mj}^{(i)}(A)), \quad \mathcal{U}_{mj}^{(i)}(A)=\begin{cases} u_{mj}^{(i)}(A), & 若\ j>m\ 且\ (m,j)\in J_i \\ 0, & 其它 \end{cases} \\ W_i(A)=(\mathcal{W}_{mj}^{(i)}(A)), \quad \mathcal{W}_{mj}^{(i)}(A)=\begin{cases} 0, & 若\ j=m\ 且\ (m,j)\in\{1,2,\cdots,n\} \\ w_{mj}^{(i)}(A), & 其它 \end{cases} \\ i=1,2,\cdots,\alpha \quad (\alpha\leq n). \end{cases}$$

满足

$$(1.2)\qquad A=D(A)-L_i(A)-U_i(A)-W_i(A)\equiv D(A)-B(A),\ i=1,2,\cdots,\alpha。$$

显见，对每个 $i\in\{1,2,\cdots,\alpha\}$，$L_i(A)\in L(R^n)$ 为严格下三角矩阵，$U_i(A)\in L(R^n)$ 为严格上三角矩阵，而 $W_i(A)\in L(R^n)$ 为对角元是零的矩阵。

取权矩阵 $E_i=diag(e_1^{(i)},e_2^{(i)},\cdots,e_n^{(i)})\in L(R^n)(i=1,2,\cdots\alpha)$ 为

$$(1.3)\qquad e_m^{(i)}=\begin{cases}0, & 若\ m\notin J_i\\ e_m^{(i)}\geqslant 0, & 若\ m\in J_i\end{cases}$$

这时，分裂(1.3)连同权矩阵 $E_i(i=1,2,\cdots,\alpha)$ 即构成矩阵 $A\in L(R^n)$ 的一个特殊但非常实用的多分裂，记其为 $(D(A)-L_i(A),D(A)-U_i(A),W_i(A),E_i)\ (i=1,2,\cdots,\alpha)$.

利用上面给出的矩阵多分裂的概念，我们现构造如下的广义并行矩阵多分裂松弛算法。

算法1.1：给定初始向量 $x^0\in R^n$，对 $p=0,1,2,\cdots$ 计算

$$\begin{cases}x_m^{p+\frac{1}{2},i}=\dfrac{1}{a_{mm}}\Big[r_m^{(1)}\sum\limits_{\substack{j<m\\(m,j)\in J_i}}l_{mj}^{(i)}(A)x_j^{p+\frac{1}{2},i}+(\omega_m^{(1)}-r_m^{(1)})\sum\limits_{\substack{j<m\\(m,j)\in J_i}}l_{mj}^{(i)}(A)x_j^p+\omega_m^{(1)}\Big(\sum\limits_{\substack{j>m\\(m,j)\in J_i}}U_{mj}^{(i)}(A)x_j^p\\ \qquad+\sum\limits_{\substack{j=1\\j\neq m}}^{n}W_{mj}^{(i)}(A)x_j^p\Big)+\omega_m^{(1)}b_m\Big]+(1-\omega_m^{(1)})x_m^p,\quad m\in J_i\\ x_m^{p+1,i}=\dfrac{1}{a_{mm}}\Big[r_m^{(2)}\sum\limits_{\substack{j>m\\(m,j)\in J_i}}U_{mj}^{(i)}(A)x_j^{p+1,i}+(\omega_m^{(2)}-r_m^{(2)})\sum\limits_{\substack{j>m\\(m,j)\in J_i}}U_{mj}^{(i)}(A)x_j^{p+\frac{1}{2},i}+\omega_m^{(2)}\Big(\sum\limits_{\substack{j<m\\(m,j)\in J_i}}l_{mj}^{(i)}(A)x_j^{p+\frac{1}{2},i}\\ \qquad+\sum\limits_{\substack{j=1\\j\neq m}}^{n}W_{mj}^{(i)}(A)x_j^{p+\frac{1}{2},i}\Big)+\omega_m^{(2)}b_m\Big]+(1-\omega_m^{(2)})x_m^{p+\frac{1}{2},i},\quad m\in J_i\\ x_m^{p+1}=\sum\limits_i e_m^{(i)}x_m^{p+1,i}\\ r_m^{(1)},r_m^{(2)}\geqslant 0,\ \omega_m^{(1)},\omega_m^{(2)}>0,\ m=1(1)n;\ i=1,2,\cdots,\alpha\end{cases}$$

算法1.1是一个实用的同步并行矩阵多分裂多参数块松弛迭代法。对于每个i，我们只需求解两个低维（维数等于集合J_i所含元素的个数）的三角线性代数方程组。

在算法1.1中，如果特殊地选取参数矩阵

$$\begin{cases} \Omega_m = \mathrm{diag}(\omega_1^{(m)}, \omega_2^{(m)}, \cdots, \omega_n^{(m)}) \\ R_m = \mathrm{diag}(r_1^{(m)}, r_2^{(m)}, \cdots, r_n^{(m)}) \end{cases}, \quad m=1,2,$$

即可得到所有已知的并行矩阵多分裂松弛算法(见[2],[3],[5],[64])，亦可产生许多新的并行矩阵多分裂松弛算法。

若再用参数$\varphi_j (j=1(1)n)$进一步松弛算法1.1,则可得算法1.1的如下外插形式：

算法1.2：给定初始向量$x^0 \in R^n$,对$p=0,1,2,\cdots$，计算

$$x_m^{p+1} = \varphi_m \sum_i e_m^{(i)} x_m^{p+1,i} + (1-\varphi_m) x_m^p$$

其中

$$\begin{cases} x_m^{p+\frac{1}{2},i} = \dfrac{1}{a_{mm}}\Big[r_m^{(1)} \sum\limits_{\substack{j<m \\ (m,j)\in J_i}} l_{mj}^{(i)}(A) x_j^{p+\frac{1}{2},i} + (\omega_m^{(1)} - r_m^{(1)}) \sum\limits_{\substack{j<m \\ (m,j)\in J_i}} l_{mj}^{(i)}(A) x_j^p + \omega_m^{(1)}\Big(\sum\limits_{\substack{j>m \\ (m,j)\in J_i}} u_{mj}^{(i)}(A) x_j^p \\ \qquad + \sum\limits_{\substack{j=1 \\ j\neq m}} W_{mj}^{(i)}(A) x_j^p\Big) + \omega_m^{(1)} b_m\Big] + (1-\omega_m^{(1)}) x_m^p, \quad m\in J_i \\ x_m^{p+1,i} = \dfrac{1}{a_{mm}}\Big[r_m^{(2)} \sum\limits_{\substack{j>m \\ (m,j)\in J_i}} u_{mj}^{(i)}(A) x_j^{p+1,i} + (\omega_m^{(2)} - r_m^{(2)}) \sum\limits_{\substack{j>m \\ (m,j)\in J_i}} u_{mj}^{(i)}(A) x_j^{p+\frac{1}{2},i} + \omega_m^{(2)}\Big(\sum\limits_{\substack{j<m \\ (m,j)\in J_i}} l_{mj}^{(i)}(A) x_j^{p+\frac{1}{2},i} \\ \qquad + \sum\limits_{\substack{j=1 \\ j\neq m}}^{n} W_{mj}^{(i)}(A) x_j^{p+\frac{1}{2},i}\Big) + \omega_m^{(2)} b_m\Big] + (1-\omega_m^{(2)}) x_m^{p+\frac{1}{2},i}, \quad m\in J_i \\ r_m^{(1)}, r_m^{(2)} \geq 0, \ \omega_m^{(1)}, \omega_m^{(2)} > 0, \ m=1(1)n; \ i=1,2,\cdots,\alpha \end{cases}$$

引进参数矩阵

$$\Phi=diag(\varphi_1,\varphi_2,\cdots,\varphi_n)$$

则类似地，相应于参数矩阵 $R_m,\Omega_m(m=1,2)$ 及 Φ 的不同选取，我们亦可由算法1.2得到许多特殊而实用的多分裂松弛算法。

§2.算法的收敛性分析—H矩阵类的情形

对于H-矩阵类，我们建立算法1.1,1.2的收敛理论。

定理2.1 设 $A\in L(R^n)$ 为H-矩阵。对任意矩阵 $G\in\Omega(A)$，$(D(G)-L_i(G),D(G)-U_i(G),W_i(G),E_i)$，$i=1,2,\cdots,\alpha$ 为 G 的一个多分裂并且满足

$$\langle G\rangle=|D(G)|-|L_i(G)|-|U_i(G)|-|W_i(G)|=|D(G)|-|B(G)|,\ i=1,2,\cdots\alpha$$

则对任何初始向量 $x^0\in R^n$，当参数矩阵 $R_m,\Omega_m(m=1,2)$ 满足

(2.1) $$0\le r_j^{(m)}\le\omega_j^{(m)},\quad 0<\omega_j^{(m)}<2/(1+\rho(|J(A)|)),\ j=1(1)n,\ m=1,2.$$

时，由算法1.1产生的求解线性方程组

(2.2) $$Gx=b$$

的迭代序列收敛到该方程组的唯一解。其中集合

$$\Omega(A)=\{G=(g_{mj})\in L(R^n)\,|\,|g_{mj}|=|a_{mj}|,\ m,j=1(1)n\}$$

表示与矩阵A相关联的等模矩阵的集合，$\langle G\rangle$，$|G|$ 分别表示矩阵G的比较矩阵和绝对值，而 $\rho(G)$ 则用来表示矩阵G的谱半径。下同。

证明：略。

定理2.2　设定理2.1的条件满足，则对任何初始向量$x^0 \in R^n$，当参数矩阵$R_m,\Omega_m(m=1,2)$满足(2.1)而更满足

$$\begin{cases} 0<\varphi_j<2/(1+\|H(\Omega_1)\|_\infty\|H(\Omega_2)\|_\infty), & j=1(1)n \\ H(\Omega_m)=|I-\Omega_m|+\Omega_m\rho(|J(A)|), & m=1,2 \end{cases}$$

时，由算法1.2产生的求解线性方程组(2.7)的迭代序列收敛到该方程组的唯一解。

证明：略。

§3.算法的收敛性分析——L-矩阵类的情形

首先，我们给出定理2.1和定理2.2的逆命题。

定理3.1　设$A\in L(R^n)$为L-矩阵，但非M-矩阵。$(D(A)-L_i(A), D(A)-U_i(A), W_i(A), E_i), i=1,2,\cdots,\alpha$为矩阵A的一个多分裂且满足

(3.1)　　$L_i(A)\geq 0,\ U_i(A)\geq 0,\ W_i(A)\geq 0,\ i=1,2,\cdots\alpha$

则对任何初始向量$x^0\in R^n$，当

a)　参数矩阵$R_m,\Omega_m(m=1,2)$满足

(3.2)　　$0\leq r_j^{(m)}\leq\omega_j^{(m)}\leq 1,\ j=1(1)n,\ m=1,2$

时，存在适当小的$R_m(m=1,2)$，使得由算法1.1产生的求解方程组(1.1.1)的迭代序列不收敛；

b)　参数矩阵$R_m,\Omega_m(m=1,2)$满足(3.2)而更满足

$0 \leq \varphi_j \leq 1, \quad j=1(1)n$

时，存在适当小的$R_m(m=1,2)$，使得由算法1.2产生的求解方程组(1.1.1)的迭代序列不收敛。

证明：略。

定理3.1亦可等价地叙述为如下形式：

定理3.2　设$A\in L(R^n)$为L-矩阵。$(D(A)-L_i(A), D(A)-U_i(A), W_i(A), E_i)$，$i=1,2,\cdots,\alpha$为矩阵A的一个多分裂并且满足(3.1). 若下列条件之一满足，则A必为M-矩阵.

a)　对$0\leq R_m\leq \Omega_m\leq I$ $(det(\Omega_m)\neq 0)$, $m=1,2$中适当小的R_m和某个$\Omega_m(m=1,2)$，由算法1.1产生的求解线性方程组(1.1.1)的迭代序列收敛；

b)　对$0\leq R_m\leq \Omega_m\leq I$ $(det(\Omega_m)\neq 0)$, $m=1,2$，$0\leq \Phi\leq I$ $(det(\Phi)\neq 0)$中适当小的R_m和某$\Omega_m(m=1,2)$, Φ，由算法1.2产生的求解线性方程组(1.1.1)的迭代序列收敛。

当系数矩阵$A\in L(R^n)$为L-矩阵时，定理2.1，定理2.2，定理3.1，和定理3.2显然蕴含了保证算法1.1，算法1.2收敛的如下充分必要条件。

定理3.3　设定理3.2的条件满足，则$\rho(J(A))<1$当且仅当

a)　对任何初始向量$x^0\in R^n$，由算法1.1产生的求解线性方程组(1.1.1)的迭代序列，当参数矩阵R_m和$\Omega_m(m=1,2)$满足

$$0 \le R_m \le \Omega_m \le I \ (\det(\Omega_m) \ne 0), \ m=1,2$$

时，均收敛到其唯一解；

b) 对任何初始向量 $x^0 \in R^n$，当参数矩阵 $R_m, \Omega_m \ (m=1,2)$ 和 Φ 满足

$$0 \le R_m \le \Omega_m \le I \ (\det(\Omega_m) \ne 0), \ m=1,2$$

$$0 \le \Phi \le I \ (\det(\Phi) \ne 0)$$

时，由算法1.2产生的求解线性方程组(1.1.1)的迭代序列收敛到其唯一解。

第三章 矩阵多重分裂与异步并行松弛算法

§1. 异步算法的建立

仍分裂数集 $\{1,2,\cdots,n\}$ 为 $\alpha \ (\alpha \le n)$ 个非空子集合 $J_i \ (i=1,2,\cdots,\alpha)$ ($J_1, J_2, \cdots, J_\alpha$ 之间可以相交)，满足(2.1.1)。考虑非奇异矩阵 $A=(a_{mj}) \in L(R^n)$，记 $D=\mathrm{diag}(A)$。对 $i=1,2,\cdots,\alpha$，分别取 $L_i=(\mathcal{L}_{mj}^{(i)})$, $U_i=(\mathcal{U}_{mj}^{(i)}) \in L(R^n)$ 为

$$\begin{cases} \mathcal{L}_{mj}^{(i)} = \begin{cases} l_{mj}^{(i)}, & \text{若 } m,j \in J_i \text{ 且 } m>j \\ 0, & \text{其它情形} \end{cases}, & m,j=1(1)n \\ \mathcal{U}_{mj}^{(i)} = \begin{cases} u_{mj}^{(i)}, & \text{若 } m \ne j \\ 0, & \text{其它情形} \end{cases}, & m,j=1(1)n \end{cases}$$

而$E_i=diag(e_1^{(i)},e_2^{(i)},\cdots,e_n^{(i)})\in L(R^n)$满足(2.1.3)。显然，这里$L_i\in L(R^n)$为严格下三角矩阵，$U_i\in L(R^n)$为对角元是零的矩阵，而$E_i\in L(R^n)$为非负对角矩阵。如果

1) $A=D-L_i-U_i,\ i=1,2,\cdots,\alpha$；

2) $\det(D)\neq 0$；

3) $\sum_i E_i=I$（$I\in L(R^n)$为单位矩阵）.

则称三元组集$(D-L_i,U_i,E_i),i=1,2,\cdots,\alpha$是矩阵$A\in L(R^n)$的一个多分裂。

假设所考虑的多处理机系统由α个CPU组成，我们引进以下符号：

i) 对$\forall i\in\{1,2,\cdots,\alpha\}$，$\forall p\in N_0:=\{0,1,2,\cdots\}$，用$J^{(i)}=\{J_i(p)\}_{p\in N_0}$表示集合$J_i$的子集合（可以为空子集$\phi$）；

ii) 对$\forall m\in\{1,2,\cdots,n\}$，$S^{(i)}=\{S_1^{(i)}(p),S_2^{(i)}(p),\cdots,S_n^{(i)}(p)\}_{p\in N_0}$表示$n$个无穷序列。

集合$J^{(i)}$和$S^{(i)}(i=1,2,\cdots,\alpha)$具如下性质：

a) 对$\forall i\in\{1,2,\cdots,\alpha\}$，$\forall m\in\{1,2,\cdots,n\}$，集合$\{p\in N_0\mid m\in J_i(p)\}$为无限；

b) 对$\forall p\in N_0$，$\bigcup_{i=1}^{\alpha}J_i(p)\neq\phi$；

c) 对$\forall i\in\{1,2,\cdots,\alpha\}$，$\forall m\in\{1,2,\cdots,n\}$，$\forall p\in N_0$，$S_m^{(i)}(p)\leq p$；

d) 对$\forall i\in\{1,2,\cdots,\alpha\}$，$\forall m\in\{1,2,\cdots,n\}$，$\lim_{p\to\infty}S_m^{(i)}(p)=\infty$.

今对于大型线性代数方程组(1.1.1)，考虑如下的异步并行矩阵多分裂AOR算法：

算法1.1：给定初始近似 $x^0\in R^n$，并且假定已获得近似序列 $x^0, x^1, \cdots, x^p$，则第 $p+1$ 次近似 $x^{p+1}=(x_1^{p+1}, x_2^{p+1}, \cdots, x_n^{p+1})^T$ 的计算过程如下：

(I) 逐次地求解关于 $\hat{x}_m^{p,i}(m\in J_i(p), m=1(1)n, i=1,2,\cdots,\alpha)$ 的线性方程组

$$(1.1)\quad \begin{cases} a_{mm}\hat{x}_m^{p,i}-\sum\limits_{\substack{m>j\\ m,j\in J_i}} l_{mj}^{(i)}\tilde{x}_j^{p,i}-\sum\limits_{\substack{j=1\\ j\neq m}} U_{mj}^{(i)}x_j^{s_j^{(i)}(p)}=b_m, \quad 若 m\in J_i(p) \\ m=1,2,\cdots,n;\quad i=1,2,\cdots,\alpha \end{cases}$$

其中 $\tilde{x}^{p,i}=(\tilde{x}_1^{p,i}, \tilde{x}_2^{p,i}, \cdots, \tilde{x}_n^{p,i})^T$ 为

$$(1.2)\quad \begin{cases} \tilde{x}_m^{p,i}=\begin{cases} r\hat{x}_m^{p,i}+(1-r)x_m^{s_m^{(i)}(p)}, & 若 m\in J_i(p) \\ x_m^{s_m^{(i)}(p)}, & 若 m\notin J_i(p) \end{cases}, m=1(1)n \\ i=1,2,\cdots,\alpha \end{cases}$$

(II) 计算 $x^{p,i}=(x_1^{p,i}, x_2^{p,i}, \cdots, x_n^{p,i})^T$，依照

$$(1.3)\quad \begin{cases} x_m^{p,i}=\begin{cases} \frac{\omega}{r}\tilde{x}_m^{p,i}+(1-\frac{\omega}{r})x_m^{s_m^{(i)}(p)}, & 若 m\in J_i(p) \\ x_m^p, & 若 m\notin J_i(p) \end{cases}, m=1(1)n \\ i=1,2,\cdots,\alpha \end{cases}$$

(III) 计算

$$x_m^{p+1}=\sum_{i=1}^{\alpha} e_m^{(i)}x_m^{p,i}, \quad m=1,2,\cdots,n$$

这里 r, ω 分别称作松弛因子和加速因子。

容易看出，对于 $\forall m\in\{1,2,\cdots,n\}$, $i\in\{1,2,\cdots,\alpha\}$，当

$$(1)\quad \begin{cases} J_i=\{1,2,\cdots,n\} \\ \forall p\in N_0,\ J_i(p)=J_i,\ s_m^{(i)}(p)=p \end{cases}$$

时，算法1.1退化为熟知的同步并行矩阵多分裂AOR算法[3]；

$$(2)\quad \begin{cases} J_i=\{1,2,\cdots,n\} \\ \forall p\in N_0,\ (J_i(p)=J_i)\vee(J_i(p)=\phi)=True,\ |N_m(p)|=1 \end{cases}$$

时，算法1.1变为Bru, Elsner和Neumann[9]的并行混乱迭代模型.

$$(3)\quad \begin{cases} J_i=\{1,2,\cdots,n\} \\ \forall p\in N_0,\ (J_i(p)=J_i)\vee(J_i(p)=\phi)=True,\ s_m^{(i)}(p)=s_i(p)\in R' \end{cases}$$

时，算法1.1成为已知的异步矩阵多分裂AOR算法[17].

另外，相应于参数对 (r,ω) 的不同选取 $(0,1)$, $(1,1)$, 和 (ω,ω) 等，即可由算法1.1得到异步矩阵多分裂Jacobi, Gauss-Seidel和SOR等许多特殊而实用的迭代算法。

如果用参数 β 进一步松弛算法1.1，则可将它改进为如下形式：

算法1.2　给定初始近似 $x^0\in R^n$，并且假设已获得近似序列 $x^0, x^1, \cdots, x^p$，则第 $p+1$ 次近似 x^{p+1} 分别由(1.1)-(1.3)以及

$$\bar{x}_m^{p,i}=\begin{cases} \beta x_m^{p,i}+(1-\beta)x_m^{s_m^{(i)}(p)}, & 若 m\in J_i(p) \\ x_m^p, & 若 m\notin J_i(p) \end{cases},\quad m=1(1)n,\ i=1,2,\cdots,\alpha$$

和

$$x_m^{p+1}=\sum_{i=1}^{\alpha} e_m^{(i)}\bar{x}_m^{p,i},\quad m=1,2,\cdots,n$$

得到。

类似地，相应于参数r,ω和β的特殊选取，我们亦可由算法1.2得到许多实用而有效的异步矩阵多分裂松弛算法。限于篇幅，我们不再一一例举。

§2. 几个引理

引理2.1　给定$\bar{x}^*\in R^n$及$\{\bar{x}^t\}_{t=0}^{p}\subset R^n(\forall p\in N_0)$。如果当$0\le t\le p$时，对某正数$\delta$和正向量$v=(v_1,v_2,\cdots,v_n)^T\in R^n$，成立

$$|\bar{x}^t-\bar{x}^*|\le\delta v,$$

则当$s_m^{(i)}(p)\le p(m=1(1)n,i=1,2,\cdots,\alpha)$时，恒有

$$|\bar{x}^{s^{(i)}(p)}-\bar{x}^*|\le\delta v,$$

其中$\bar{x}^{s^{(i)}(p)}=(\bar{x}_1^{s_1^{(i)}(p)},\bar{x}_2^{s_2^{(i)}(p)},\cdots,\bar{x}_n^{s_n^{(i)}(p)})^T$。

证明：略。

现引进非负数列$\{i_m^p\}_{p\in N_0}$和$\{j_m^p\}_{p\in N_0}(m=1(1)n)$分别为

$$i_m^p=\sum_{i\in N_m(p)} e_m^{(i)},\quad j_m^{(p)}=\sum_{i\notin N_m(p)} e_m^{(i)},\ p=0,1,2,\cdots.$$

引理2.2　设$\xi_m>0\ (m=1(1)n)$。如果序列$\{\varepsilon_m^p\}_{p\in N_0}(m=1(1)n)$满足

$$|\varepsilon_m^{p+1}|\le i_m^p\xi_m+j_m^p|\varepsilon_m^p|,\quad p=0,1,2,\cdots;\ m=1(1)n,$$

则对任何非负整数 $g \le p-1$，成立

$$|\varepsilon_m^{p+1}| \le (1-\prod_{k=p-g-1}^{p} j_m^k)\xi_m + \prod_{k=p-g-1}^{p} j_m^k |\varepsilon_m^{p-g-1}|,\ m=1(1)n.$$

证明：略。

引理2.3 设

$$j_m^{(0)} = \prod_{p=0}^{m_0-1} j_m^p,\quad j_m^{(\ell+1)} = \prod_{p=m_\ell}^{m_{\ell+1}-1} j_m^p,\ \ell=0,1,2,\cdots;\ m=1(1)n.$$

则 $\{j_m^{(\ell)}\}_{\ell\in N_0} \subset [0,1)$ $(m=1(1)n)$。

证明：略。

§3. 异步算法的收敛性

现在，我们建立算法1.1与算法1.2的收敛性理论。

定理3.1 设 $A\in L(R^n)$ 是H-矩阵，$(D-L_i, U_i, E_i)$ $(i=1,2,\cdots,\alpha)$ 为其一多分裂，并且

$$\langle A\rangle = |D|-|L_i|-|U_i| = |D|-|B|,\ i=1,2,\cdots,\alpha$$

则当参数 r 和 ω 满足

(3.1) $$0\le r\le \omega,\quad 0<\omega<2/(1+\rho(|D|^{-1}|B|))$$

时，从任何初始 $x^0\in R^n$ 出发，由算法1.1产生的序列 $\{x^p\}_{p\in N_0}$ 均收敛到线性方程组(1.1.1)的唯一解 $x^*\in R^n$。

证明：略。

定理3.2 设定理3.1的条件满足。则当参数 r,ω 满足(3.1)而 β 满足 $0<\beta<2/(1+\rho(H(\omega)))$，$H(\omega)=|1-\omega|+\omega|D|^{-1}|B|$

时，从任何初始 $x^0\in R^n$ 出发，由算法1.2产生的序列 $\{x^p\}_{p\in N_0}$ 均收敛到线性方程组(1.1.1)的唯一解 $x^*\in R^n$。

证明：略。

第四章　并行多层迭代算法

§1. 问题的背景

二阶自共轭椭圆型边值问题

$$
(1.1)\quad \begin{cases} -\sum_{m,j=1}^{K}\frac{\partial}{\partial x_m}\left(a_{mj}(x)\frac{\partial u}{\partial x_j}\right)=f(x), & \forall x\in \mathcal{D} \\ u|_{\Gamma_{\mathcal{D}}}=0, \quad \frac{\partial u}{\partial n}\Big|_{\Gamma_{\mathcal{D}}'}=0 \end{cases}
$$

经有限元离散，即可产生线性代数方程组(1.1.1)，其中 $A\in L(R^n)$ 是对称正定矩阵，且具分块形式：

$$
(1.2)\quad A=\begin{bmatrix} A_{11} & \cdots & A_{1\alpha} \\ \vdots & \ddots & \vdots \\ A_{\alpha 1} & \cdots & A_{\alpha\alpha} \end{bmatrix}
$$

这里，$A_{mj}\in L(R^{n_j},R^{n_m})$，$m,j=1,2,\cdots,\alpha$，而 $n_m\le n$ $(m=1,2,\cdots,\alpha)$ 为满足 $\sum_{i=1}^{\alpha}n_i=n$ 的 α 个给定的正整数。

$$
b=(b_1^T,b_2^T,\cdots,b_\alpha^T)^T,\ b_i\in R^{n_i},\ i=1,2,\cdots,\alpha
$$

为已知向量，而

$$
x=(x_1^T,x_2^T,\cdots,x_\alpha^T)^T,\ x_i\in R^{n_i},\ i=1,2,\cdots,\alpha
$$

为未知向量。

鉴于原偏微分方程(1.1)的具体特性，在实际应用中，以(1.2)为系数矩阵的线性方程组(1.1.1)往往是大型的，并且矩阵A的条件数为$O(n^2)$。因此，如何恰当地设计并行迭代算法，充分地利用多处理机系统的并行计算功能，来有效地求解(1.1.1)—(1.2)，是一项具有理论价值和实际意义的研究课题。

本章就是基于这样一个前提，运用多层迭代的思想(见[19]—[22])，构造了关于矩阵A的具良好条件数的预处理矩阵，并建立了求解这类方程组的并行多层迭代算法，并对这类算法进行了完整而细致的理论分析

§2. 预处理矩阵的构造与并行多层迭代算法的设计

首先，我们由矩阵$A\in L(R^n)$出发，对任取定的正整数l，构造矩阵序列$\{A^{(k)}\}_{k=0}^{l}$如下：

$$(2.1)\quad \begin{cases} A^{(l)}=A \\ A^{(k)}=\begin{bmatrix} A_{11}^{(k)} & \cdots & A_{1\alpha}^{(k)} \\ \vdots & \ddots & \vdots \\ A_{\alpha 1}^{(k)} & \cdots & A_{\alpha\alpha}^{(k)} \end{bmatrix},\ A_{ij}^{(k)}\in L(R^{n_j^{(k)}},R^{n_i^{(k)}}),\ i,j=1,2,\cdots,\alpha \\ k=0,1,\cdots,l-1 \end{cases}$$

这里，$\{n_i^{(k)}\}_{k=0}^{l}(i=1,2,\cdots,\alpha)$为给定的正整数，满足

$$(2.2)\quad n_i=n_i^{(l)}>n_i^{(l-1)}>\cdots>n_i^{(0)},\quad i=1,2,\cdots,\alpha$$

而

$$\begin{cases} A_{ii}^{(k)}=\begin{bmatrix} C_{ii}^{(k)} & E_{ii}^{(k)} \\ E_{ii}^{(k)T} & A_{ii}^{(k-1)} \end{bmatrix}, \quad i=1,2,\cdots,\alpha \\ A_{ij}^{(k)}=\begin{bmatrix} C_{ij}^{(k)} & E_{ij}^{(k)} \\ F_{ij}^{(k)} & A_{ij}^{(k-1)} \end{bmatrix}, \quad i\neq j,\ i,j=1,2,\cdots,\alpha \\ k=1,2,\cdots,l \end{cases}$$

根据矩阵A的对称正定性及$\{A^{(k)}\}_{k=0}^{l}$的具体结构易知，对于$k=0,1,\cdots,l$，$A_{ij}^{(k)T}=A_{ji}^{(k)}$，$i\neq j$，$i,j=1,2,\cdots,\alpha$，且$A^{(k)}$对称正定。因而，每个$A_{ii}^{(k)}(i=1,2,\cdots,\alpha)$也均是对称正定的。记$A_{ii}^{(k)}(i=1,2,\cdots,\alpha)$的Schur分解为

$$A_{ii}^{(k)}=\begin{bmatrix} C_{ii}^{(k)} & 0 \\ E_{ii}^{(k)T} & S_{ii}^{(k-1)} \end{bmatrix}\begin{bmatrix} I & C_{ii}^{(k)^{-1}}E_{ii}^{(k)} \\ 0 & I \end{bmatrix}, \quad k=1,2,\cdots,l$$

其中$S_{ii}^{(k-1)}$为$A_{ii}^{(k)}$的Schur补，

(2.5) $$S_{ii}^{(k-1)}=A_{ii}^{(k-1)}-E_{ii}^{(k)T}C_{ii}^{(k)^{-1}}E_{ii}^{(k)}。$$

对于$i=1,2,\cdots,\alpha$，分别取$\nu_i(\nu_i\geq 1)$次多项式$P_{\nu_i}^{(i)}(t)$，满足

(2.6) $$0\leq P_{\nu_i}^{(i)}(t)<1\ (0<t\leq 1),\quad P_{\nu_i}^{(i)}(0)=1,$$

并设$B_{ii}^{(k)}$为$C_{ii}^{(k)}(i=1,2,\cdots,\alpha)$的近似矩阵，则我们可定义矩阵$A_{ii}^{(k)}$的预处理矩阵为

(2.7) $$M_{ii}^{(k)}=\begin{bmatrix} B_{ii}^{(k)} & 0 \\ E_{ii}^{(k)T} & \tilde{S}_{ii}^{(k-1)} \end{bmatrix}\begin{bmatrix} I & B_{ii}^{(k)^{-1}}E_{ii}^{(k)} \\ 0 & I \end{bmatrix},\ M_{ii}^{(0)}=A_{ii}^{(0)},\ M_{ii}=M_{ii}^{(l)};\ k=1,2,\cdots,l,$$

其中

$$(2.8)\qquad \tilde{S}_{ii}^{(k-1)^{-1}}=\begin{cases}[I-P_{\nu_i}^{(i)}(M_{ii}^{(k-1)^{-1}}\hat{S}_{ii}^{(k-1)})]\hat{S}_{ii}^{(k-1)^{-1}}, & 情形(I)\\ [I-P_{\nu_i}^{(i)}(M_{ii}^{(k-1)^{-1}}A_{ii}^{(k-1)})]A_{ii}^{(k-1)^{-1}}, & 情形(II)\end{cases}$$

而 $\hat{S}_{ii}^{(k-1)}$ 为矩阵 $S_{ii}^{(k-1)}$ 的近似矩阵，即

$$(2.9)\qquad \hat{S}_{ii}^{(k-1)}=A_{ii}^{(k-1)}-E_{ii}^{(k)^T}B_{ii}^{(k)^{-1}}E_{ii}^{(k)}.$$

由此，我们可建立求解大型分块线性代数方程组(1.1.1)的并行多层迭代算法如下：

算法2.1：给定初始向量 $x^0\in R^n$，对 $p=0,1,2,\cdots$，计算

$$\begin{cases}r_i^p=b_i-\sum_{j=1}^{\alpha}A_{ij}x_j^p\\ x_i^{p+1}=x_i^p+\Delta x_i^p\end{cases},\quad i=1,2,\cdots,\alpha$$

其中 Δx_i^p 由方程组

$$M_{ii}\Delta x_i^p=\omega r_i^p$$

确定，而 $\omega>0$ 为松弛因子。

显然，上述算法中每个 r_i^p, Δx_i^p 及 $x_i^{p+1}(i=1,2,\cdots,\alpha)$ 的计算对于各个不同的 i 相互独立，因而，可以并行实现。

§3. 算法的具体实现及其计算工作量的估计

本节我们详细地给出了算法2.1在多处理机系统上的具体实施方案，并对于每个处理机，当

$$n_i^{(k)}\approx g_i n_i^{(k-1)},\ i=1,2,\cdots,\alpha,\ k=1,2,\cdots,l;\ 1\le\nu_i<g_i\ (i=1,2,\cdots,\alpha)$$

时，得到了其计算工作量不超过$O(\max_{1\le i\le\alpha} n_i)$的定量估计。

§4. 子矩阵相对条件数的计算

现在，对于每个$k\in\{0,1,2,\cdots,l\}$, $i\in\{1,2,\cdots,\alpha\}$, 我们计算矩阵$M_{ii}^{(k)}$关于矩阵$A_{ii}^{(k)}$的相对条件数。为此，我们做如下基本假设：

(A₁) 强Cauchy型不等式：

$$x^T E_{ii}^{(k)} y \le \gamma_{ii}(x^T C_{ii}^{(k)} x)^{\frac{1}{2}}(y^T A_{ii}^{(k-1)} y)^{\frac{1}{2}}, \gamma_{ii}\in(0,1), x\in R^{n_i^{(k)}-n_i^{(k-1)}}, \quad y\in R^{n_i^{(k-1)}};$$

(A₂) $B_{ii}^{(k)}$是对称正定矩阵，并且满足

$$x^T C_{ii}^{(k)} x \le x^T B_{ii}^{(k)} x \le (1+\beta_i) x^T C_{ii}^{(k)} x,\quad x\in R^{n_i^{(k)}-n_i^{(k-1)}},\ \beta_i\ge 0.$$

引理4.1 设强Cauchy型不等式成立，则

$$(1-\gamma_{ii}^2) y^T A_{ii}^{(k-1)} y \le y^T S_{ii}^{(k-1)} y \le y^T A_{ii}^{(k-1)} y,\quad \forall y\in R^{n_i^{(k-1)}}.$$

证明：略。

定义

$$S=\begin{cases}\hat{S}_{ii}^{(k-1)}, & \text{情形(I)}\\ A_{ii}^{(k-1)}, & \text{情形(II)}\end{cases} \tag{4.1}$$

记

$$T_{ii}^{(k-1)} = S^{\frac{1}{2}} M_{ii}^{(k-1)^{-1}} S^{\frac{1}{2}}, \tag{4.2}$$

我们有

引理4.2　若基本假设(A_1)-(A_2)满足，则

(1) $$\frac{y^T A_{ii}^{(k)} y}{y^T M_{ii}^{(k)} y} \leq 1,\ i=1,2,\cdots,\alpha;\ k=0,1,\cdots,l,\ y \in R^{n_i^{(k)}};$$

(2) 由(4.2)定义的$T_{ii}^{(k)}$的特征值均属于$[\alpha_i^{(k)},1]$。其中

$$\alpha_i^{(k)} = \begin{cases} (1-\gamma_{ii}^2)/\lambda_i^{(k)}, & 情形(I) \\ 1/\lambda_i^{(k)}, & 情形(II) \end{cases},$$

而

$$\lambda_i^{(k)} = \sup_{\bar{x} \neq 0} \frac{\bar{x}^T M_{ii}^{(k)} \bar{x}}{\bar{x}^T A_{ii}^{(k)} \bar{x}} \in [1,\infty).$$

证明：略。

引理4.3　若基本假设(A_1)满足，则对任$\xi>0$，成立

$$(1-\xi^{-1}\gamma_{ii}) y_1^T C_{ii}^{(k+1)} y_1 + (1-\xi\gamma_{ii}) y_2^T A_{ii}^{(k)} y_2 \leq y^T A_{ii}^{(k+1)} y.$$

特别，我们有

$$y^T A_{ii}^{(k+1)} y \geq (1-\gamma_{ii}^2) y_1^T C_{ii}^{(k+1)} y_1.$$

证明：略

引理4.4　若基本假设(A_1)-(A_2)满足，则

$$\frac{y^T S y}{y^T A_{ii}^{(k)} y} \leq \begin{cases} (1+\frac{\beta_i}{1-\gamma_{ii}^2})/(1+\beta_i), & 情形(I) \\ 1/(1-\gamma_{ii}^2), & 情形(II) \end{cases}$$

其中S由(4.1)定义。

证明：略。

定理4.1 若基本假设(A_1)-(A_2)满足，则

$$\lambda_i^{(k+1)} \leqslant \begin{cases} 1+\dfrac{\beta_i}{1-\gamma_{ii}^2}+\dfrac{d_i^{(k)}}{1+\beta_i}\left(1+\dfrac{\beta_i}{1-\gamma_{ii}^2}\right), & \text{情形}(\mathrm{I}) \\ 1+\dfrac{1}{1-\gamma_{ii}^2}\left(\beta_i+d_i^{(k)}\right) & \text{情形}(\mathrm{II}) \end{cases}$$

其中

$$\lambda_i^{(k)}=\sup_{\bar{x}\neq 0}\frac{\bar{x}^T M_{ii}^{(k)}\bar{x}}{\bar{x}^T A_{ii}^{(k)}\bar{x}}, \quad P_{\nu_i}^{(i)}(t_i^{(k)})=\max_{t\in[\alpha_i^{(k)},1]} P_{\nu_i}^{(i)}(t),$$

$$d_i^{(k)}=\begin{cases} P_{\nu_i}^{(i)}(t_i^{(k)})/(1-P_{\nu_i}^{(i)}(t_i^{(k)})), & \text{情形}(\mathrm{I}) \\ \gamma_{ii}^2+P_{\nu_i}^{(i)}(t_i^{(k)})/(1-P_{\nu_i}^{(i)}(t_i^{(k)})), & \text{情形}(\mathrm{II}) \end{cases}$$

证明：略。

下面，我们进一步去得到预处理矩阵$M_{ii}^{(k)}$关于矩阵$A_{ii}^{(k)}$的与层数无关的相对条件数估计。为此，我们再设每个多项式$P_{\nu_i}^{(i)}(t)$在区间$[0,1]$是单调不增的。令

$$Q_{\nu_i-1}^{(i)}(t)=(1-P_{\nu_i}^{(i)}(t))/t, \quad i=1,2,\cdots,\alpha$$

则由(2.6)知，每个$Q_{\nu_i-1}^{(i)}(t)$是ν_i-1次多项式，并且有

$$Q_{\nu_i-1}^{(i)}(t)>0 \ (0<t<1), \quad Q_{\nu_i-1}^{(i)}(0)=-P_{\nu_i}^{(i)\prime}(0)。$$

定理4.2 若基本假设(A_1)-(A_2)满足，且对于$i=1,2,\cdots,\alpha$，多项式$P_{\nu_i}^{(i)}(t)$在区间$[0,1]$中均是单调不增的，则对于每个$i\in\{1,2,\cdots,\alpha\}$

(1) 如果定义

$$\hat{\lambda}_i^{(k+1)}=\left(1+\frac{\beta_i}{1-\gamma_{ii}^2}\right)\left(1+\frac{\hat{d}_i^{(k)}}{1+\beta_i}\right), \quad \hat{\lambda}_i^{(0)}=1, \ k=0,1,\cdots,l-1,$$

其中

$$\hat{d}_i^{(k)} = p_{\nu_i}^{(i)}(\hat{\alpha}_i^{(k)})/(1-p_{\nu_i}^{(i)}(\hat{\alpha}_i^{(k)})),\ \hat{\alpha}_i^{(k)} = (1-\gamma_{ii}^2)/\hat{\lambda}_i^{(k)},\ k=0,1,\cdots l,$$

$\{\hat{\lambda}_i^{(k)}\}_{k=0}^{l}$ 即构成 $\{\lambda_i^{(k)}\}_{k=0}^{l}$ 的优序列。这时，我们有

$$\hat{\alpha}_i^{(k+1)} = \frac{(1-\gamma_{ii}^2)^2(1+\beta_i)}{1+\beta_i-\gamma_{ii}^2}\cdot\frac{\hat{\alpha}_i^{(k)}Q_{\nu_i-1}^{(i)}(\hat{\alpha}_i^{(k)})}{1+\beta_i\hat{\alpha}^{(k)}Q_{\nu_i-1}^{(i)}(\hat{\alpha}_i^{(k)})},\ k=0,1,\cdots,l-1.$$

从而当

(4.3) $$\frac{(1-\gamma_{ii}^2)^2(1+\beta_i)}{1+\beta_i-\gamma_{ii}^2}Q_{\nu_i-1}^{(i)}(0)>1$$

时，序列 $\{\hat{\alpha}_i^{(k)}\}_{k=0}^{\infty}$ 具唯一极限点 $\alpha_i^*\in(0,1)$，且满足

$$\hat{\alpha}_i^{(k)} \geq \alpha_i^*,\ k=0,1,2,\cdots.$$

这里，α_i^* 是方程

(4.4) $$\frac{(1-\gamma_{ii}^2)^2(1+\beta_i)}{1+\beta_i-\gamma_{ii}^2}Q_{\nu_i-1}^{(i)}(t)-\beta_i t\,Q_{\nu_i-1}^{(i)}(t)-1=0$$

在 (0,1) 中的唯一实根；

(2) 如果定义

$$\hat{\lambda}_i^{(k+1)} = 1+\frac{\beta_i+\hat{d}_i^{(k)}}{1-\gamma_{ii}^2},\quad \hat{\lambda}_i^{(0)}=1,\quad k=0,1,\cdots,l-1,$$

其中

$$\hat{d}_i^{(k)} = \gamma_{ii}^2+p_{\nu_i}^{(i)}(\hat{\alpha}_i^{(k)})/(1-p_{\nu_i}^{(i)}(\hat{\alpha}_i^{(k)})),\ \hat{\alpha}_i^{(k)} = 1/\hat{\lambda}_i^{(k)},\ k=0,1,\cdots,l.$$

$\{\hat{\lambda}_i^{(k)}\}_{k=0}^{l}$ 即构成 $\{\lambda_i^{(k)}\}_{k=0}^{l}$ 的优序列。这时，我们有

$$\hat{\alpha}_i^{(k+1)} = \frac{(1-\gamma_{ii}^2)\hat{\alpha}_i^{(k)}Q_{\nu_i-1}(\hat{\alpha}_i^{(k)})}{1+\beta_i\hat{\alpha}_i^{(k)}Q_{\nu_i-1}^{(i)}(\hat{\alpha}_i^{(k)})},\quad k=0,1,\cdots,l-1$$

从而当

(4.5) $$(1-\gamma_{ii}^2)Q_{\nu_i-1}^{(i)}(0)>1$$

时，序列 $\{\hat{\alpha}_i^{(k)}\}_{k=0}^{\infty}$ 具唯一极限点 $\alpha_i^*\in(0,1)$，且满足

$$\hat{\alpha}_i^{(k)} \geq \alpha_i^*, \quad k=0,1,2,\cdots。$$

这里，α_i^*是方程

$$(1-\gamma_{ii}^2)Q_{\nu_i-1}^{(i)}(t)-\beta_i t Q_{\nu_i-1}^{(i)}(t)-1=0 \tag{4.6}$$

在$(0,1)$中的唯一实根。

证明：略。

由定理4.1和定理4.2，我们易得如下结论.

定理4.3 若基本假设(A_1)-(A_2)满足，且多项式$P_{\nu_i}^{(i)}(t)(i=1,2,\cdots,\alpha)$在$[0,1]$中均是单调不增的，则对于每个$i\in\{1,2,\cdots,\alpha\}$.

(1) 如果(4.3)满足，对于情形(I)，我们有

$$\lambda_i^{(k)} \leq \hat{\lambda}_i^{(k)} \leq \frac{1-\gamma_{ii}^2}{\alpha_i^*},$$

其中$\alpha_i^*\in(0,1)$由方程(4.4)唯一确定；

(2) 如果(4.5)满足，对于情形(II)，我们有

$$\lambda_i^{(k)} \leq \hat{\lambda}_i^{(k)} \leq \frac{1}{\alpha_i^*},$$

其中$\alpha_i^*\in(0,1)$由方程(4.6)唯一确定.

§5 整体矩阵相对条件数的计算

令

$$M^{(k)}=\mathrm{diag}(M_{11}^{(k)},M_{22}^{(k)},\cdots,M_{\alpha\alpha}^{(k)}),\ k=0,1,\cdots,l;\ M=M^{(l)},$$

则§2定义的并行多层迭代算法亦可等价地表示为

$$\begin{cases} r^p = b - Ax^p \\ M\Delta x^p = \omega r^p, \quad p = 0,1,2\cdots. \\ x^{p+1} = x^p + \Delta x^p \end{cases}$$

由于该算法的收敛性态与矩阵$M^{-1}A$的最大、最小特征值密切相关，因此，我们下面估计$M^{(k)}$关于$A^{(k)}(k=0,1,\cdots,l)$的相对条件数。为此，我们再做如下基本假设：

(A3) 强Cauchy型不等式：

$$|y_i^T A_{ij}^{(k)} y_j| \le \gamma_{ij}(y_i^T A_{ii}^{(k)} y_i)^{\frac{1}{2}}(y_j^T A_{jj}^{(k)} y_j)^{\frac{1}{2}}, \quad \gamma_{ij} = \gamma_{ji} \in [0,1), \quad y_i \in R^{n_i^{(k)}},$$

$$i \ne j, \quad i,j = 1,2,\cdots,\alpha.$$

引理5.1　设$\{\xi_i\}_{i=1}^{\alpha}$，$\{\eta_i\}_{i=1}^{\alpha}$和$\{\zeta_i\}_{i=1}^{\alpha}$均为正实数，则

$$\min_{1\le i\le\alpha}\{\xi_i/\eta_i\} \le \sum_{i=1}^{\alpha}\xi_i\zeta_i \Big/ \sum_{i=1}^{\alpha}\eta_i\zeta_i \le \max_{1\le i\le\alpha}\{\xi_i/\eta_i\}.$$

证明：略。

引理5.2　若基本假设(A3)满足，则

$$\sum_{i=1}^{\alpha}(1-\gamma_i)y_i^T A_{ii}^{(k)} y_i \le y^T A^{(k)} y \le \sum_{i=1}^{\alpha}(1+\gamma_i)y_i^T A_{ii}^{(k)} y_i, \quad k=0,1,\cdots,l,$$

其中

$$\gamma_i = \sum_{j\ne i}\gamma_{ij}.$$

证明：略。

定理5.1　若基本假设(A1)-(A3)满足，且$\gamma_i \in [0,1)$，则

$$\min_{1\le i\le\alpha}\frac{1}{1+\gamma_i} \le \frac{y^T M^{(k)} y}{y^T A^{(k)} y} \le \max_{1\le i\le\alpha}\frac{\lambda_i^{(k)}}{1-\gamma_i}, \quad k=0,1,\cdots,l.$$

证明：略。

结合定理4.3和定理4.4，我们可得$M^{(k)}$关于$A^{(k)}(k=0,1,\cdots,l)$的与层数无关的相对条件数的估计。

定理5.2　若基本假设(A1)-(A3)满足，$P_{\nu_i}^{(i)}(t)\ (i=1,2,\cdots,\alpha)$在[0,1]中均是单调不增的，且$\gamma_i\in[0,1)$。则

(1)　如果(4.3)满足，对于情形(I)，我们有

$$\min_{1\le i\le\alpha}\frac{1}{1+\gamma_i}\le\frac{y^TM^{(k)}y}{y^TA^{(k)}y}\le\max_{1\le i\le\alpha}\frac{1}{(1-\gamma_i)\alpha_i^*},\ k=0,1,\cdots,l.$$

其中$\alpha_i^*\in(0,1)$由方程(4.6)唯一确定

§6若干应用

我们现将§4.§5中的理论结果具体化到多项式

(P_1)　$P_{\nu_i}^{(i)}(t)=(1-t)^{\nu_i}$

和

$$(P_2)\quad\begin{cases}P_{\nu_i}^{(i)}(t)=\left[T_{\nu_i}\left(\frac{1+\alpha_i-2t}{1-\alpha_i}\right)+1\right]\Big/\left[1+T_{\nu_i}\left(\frac{1+\alpha_i}{1-\alpha_i}\right)\right]\\ T_{\nu_i}\text{ 为 }\nu_i\text{ 次 Chebyshev 多项式}\end{cases}$$

得到一些实用的结论。

显见，(P_2)中的多项式$P_{\nu_i}^{(i)}(t)$是满足$0\le P_{\nu_i}^{(i)}(t)<1\ (0<t\le1)$，$P_{\nu_i}^{(i)}(0)=1$的在区间$[\alpha_i,1]\ (0<\alpha_i<1)$上的极小局部极大多项式，且有

$$P_{\nu_i}^{(i)}(\alpha_i)=2\Big/\left[T_{\nu_i}\left(\frac{1+\alpha_i}{1-\alpha_i}\right)+1\right],\quad P_{\nu_i}^{(i)}(1)=\left[(-1)^{\nu_i}+1\right]\Big/\left[T_{\nu_i}\left(\frac{1+\alpha_i}{1-\alpha_i}\right)+1\right].$$

注意到由(P_1)-(P_2)给出的多项式$p^{(i)}_{\nu_i}(t)$ $(i=1,2,\cdots,\alpha)$在区间$[0,1]$中均是单调不增的，利用定理4.3和定理5.2，我们可得

定理6.1. 若基本假设(A_1)-(A_3)满足，$\gamma_i\in[0,1)$ $(i=1,2,\cdots,\alpha)$且$p^{(i)}_{2_i}(t)$ $(i=1,2,\cdots,\alpha)$由(P_1)给出，则

(1) 如果

$$\frac{(1-\gamma_{ii}^2)^2(1+\beta_i)}{1+\beta_i-\gamma_{ii}^2}>\frac{1}{\nu_i},\quad i=1,2,\cdots,\alpha,$$

对于情形(I)，我们有

$$\lambda_i^{(k)}\le\frac{1-\gamma_{ii}^2}{\alpha_i}=\left(1+\frac{\beta_i}{1-\gamma_{ii}^2}\right)\left(1+\frac{1}{1+\beta_i}\frac{(1-\alpha_i)^{\nu_i}}{1-(1-\alpha_i)^{\nu_i}}\right),\ i=1,2,\cdots,\alpha;\ k=0,1,\cdots,l.$$

从而

$$\min_{1\le i\le\alpha}\frac{1}{1+\gamma_i}\le\frac{y^TM^{(k)}y}{y^TA^{(k)}y}\le\max_{1\le i\le\alpha}\frac{1}{1-\gamma_i}\left(1+\frac{\beta_i}{1-\gamma_{ii}^2}\right)\left(1+\frac{1}{1+\beta_i}\cdot\frac{(1-\alpha_i)^{\nu_i}}{1-(1-\alpha_i)^{\nu_i}}\right),k=0,1,\cdots,l,$$

其中$\alpha_i\in(0,1)$ $(i=1,2,\cdots,\alpha)$是方程

$$1-\gamma_{ii}^2=\left(1+\frac{\beta_i}{1-\gamma_{ii}^2}\right)t\left(1+\frac{1}{1+\beta_i}\frac{(1-t)^{\nu_i}}{1-(1-t)^{\nu_i}}\right),\quad i=1,2,\cdots,\alpha$$

的最小正实根。

特别，当

$$\nu_i=2,\ \frac{(1-\gamma_{ii}^2)^2(1+\beta_i)}{1+\beta_i-\gamma_{ii}^2}>\frac{1}{2},\quad i=1,2,\cdots,\alpha$$

时，可得

$$\begin{cases}\alpha_i=\dfrac{2(1+\beta_i)\sigma_i-1}{\beta_i+(1+\beta_i)\sigma_i/2+\sqrt{(\beta_i-(1+\beta_i)\sigma_i/2)^2+\beta_i}},\ \sigma_i=\dfrac{(1-\gamma_{ii}^2)^2}{1+\beta_i-\gamma_{ii}^2},\ i=1,2,\cdots,\alpha\\[2ex] \lambda_i^{(k)}\le(1-\gamma_{ii}^2)\dfrac{\beta_i+(1+\beta_i)\sigma_i/2+\sqrt{(\beta_i-(1+\beta_i)\sigma_i/2)^2+\beta_i}}{2(1+\beta_i)\sigma_i-1},\ i=1,2,\cdots,\alpha;\ k=0,1,\cdots,l\\[2ex] \min\limits_{1\le i\le\alpha}\dfrac{1}{1+\gamma_i}\le\dfrac{y^TM^{(k)}y}{y^TA^{(k)}y}\le\max\limits_{1\le i\le\alpha}\dfrac{1-\gamma_{ii}^2}{1-\gamma_i}\cdot\dfrac{\beta_i+(1+\beta_i)\sigma_i/2+\sqrt{(\beta_i-(1+\beta_i)\sigma_i/2)^2+\beta_i}}{2(1+\beta_i)\sigma_i-1}\\ \qquad k=0,1,\cdots,l;\end{cases}$$

(2) 如果

$$1-\gamma_{ii}^2>\frac{1}{\nu_i},\quad i=1,2,\cdots,\alpha,$$

对于情形(Ⅲ)，我们有

$$\lambda_i^{(k)}\leq\frac{1}{\alpha_i}=\frac{1}{1-\gamma_{ii}^2}\left(\beta_i+\frac{1}{1-(1-\alpha_i)^{\nu_i}}\right),\ i=1,2,\cdots,\alpha;\ k=0,1,\cdots,l.$$

从而

$$\min_{1\leq i\leq\alpha}\frac{1}{1+\gamma_i}\leq\frac{y^TM^{(k)}y}{y^TA^{(k)}y}\leq\max_{1\leq i\leq\alpha}\frac{1}{(1-\gamma_i)(1-\gamma_{ii}^2)}\left(\beta_i+\frac{1}{1-(1-\alpha_i)^{\nu_i}}\right),\ k=0,1,\cdots,l,$$

其中 $\alpha_i\in(0,1)\ (i=1,2,\cdots,\alpha)$ 是方程

$$1-\gamma_{ii}^2=t\left(\beta_i+\frac{1}{1-(1-t)^{\nu_i}}\right),\quad i=1,2,\cdots,\alpha$$

的最小正实根。

特别，当

$$\nu_i=2,\quad\gamma_{ii}^2<\frac{1}{2},\quad i=1,2,\cdots,\alpha$$

时，可得

$$\begin{cases}\alpha_i=\left[2\beta_i+(1-\gamma_{ii}^2)-\sqrt{(2\beta_i+(1-\gamma_{ii}^2))^2-4\beta_i(1-2\gamma_{ii}^2)}\right]/(2\beta_i),\ i=1,2,\cdots,\alpha\\ \lambda_i^{(k)}\leq\dfrac{2\beta_i}{2\beta_i+(1-\gamma_{ii}^2)-\sqrt{(2\beta_i+(1-\gamma_{ii}^2))^2-4\beta_i(1-2\gamma_{ii}^2)}},\ i=1,2,\cdots,\alpha;\ k=0,1,\cdots,l\\ \min\limits_{1\leq i\leq\alpha}\dfrac{1}{1+\gamma_i}\leq\dfrac{y^TM^{(k)}y}{y^TA^{(k)}y}\leq\max\limits_{1\leq i\leq\alpha}\dfrac{1}{1-\gamma_i}\cdot\dfrac{2\beta_i+(1-\gamma_{ii}^2)+\sqrt{(2\beta_i+(1-\gamma_{ii}^2))^2-4\beta_i(1-2\gamma_{ii}^2)}}{2(1-2\gamma_{ii}^2)}\\ \qquad\qquad k=0,1,2,\cdots,l.\end{cases}$$

证明：略。

定理6.2 若基本假设(A₁)-(A₃)满足，$\gamma_i\in(0,1]\ (i=1,2,\cdots,\alpha)$，且 $P_{\nu_i}^{(i)}(t)\ (i=1,2,\cdots,\alpha)$ 由 (P_2) 给出，则

(1) 如果

$$(6.1)\qquad \frac{(1-\gamma_{ii}^2)^2(1+\beta_i)}{1+\beta_i-\gamma_{ii}^2} > \frac{1}{\nu_i}\frac{[(1+\sqrt{\alpha_i})^{\nu_i}+(1-\sqrt{\alpha_i})^{\nu_i}]^2}{2\sum_{s=0}^{[\frac{\nu_i-1}{2}]}\binom{\nu_i}{2s+1}\alpha_i^s(1-\alpha_i)^{\nu_i-1}},\quad i=1,2,\cdots,\alpha,$$

对于情形(I)，我们有

$$(6.2)\qquad \lambda_i^{(k)} \le \frac{1-\gamma_{ii}^2}{\alpha_i} = (1+\frac{\beta_i}{1-\gamma_{ii}^2})(1+\frac{1}{1+\beta_i}\frac{4(1-\alpha_i)^{\nu_i}}{[(1+\sqrt{\alpha_i})^{\nu_i}-(1-\sqrt{\alpha_i})^{\nu_i}]^2},\ i=1,2,\cdots,\alpha,\ k=0,1,\cdots l.$$

从而，

$$(6.3)\qquad \min_{1\le i\le\alpha}\frac{1}{1+\gamma_i} \le \frac{y^T M^{(k)} y}{y^T A^{(k)} y} \le \max_{1\le i\le\alpha}\left\{\frac{1}{1-\gamma_i}(1+\frac{\beta_i}{1-\gamma_{ii}^2})(1+\frac{1}{1+\beta_i}\frac{4(1-\alpha_i)^{\nu_i}}{[(1+\sqrt{\alpha_i})^{\nu_i}-(1-\sqrt{\alpha_i})^{\nu_i}]^2}\right\},$$

$$k=0,1,\cdots,l,$$

其中 $\alpha_i\in(0,1)$ $(i=1,2,\cdots,\alpha)$ 是方程

$$(6.4)\qquad 1-\gamma_{ii}^2 = t(1+\frac{\beta_i}{1-\gamma_{ii}^2})(1+\frac{1}{1+\beta_i}\frac{4(1-t)^{\nu_i}}{[(1+\sqrt{t})^{\nu_i}-(1-\sqrt{t})^{\nu_i}]^2},\quad i=1,2,\cdots,\alpha$$

的最小正实根.

如果

$$\frac{(1-\gamma_{ii}^2)^2(1+\beta_i)}{1+\beta_i-\gamma_{ii}^2} > \frac{1}{\nu_i^2},\quad i=1,2,\cdots,\alpha,$$

则方程(6.4)的最小正实根 α_i $(i=1,2,\cdots,\alpha)$ 可保证(6.1)成立.因而亦有(6.2)-(6.3).

特别，当

$$\nu_i=2,\ \frac{(1-\gamma_{ii}^2)^2(1+\beta_i)}{1+\beta_i-\gamma_{ii}^2} > \frac{1}{4},\quad i=1,2,\cdots,\alpha$$

时，可得

$$\begin{cases} \alpha_i = \dfrac{4(1+\beta_i)\sigma_i - 1}{2\sqrt{(1+\beta_i)(\beta_i+\sigma_i)}+2\beta_i+1}, & i=1,2,\cdots,\alpha \\ \lambda_i^{(k)} \le (1-\gamma_{ii}^2)\dfrac{2\sqrt{(1+\beta_i)(\beta_i+\sigma_i)}+2\beta_i+1}{4(1+\beta_i)\sigma_i-1}, & i=1,2,\cdots,\alpha;\ k=0,1,\cdots,l \\ \min\limits_{1\le i\le\alpha}\dfrac{1}{1+\gamma_i} \le \dfrac{y^T M^{(k)} y}{y^T A^{(k)} y} \le \max\limits_{1\le i\le\alpha}\dfrac{1-\gamma_{ii}^2}{1-\gamma_i}\,\dfrac{2\sqrt{(1+\beta_i)(\beta_i+\sigma_i)}+2\beta_i+1}{4(1+\beta_i)\sigma_i-1}, & k=0,1,\cdots,l \end{cases}$$

(2) 如果

(6.5) $$\gamma_{ii}^2 < 1-\left(\frac{(1+\sqrt{\alpha_i})^{\nu_i}+(1-\sqrt{\alpha_i})^{\nu_i}}{2\sum\limits_{s=1}^{\nu_i}(1+\sqrt{\alpha_i})^{2\nu_i-s}(1-\sqrt{\alpha_i})^{s-1}}\right)^2, \quad i=1,2,\cdots,\alpha,$$

对于情形(1)，我们有

(6.6) $$\lambda_i^{(k)} \le \frac{1}{\alpha_i} = \frac{1}{1-\gamma_{ii}^2}\left[\beta_i+\left(\frac{(1+\sqrt{\alpha_i})^{\nu_i}+(1-\sqrt{\alpha_i})^{\nu_i}}{(1+\sqrt{\alpha_i})^{\nu_i}-(1-\sqrt{\alpha_i})^{\nu_i}}\right)^2\right],\quad i=1,2,\cdots,\alpha;\ k=0,1,\cdots,l.$$

从而

(6.7) $$\min_{1\le i\le\alpha}\frac{1}{1+\gamma_i} \le \frac{y^T M^{(k)} y}{y^T A^{(k)} y} \le \max_{1\le i\le\alpha}\frac{1}{(1-\gamma_i)(1-\gamma_{ii}^2)}\left[\beta_i+\left(\frac{(1+\sqrt{\alpha_i})^{\nu_i}+(1-\sqrt{\alpha_i})^{\nu_i}}{(1+\sqrt{\alpha_i})^{\nu_i}-(1-\sqrt{\alpha_i})^{\nu_i}}\right)^2\right],\ k=0,1,\cdots,l,$$

其中 $\alpha_i\in(0,1)$ $(i=1,2,\cdots,\alpha)$ 是方程

(6.8) $$t\left[\beta_i+\left(\frac{(1+\sqrt{t})^{\nu_i}+(1-\sqrt{t})^{\nu_i}}{(1+\sqrt{t})^{\nu_i}-(1-\sqrt{t})^{\nu_i}}\right)^2\right]=1-\gamma_{ii}^2,\quad i=1,2,\cdots,\alpha$$

的最小正实根。

如果

$$1-\gamma_{ii}^2 > \frac{1}{\nu_i^2},\quad i=1,2,\cdots,\alpha,$$

则方程(6.8)的最小正根 α_i $(i=1,2,\cdots,\alpha)$ 可保证(6.5)成立，因而亦有(6.6)-(6.7)。

特别，当

$$\nu_i=2,\quad \gamma_{ii}<\frac{\sqrt{3}}{2},\quad i=1,2,\cdots,\alpha$$

时，可得

$$
\begin{cases}
\alpha_i=\sqrt{(1+2\beta_i)^2+(3-4\gamma_{ii}^2)}-(1+2\beta_i), \quad i=1,2,\cdots,\alpha \\
\lambda_i^{(k)}\leqslant \dfrac{\sqrt{(1+2\beta_i)^2+(3-4\gamma_{ii}^2)}+(1+2\beta_i)}{3-4\gamma_{ii}^2}, \quad i=1,2,\cdots,\alpha;\ k=0,1,\cdots,l \\
\min\limits_{1\leqslant i\leqslant\alpha}\dfrac{1}{1+\gamma_i}\leqslant\dfrac{y^TM^{(k)}y}{y^TA^{(k)}y}\leqslant\max\limits_{1\leqslant i\leqslant\alpha}\dfrac{1}{1-\gamma_i}\cdot\dfrac{\sqrt{(1+2\beta_i^2)+(3-4\gamma_{ii}^2)}+(1+2\beta_i)}{3-4\gamma_{ii}^2} \\
\qquad\qquad k=0,1,\cdots,l.
\end{cases}
$$

证明：略。

最后，基于不完全三角分解，我们讨论了$C_{ii}^{(k)}$的近似矩阵$B_{ii}^{(k)}$($k=1,2,\cdots,\alpha$; $k=0,1,\cdots,l$)的满足基本假设(A_2)的两种取法(从略)。

注解：在对于有限元空间和三角剖分的一定正则性假设下，由(1.1)经离散所得到的线性代数方程组自然满足基本假设(A_1)-(A_3)[23]。

第二部分 大型非线性代数方程组

对于大型非线性代数方程组(1.1.4)，在非线性多分裂的意义下，我们设计了具很强概括性的同步与异步并行松弛迭代算法，并且完整地建立了这些新算法的局部收敛性理论。特别，我们深入细致地研究了同步并行非线性多分裂松弛算法的单调收敛性和全局收敛性，得到了比较深刻的结果。大量的数值实验表明，我们的算法是可行、有效的，所得数值结果均与理论分析完全相符。

第五章 非线性多重分裂与并行松弛算法

§1. 非线性多分裂及松弛算法

令取数集 $J_i (i=1,2,\cdots,\alpha,\ \alpha\le n)$，它们分别为数集 $\{1,2,\cdots,n\}$ 的 α 个非空子集合，且满足(2.1.1)，这些 $J_i (i=1,2,\cdots,\alpha)$ 彼此之间可以重复。

又有 α 个非负对角矩阵 $E_i=diag(e_1^{(i)},e_2^{(i)},\cdots,e_n^{(i)})\in L(R^n)\ (i=1,2,\cdots,\alpha)$ 它们满足(2.1.3)，且有 $\sum\limits_i E_i=I$。

今对映射 $F:\mathcal{D}\subset R^n\to R^n$ 定义多分裂：设有映射 $f^{(i)}:R^n\times R^n\to R^n$，$i=1,2,\cdots,\alpha$，使满足等式

$$f^{(i)}(x;x)=F(x),\ \forall x\in R^n,\ i=1,2,\cdots,\alpha$$

则称 $(f^{(i)},E_i)\ (i=1,2,\cdots,\alpha)$ 为 F 的非线性多分裂。

作为特例，可取

(1.1) $$f_m^{(i)}(x;y)=\begin{cases}F_m((I-P^{(i)})x+P^{(i)}y), & \text{当 } m\in J_i \text{ 时}\\ F_m(x_1,\cdots,x_{m-1},y_m,x_{m+1},\cdots,x_n), & \text{其它}\end{cases},\quad m=1(1)n,$$

其中$f_m^{(i)}$，F_m分别为$f^{(i)}$，F的第m个分量，而$P^{(i)}(i=1,2,\cdots,\alpha)$则定义为如下投影算子

(1.2) $$\begin{cases}P^{(i)}: R^n\longrightarrow R^n\\ P_m^{(i)}(x)=(P^{(i)}(x))_m=\begin{cases}x_m, & \text{当 } m\in J_i \text{ 时}\\ 0, & \text{其它}\end{cases}\end{cases}$$

利用前述非线性多分裂的概念，我们可建立如下求解非线性方程组(1.1.4)的加速松弛型算法。

算法1.1：给定初始近似$x^0\in\mathscr{D}\subset R^n$，对于$p=0,1,\cdots$，计算

$$\begin{cases}x^{p+1}=\frac{\omega}{r}\sum_i E_i x^{p,i}+(1-\frac{\omega}{r})x^p\\ r\in(0,+\infty),\ \omega\in(0,+\infty)\end{cases}$$

其中$x^{p,i}$的分量$x_m^{p,i}(m=1,2,\cdots,n)$满足

$$x_m^{p,i}=\begin{cases}r\tilde{x}_m^{p,i}+(1-r)x_m^p, & \text{当 } m\in J_i \text{ 时}\\ x_m^p, & \text{当 } m\notin J_i \text{ 时}\end{cases},\quad m=1,2,\cdots,n;\ i=1,2,\cdots,\alpha$$

而$\tilde{x}_m^{p,i}$满足方程

(1.3) $$\begin{cases}f_m^{(i)}(x^p;x_1^{p,i},\cdots,x_{m-1}^{p,i},\tilde{x}_m^{p,i},x_{m+1}^p,\cdots x_n^p)=0,\ m\in J_i\\ m=1,2,\cdots,n;\ i=1,2,\cdots,\alpha.\end{cases}$$

这里，我们称r为松弛因子，ω为加速因子

我们称算法1.1为非线性多分裂AOR型算法。为今后讨

论方便，简记为 $\mathcal{T}_{NMAOR}(r,\omega)$—算法.

如果映射 $f^{(i)}:\mathcal{D}\times\mathcal{D}\subset R^n\times R^n\to R^n\ (i=1,2,\cdots,\alpha)$ 关于 y 的偏导数存在，我们将它们记为

$$\partial_2 f^{(i)}(x;y)=(\partial_2 f^{(i)}(x;y))_{l,j}\ ,\ l,j=1,2,\cdots,n.$$

此时，我们可给出下述更便于应用的非线性多分裂算法：

算法1.2：给定初始近似 $x^0\in\mathcal{D}$，对于 $p=0,1,2,\cdots$，计算

$$\begin{cases}x^{p+1}=\frac{\omega}{r}\sum_i E_i x^{p,i}+(1-\frac{\omega}{r})x^p\\ r\in(0,+\infty),\ \omega\in(0,+\infty)\end{cases}$$

其中

$$\begin{cases}x^p=(x_1^p,x_2^p,\cdots,x_n^p)^T\\ x^{p,i}=(x_1^{p,i},x_2^{p,i},\cdots,x_n^{p,i})^T\\ x_m^{p,i}=\begin{cases}x_m^p-rf_m^{(i)}(x^p,U_m^{p,i})/H_{mm}^{(i)}(x^p,U_m^{p,i}),\ m\in J_i\\ x_m^p,\ m\notin J_i\end{cases}\\ U_m^{p,i}=(x_1^{p,i},x_2^{p,i},\cdots,x_{m-1}^{p,i},x_m^p,\cdots,x_n^p)^T\\ m=1,2,\cdots,n;\ i=1,2,\cdots,\alpha\end{cases}$$

这里，$H_{mm}^{(i)}(x;y)$ 为某矩阵 $H^{(i)}(x;y)\in L(R^n)$，$\forall x,y\in\mathcal{D}$ 的第 m 个对角元。具体可选择：

1) 当 $H_{mm}^{(i)}(x^p;U_m^{p,i})$ 取为

$$H_{mm}^{(i)}(x^p;U_m^{p,i})=\partial_2 f_{mm}^{(i)}(x^p;U_m^{p,i})$$

时，算法1.2成为非线性多分裂AOR-Newton算法，因此时我们

应用Newton法近似求解了方程(1.3)，其中$\partial_2 f_{mm}^{(i)}(x^p;U_m^{p,i})$表示$f_m^{(i)}$关于第二个变量的第m个分量的偏导数在点$(x^p;U_m^{p,i})$处的值。

2) 当$H_{mm}^{(i)}(x^p;U_m^{p,i})$取为

$$H_{mm}^{(i)}(x^p,U_m^{p,i})=\frac{f_m^{(i)}(x^p;U_m^{p,i}+(x_m^{p-1}-x_m^p)e_m)-f_m^{(i)}(x^p;U_m^{p,i})}{x_m^{p-1}-x_m^p}\quad,m\in J_i$$

时，算法1.2成为非线性多分裂AOR-Chord算法。因为，此时我们用弦法近似求解了方程(1.3).

3) 当$H_{mm}^{(i)}(x^p;U_m^{p,i})$取为

$$H_{mm}^{(i)}(x^p;U_m^{p,i})=\frac{f_m^{(i)}(x^p;U_m^{p,i}+f_m^{(i)}(x^p;U_m^{p,i})e_m)-f_m^{(i)}(x^p;U_m^{p,i})}{f_m^{(i)}(x^p;U_m^{p,i})},\ m\in J_i$$

时，算法1.2成为非线性多分裂AOR-Steffensen算法。因为，此时我们用Steffensen方法近似求解了方程(1.3)。

算法1.2实际就是文[13]在非线性多分裂意义下的推广与改进，且具有并行计算之功能.

§2. 算法的局部收敛性分析(I)

本节我们建立关于算法1.1的局部收敛性理论。

今假定$f^{(i)}(i=1,2,\cdots,\alpha)$关于x,y有偏导数存在，于是，对任意$x\in\mathbb{D}$，我们有

$$F'(x)=\partial_1 f^{(i)}(x;x)+\partial_2 f^{(i)}(x;x)\in L(R^n)$$

其中$\partial_1 f^{(i)}$表示关于第一个变量的偏导数，$\partial_2 f^{(i)}$表示关于第二个变量的偏导数. 显然分别有$\partial_1 f^{(i)}(x;x),\partial_2 f^{(i)}(x;x)\in L(R^n)$. 今令

$$B_i=\partial_2 f^{(i)}(x^*;x^*)=(b_{\ell j}^{(i)}),\quad C_i=-\partial_1 f^{(i)}(x^*;x^*)=(C_{\ell j}^{(i)})$$

其中$x^*\in\mathscr{D}$是$F(x)=0$的一个根. 于是，若取

$$\begin{cases} D_i=diag(B_i),\ det(D_i)\neq 0 \\ L_i=(l_{\ell j}^{(i)}),\quad l_{\ell j}^{(i)}=\begin{cases}-b_{\ell j}^{(i)}, & l>j,\ (l,j)\in J_i \\ 0, & 其它\end{cases} \\ U_i=(u_{\ell j}^{(i)}) \end{cases}$$

并满足

$$B_i=D_i-L_i-U_i,\quad i=1,2,\cdots,\alpha$$

则有

$$F'(x^*)=D_i-L_i-(U_i+C_i),\quad i=1,2,\cdots,\alpha$$

同时,(D_i-L_i,U_i+C_i,E_i), $i=1,2,\cdots,\alpha$, 即构成矩阵$F'(x^*)$的一个多分裂。

定理2.1. 设$(f^{(i)},E_i)$, $i=1,2,\cdots,\alpha$, 为F的一个非线性多分裂，$x^*\in\mathscr{D}$满足$F(x^*)=0$. 对每个$i\in\{1,2,\cdots,\alpha\}$, 映射$f^{(i)}:\mathscr{D}\times\mathscr{D}\subset R^n\times R^n\to R^n$在$(x^*;x^*)$的某邻域$S_0$中连续可微，且$(D_i-L_i,U_i+C_i,E_i)(i=1,2,\cdots,\alpha)$为矩阵$F'(x^*)$的一个多分裂、则当

$$\begin{cases}\rho(\mathcal{L}_{r,\omega}(x^*))<1\\ \mathcal{L}_{r,\omega}(x^*)=\sum_i E_i(D_i-rL_i)^{-1}[(1-\omega)D_i+(\omega-r)L_i+\omega(U_i+C_i)]\end{cases}$$

时，$x^*\in\mathcal{D}$是$\mathcal{F}_{NMAOR}(r,\omega)$—算法的一个吸引点，且$R_1$—收敛因子为$\rho(\mathcal{L}_{r,\omega}(x^*))$.

证明：略。

若记

$$\bar{D}_i=diag(C_i),\quad D=diag(F'(x^*))$$

则有

$$D=D_i-\bar{D}_i,\quad \bar{C}=C_i-diag(C_i)$$

于是，我们容易得到定理2.1的一个有用的推论.

推论2.1　设定理2.1的条件满足，且$F'(x^*)$为H-矩阵，满足

$$\langle F'(x^*)\rangle=|D_i|-|L_i|-|U_i+C_i|=|D|-|B|,\quad i=1,2,\cdots,\alpha$$

则当$0<r\le\omega$, $\omega\in(0,2/(1+\rho(|D|^{-1}|B|)))$时，$x^*$为$\mathcal{F}_{NMAOR}(r,\omega)$—算法的一个吸引点。

对于由(1.1)-(1.2)所定义的特殊分裂，我们可得下述推论。

推论2.2. 定理2.1的条件满足，$F'(x^*)$为H-矩阵，

$$D=diag(F'(x^*)),\quad F'(x^*)=D-B.$$

则当$0<r\le\omega$, $\omega\in(0,2/(1+\rho(|D|^{-1}|B|)))$时，$x^*$为$\mathcal{F}_{NMAOR}(r,\omega)$—算法的一个吸引点。

§3 算法的局部收敛性分析(II)

本节我们以非线性多分裂AOR-Newton算法为例，建立关于算法1.2的局部收敛性理论。至于非线性多分裂AOR-Chord算法以及非线性多分裂AOR-Steffensen算法的收敛性证明，可相似地得出。

定理3.1. 定理2.1的条件满足，则存在x^*的开邻域$S\subset\mathcal{D}$，使$\mathcal{F}_{NMAOR}(r,\omega)$-Newton算法在开邻域$S$上有意义，且$x^*$为该算法的吸引点，其$R_1$-收敛因子为

$$\rho(\mathcal{L}_{r,\omega}(x^*))=\rho(G(x^*))$$

其中

$$\mathcal{L}_{r,\omega}(x^*)=\sum_i E_i(D_i-rL_i)^{-1}[(1-\omega)D_i+(\omega-r)L_i+\omega(U_i+C_i)].$$

证明：略。

§4. 数值例子

考虑半线性椭圆方程：

$$\begin{cases}-\frac{\partial}{\partial x}\left((1+x^2+y^2)\frac{\partial u}{\partial x}\right)-\frac{\partial}{\partial y}\left((1+e^x+e^y)\frac{\partial u}{\partial y}\right)=-g(x,y)e^u, & (x,y)\in\Omega\\ u=x^2+y^2, & (x,y)\in\partial\Omega\end{cases}$$

其中

$$\begin{cases}g(x,y)=2(2+3x^2+y^2+e^x+(1+y)e^y)e^{x^2-y^2}\\ \Omega=(0,1)\times(0,1)\end{cases}$$

按照文[8]同样的要求进行离散化，同时，按文[8]同样的分裂方法，考虑五种多分裂，即$i=2,4,10,20,100$的情形，J_i的取法亦同文[8]。我们利用多分裂AOR-Newton法，取不同的参数r,ω进行了大量的数值实验。数值结果表明，分裂多不一定比分裂少来得好，$i=4$的情形为最好。还可看到当$\omega=r$时往往能得到较好的收敛效果，但是松弛参数ω选择范围较小，而在$\omega\neq r$的情况下，适当调节ω与r，可在ω和r的较大的取值范围内得到算法的收敛性。由此，说明非线性多分裂AOR型算法的收敛性质要优于非线性多分裂SOR型算法的收敛性质，而且收敛速度一般亦不低于SOR型算法。

第六章 非线性多分裂松弛算法的单调收敛性

本章进一步研究在第五章中已建立的求解大型非线性代数方程组(1.1.4)的非线性多分裂AOR型算法的单调收敛性和全局收敛性。同时，我们完整地得到了求解大型线性代数方程组(1.1.1)的矩阵多分裂迭代算法的单调收敛性理论，以及在单调意义下收敛速度的比较定理。大量数值结果表明，所述理论与实际完全相符。

§1. 单调收敛性分析

本节我们主要建立非线性多分裂AOR型算法5.1.1的双侧逼近性质。

定理1.1 设$(f^{(i)},E_i)$, $i=1,2,\cdots,\alpha$ 是$F:\mathscr{D}\subset R^n\to R^n$的一个非线性多分裂，对于每个$i\in\{1,2,\cdots,\alpha\}$, $f^{(i)}(x;y):\mathscr{D}\times\mathscr{D}\subset R^n\times R^n\to R^n$连续，且关于变量$x$反序，关于变量$y$非对角反序但严格对角保序。又存在$x^0,y^0\in\mathscr{D}$，使得

$$x^0\le y^0,\ J=\{x\in R^n|x^0\le x\le y^0\}\subset\mathscr{D}$$

$$F(x^0)\le 0\le F(y^0)$$

则从初始点x^0,y^0出发，由非线性多分裂AOR算法5.1.1产生的序列$\{x^p\}$, $\{y^p\}$均在J中有意义，且对任何$\{r,\omega\}\subset(0,1]$成立

(a) $x^0\le x^p\le x^{p+1}\le y^{p+1}\le y^p\le y^0$；

(b) $\lim\limits_{p\to\infty}x^p=x^*\le y^*=\lim\limits_{p\to\infty}y^p$；

(c) $x^*,y^*\in J$ 均是$F(x)=0$的解；

(d) 若$\tilde{x}$是$F(x)=0$在J中的一个解，则必有$x^*\le\tilde{x}\le y^*$.

证明：略。

§2. 收敛速度的比较与全局收敛性定理

下面的定理刻划了非线性多分裂AOR算法5.1.1的收敛

速度对于参数(r,ω)的依赖性.

定理2.1　在定理1.1的条件下，设$\{\omega,\omega'\}\subset(0,1]$分别为给定的参数。记从$y^0$出发由非线性多分裂AOR算法利用参数对$(r,\omega)$和$(r,\omega')$所产生的序列分别为$\{y^p\}$和$\{y'^p\}$。则当

$$0<\omega\leq\omega'\leq 1$$

时，成立

$$y^p\geq y'^p,\quad p=0,1,2,\cdots$$

并且$\{y^p\}$和$\{y'^p\}$分别单调下降收敛到方程组$F(x)=0$在J中的同一解$y^*\in J$。另外，若$\tilde{x}\in J$是$F(x)=0$的解，则有$\tilde{x}\leq y^*$.

对于由初始x^0出发由非线性多分裂AOR算法利用参数对(r,ω)和(r,ω')所产生的序列$\{x^p\}$和$\{x'^p\}$，相应的结论亦成立。

证明：略。

定理2.2　设$F:\mathscr{D}\subset R^n\to R^n$是逆保序映射，$f:\mathscr{D}\times\mathscr{D}\subset R^n\times R^n\to R^n$满足$f(x;x)=F(x)$，$\forall x\in\mathscr{D}$，并且$f(x;y)$关于$x$反序，关于$y$非对角反序，则$F:\mathscr{D}\subset R^n\to R^n$是M-映射。

证明：略。

利用前面已建立的事实，我们给出如下关于非线性多分裂AOR算法的全局收敛性定理。

定理2.3　在定理1.1的条件下，再设$\mathscr{D}=R^n$，$F:R^n\to R^n$是

满的逆保序映射，且对于每个$i\in\{1,2,\cdots,\alpha\}$，$f^{(i)}(x;y)$是关于$y$的满的对角保序映射。则对任$\zeta\in R^n$，从任何初始点$x^0\in R^n$出发，由非线性多分裂AOR算法产生的序列$\{x^p\}$收敛到方程组$F(x)=\zeta$在$R^n$中的唯一解$x^*$.

证明：略

最后，我们以如下两个注解结束本节。

注解1°. 对于特殊的多重分裂(5.1.1)-(5.1.2)，定理1.1，定理2.1中的条件"对于每个$i\in\{1,2,\cdots,\alpha\}$，$f^{(i)}(x;y):\mathcal{D}\times\mathcal{D}\subset R^n\times R^n\to R^n$关于变量$x$反序，关于变量$y$非对角反序且严格对角保序"可代之以"F非对角反序"。

注解2°. 对于上述特殊的多重分裂，定理2.3中的所有条件均可用"$F:R^n\to R^n$为满的M-函数"代替。

§3. 矩阵多分裂迭代法的单调收敛理论

当$F:R^n\to R^n$为线性映射，即

$$F(x)=Ax-b,\quad A\in L(R^n)\text{ 非奇异},\quad x,b\in R^n$$

时，算法5.1.1成为求解线性方程组(1.1.1)的矩阵多分裂AOR算法[3]。作为§1，§2中理论结果的具体化，我们立得关于该算法的单调收敛性定理和单调性的比较定理。对于矩阵多分裂迭代算法1.1.1的单调收敛性，和在单调意义下，由多

重分裂与单分裂、以及不同的多重分裂所构成的迭代算法之间的敛速比较，我们亦进行了详细的论证。限于篇幅，在此，我们均予以省略。

§4. 数值结果

为验证前述结果的有效性和正确性，我们考虑弱线性常微分方程的两点边值问题

$$(4.1)\qquad \begin{cases} \dfrac{d^2u}{dt^2}=\dfrac{1}{2}u^2, & 0<t<1 \\ u(0)=1,\quad u(1)=2. \end{cases}$$

用等步长 $h=\frac{1}{n+1}$ 对其进行离散，我们可得关于未知量 $u=(u_1,u_2,\cdots,u_n)^T$ 的型如

$$(4.2)\qquad F(u)=Au+\varphi(u)=0$$

的非线性方程组，其中 A 是 M-矩阵而 $\varphi: R^n\to R^n$ 是严格对角的保序映射，因此，方程组 (4.2) 在 R^n 中有唯一解。对于方程组 (4.2)，我们取 $n=6$ 和三种不同的多分裂情形利用Newton法近似求解非线性方程 (5.1.3) 对于不同的参数对 (γ,ω)，运用 $\mathcal{B}_{NMAOR}(\gamma,\omega)$—算法数值求解 (4.2)，进行了大量的数值实验，所得的所有数值结果均与理论相符。

第七章 异步非线性多分裂松弛算法

§1. 异步松弛算法模型

采用第三章§1中引进的符号系统和第五章§1中定义的非线性多重分裂的概念，对于大型非线性代数方程组(1.1.4)，我们构造求其解的如下异步非线性多分裂松弛算法：

算法1.1：设$x^0\in R^n$为(1.1.4)的解的初始近似，并且假定已获得近似序列$x^0,x^1,\cdots,x^p$，则第$p+1$次近似x^{p+1}的计算过程为：

(I) 逐次求解非线性方程组

$$f_m^{(i)}(x^{s^{(i)}(p)};\tilde{x}_1^{p,i},\cdots,\tilde{x}_{m-1}^{p,i},\hat{x}_m^{p,i},x_{m+1}^{s_{m+1}^{(i)}(p)},\cdots,x_n^{s_n^{(i)}(p)})=0,\text{若 } m\in J_i(p), \tag{1.1}$$

$$m=1(1)n;\ i=1,2,\cdots,\alpha,$$

得$\hat{x}_m^{p,i}(m\in J_i(p),i=1,2,\cdots,\alpha)$，其中

$$x^{s^{(i)}(p)}=(x_1^{s_1^{(i)}(p)},x_2^{s_2^{(i)}(p)},\cdots,x_n^{s_n^{(i)}(p)})^T,\ i=1,2,\cdots,\alpha, \tag{1.2}$$

而$\tilde{x}^{p,i}=(\tilde{x}_1^{p,i},\tilde{x}_2^{p,i},\cdots,\tilde{x}_n^{p,i})^T$为

$$\tilde{x}_m^{p,i}=\begin{cases}r\hat{x}_m^{p,i}+(1-r)x_m^{s_m^{(i)}(p)}, & \text{若 } m\in J_i(p)\\ x_m^{s_m^{(i)}(p)}, & \text{若 } m\notin J_i(p)\end{cases},m=1(1)n;\ i=1,2,\cdots,\alpha; \tag{1.3}$$

(II) 计算$x^{p,i}=(x_1^{p,i},x_2^{p,i},\cdots,x_n^{p,i})^T$，

$$x_m^{p,i}=\begin{cases}\frac{\omega}{r}\tilde{x}_m^{p,i}+(1-\frac{\omega}{r})x_m^{s_m^{(i)}(p)}, & \text{若 } m\in J_i(p)\\ x_m^p, & \text{若 } m\notin J_i(p)\end{cases},m=1(1)n;\ i=1,2,\cdots,\alpha; \tag{1.4}$$

(III) 计算

$$x_m^{p+1}=\sum_{i=1}^{\alpha}e_m^{(i)}x_m^{p,i},\ m=1,2,\cdots,n. \tag{1.5}$$

这里，$r\in(0,+\infty)$为松弛因子，$\omega\in(0,+\infty)$为加速因子。

显然，利用(1.3)，(1.4)可等价地表示为

$$(1.6)\qquad x_m^{p,i}=\begin{cases}\omega\hat{x}_m^{p,i}+(1-\omega)x_m^{s_m^{(i)}(p)}, & 若\ m\in J_i(p)\\ x_m^p, & 若\ m\notin J_i(p)\end{cases},\ m=1,2,\cdots,n;\ i=1,2,\cdots,\alpha.$$

由(1.1)－(1.3)，(1.5)－(1.6)容易看出，相应于参数对(r,ω)的特殊选取$(0,1)$，$(0,\omega)$，$(1,1)$，$(1,\omega)$和(ω,ω)，即分别得到异步并行非线性多分裂Jacobi，外插型Jacobi，Gauss-Seidel，外插型Gauss-Seidel以及SOR等实用而有效的算法。另外，对于$m\in\{1,2,\cdots,n\}$，$i\in\{1,2,\cdots,\alpha\}$，当

$$(1)\qquad\begin{cases}J_i=\{1,2,\cdots,n\}\\ \forall p\in N_0,\ J_i(p)=J_i,\ s_m^{(i)}(p)=p\end{cases}$$

时，算法1.1即退化为同步并行非线性多分裂AOR算法5.1.1；

$$(2)\qquad\begin{cases}J_i=\{1,2,\cdots,n\}\\ \forall p\in N_0,\ (J_i(p)=J_i)\vee(J_i(p)=\phi)=TRUE,\ s_m^{(i)}(p)=s_i(p)\in R^1\end{cases}$$

时，算法1.1成为文[69]中给出的异步并行非线性多分裂AOR算法。

算法1.1中的隐式非线性方程组(1.1)的精确求解，往往是不可能的。在具体应用时，通常采用近似方法来求得(1.1)的近似解。

算法1.2：给定(1.1.4)的解的初始近似$x^0\in R^n$，并且假定已

获得近似序列$x^0, x^1, \cdots, x^p$，则第$p+1$次近似$x^{p+1}=(x_1^{p+1}, x_2^{p+1}, \cdots, x_n^{p+1})^T$分别由

$$(1.7)\qquad \hat{x}_m^{p,i}=x_m^{s_m^{(i)}(p)}-f_m^{(i)}(x^{s^{(i)}(p)};U_m^{p,i})/H_{mm}^{(i)}(x^{s^{(i)}(p)};U_m^{p,i}),\ m\in J_i(p),$$
$$i=1,2,\cdots,\alpha$$

以及(1.2)—(1.5)确定。这里，对$i=1,2,\cdots,\alpha$, $m=1(1)n$,

$$U_m^{p,i}=(\tilde{x}_1^{p,i},\cdots,\tilde{x}_{m-1}^{p,i},x_m^{s_m^{(i)}(p)},\cdots,x_n^{s_n^{(i)}(p)})^T,$$

$H_{mm}^{(i)}(x;y)$为$\partial_2 f^{(i)}(x;y)$的某近似矩阵$H^{(i)}(x;y)$的第m个对角元，而$\partial_2 f^{(i)}(x;y)$, $\partial_1 f^{(i)}(x;y)$分别表示$f^{(i)}(x;y)$关于自变量y、x的偏导数。

在算法1.2中，相应于$H^{(i)}(x;y)$ $(i=1,2,\cdots,\alpha)$的不同选取，即可导出各种实用而有效的程序。例如，当

(1) $H_{mm}^{(i)}(x;y)=\partial_2 f_{mm}^{(i)}(x^{s^{(i)}(p)};U_m^{p,i})$ $(m\in J_i(p),\ i=1,2,\cdots,\alpha)$

时，可得异步非线性多分裂AOR-Newton法；

(2) $H_{mm}^{(i)}(x^{s^{(i)}(p)};U_m^{p,i})=\dfrac{f_m^{(i)}(x^{s^{(i)}(p)};U_m^{p,i}+h_m^{p,i}e_m)-f_m^{(i)}(x^{s^{(i)}(p)};U_m^{p,i})}{h_m^{p,i}}$ $(m\in J_i(p), i=1,2,\cdots,\alpha)$

时，可得异步并行非线性多分裂AOR-Chord法。这里，$h_m^{p,i}$ $(m\in J_i(p), i=1,2,\cdots,\alpha, p\in N_0)$为给定的差分步长；

(3) (2)中$h_m^{p,i}=f_m^{(i)}(x^{s^{(i)}(p)};U_m^{p,i})$ $(m\in J_i(p),\ m=1(1)n, i=1,2,\cdots,\alpha, \forall p\in N_0)$时，可得异步并行非线性多分裂AOR-Steffensen法。

§2. 算法的局部收敛性分析(I)

本节我们建立关于算法1.1的局部收敛性理论。所用符号的意义，请参看第五章§2。

定理2.1 设$x^*\in\mathscr{D}$是非线性方程组(1.1.4)的解，$(f^{(i)},E_i)$ $(i=1,2,\cdots,\alpha)$是$F:\mathscr{D}\subset R^n\to R^n$的一个非线性多分裂，并且对于$i=1,2,\cdots,\alpha$，$f^{(i)}:\mathscr{D}\times\mathscr{D}\subset R^n\times R^n\to R^n$在$(x^*;x^*)$的某邻域连续可微。再设$(D_i-L_i,U_i+C_i,E_i)$，$i=1,2,\cdots,\alpha$，是$F'(x^*)$的一个多分裂且满足

$$\langle F'(x^*)\rangle=|D_i|-|L_i|-|U_i+C_i|=|D|-|B|,\quad i=1,2,\cdots,\alpha.$$

则当$F'(x^*)\in L(R^n)$是H-矩阵时，如果

$$0<r\leq\omega,\quad 0<\omega<2/(1+\rho(|D|^{-1}|B|)),$$

存在x^*的邻域$S(x^*,\delta)$，使得对任何初始近似$x^0\in S(x^*,\delta)$，由算法1.1产生的序列$\{x^p\}$有意义，且收敛到非线性方程组(1.1.4)的解$x^*\in R^n$。

证明：略。

§3. 算法的局部收敛性分析(II)

我们建立关于算法1.2的局部收敛性定理。

定理3.1 在定理2.1的条件下，再设对每个$i\in\{1,2,\cdots,\alpha\}$，$H^{(i)}(x;y)$在$(x^*;x^*)$的某邻域连续可微，且满足

$$\lim_{(x;y)\to(x^*;x^*)}H^{(i)}_{mm}(x;y)=\partial_2 f^{(i)}_{mm}(x^*;x^*),\quad m=1(1)n,$$

则存在x^*的邻域$S(x^*,\delta)$，使得对于任何初始近似$x^0\in S(x^*,\delta)$，

由算法1.2产生的迭代序列$\{x^p\}$有意义，且收敛到非线性方程组(1.1.4)的解$x^*\in R^n$.

证明：略。

注解：对于异步并行非线性多分裂AOR-Newton、-Chord，-Steffensen算法的收敛性理论，可作为定理3.1的特殊情形得到。

References

[1] O'Leary,D.P., White,R.E.: Multisplittings of matrices and parallel solution of linear systems, SIAM J.Alg.Disc.Methods, 6(1985), 630-640.

[2] Frommer,A., Mayer,G.: Convergence of relaxed parallel multisplitting methods, Linear Algebra Appl., 119(1989), 141-152.

[3] Wang Deren: On the convergence of parallel multisplitting AOR algorithm, Linear Algebra Appl., 154-156(1991), 473-486.

[4] White,R.E.: Multisplitting with different weighting schemes, SIAM J. MatrixAnal.Appl., 10(1989), 481-493.

[5] White,R.E.:Multisplitting of a symmetric positive definite matrix, SIAM J.Matrix Anal.Appl., 11(1990), 69-82.

[6] Neumann,M.,Plemmons,R.G.: Convergence of parallel multisplitting iterative methods for M-matrices, Linear Algebra Appl., 88(1987), 559-573.

[7] Elsner,L.: Comparisons of weak regular splittings and multisplitting methods, Numer.Math., 56(1989), 283-289.

[8] Frommer,A.: Parallel nonlinear multisplitting methods, Numer. Math.,56(1989), 269-282.

[9] Bru,R., Elsner,L., Neumann,M.: Models of parallel chaotic iterative methods,Linear Algebra Appl., 103(1988), 175-192.

[10] Chazan,D., Miranker,W.L.: Chaotic relaxation, Linear Algebra Appl.,2(1969), 199-222.

[11] Su Yangfeng, Zhu shidong: Models of parallel multisplitting chaotic iterations, J.Fudan University(Natural Science), 4(1991), 444-450(in Chinese).

[12] Schechter,M.: Relaxation methods for convex problems, SIAM J.Numer.Anal., 5(1968), 601-612.

[13] Feng Guochen: On the global convergence of relaxation methods for solving systems of nonlinear equations, Numer.Math. A J. of Chinese Univ., 1(1980), 7-14(in Chinese).

[14] Varga,R.S.: Matrix Iterative Analysis, Englewood Cliffs, N.J.: Prentice-Hall, 1961.

[15] Hu Jiagan: Scaling transformation and convergence of splittings

of matrix, Math.Numer.Sinica, 1(1983), 72-78(in Chinese).

[16] Ortega,J.M, Rheinboldt,W.C.: Iterative Solution of Nonlinear Equations in Several Variables, New York: Academic Press, 1970.

[17] Wang Deren, Bai Zhongzhi, Evans,D.J.: A class of asynchronous parallel matrix multisplitting relaxation methods, to appear (1992).

[18] Bank,R.E., Dupont,T.F., Yserentant,H.: The hierarchical basis multigrid methods, Numer.Math., 52(1988), 427-458.

[19] Axelsson,O., Vassilevski,P.S.: Algebraic multilevel precondition-ing methods I, Numer.Math., 56(1989), 157-177.

[20] Axelsson,O., Vassilevski,P.S.: Algebraic multilevel precondition-ing methods II, SIAM J.Numer.Anal., 27(1990), 1569-1590.

[21] Axelsson,O., Vassilevski,P.S.: A survey of multilevel precondi-tioned iterative methods, BIT, 29(1989), 769-793.

[22] Ewing,R.E., Lazarov,R.D., Vassilevski,P.S.: Local refinement techniques for elliptic problems on cell-centered grids III-Algebraic multilevel BEPS preconditioners, Numer.Math., 59(1991), 431-452.

[23] Ciarlet,P.G.: The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam, 1978.

[24] Wang Deren, Sun Baoyung: A parallel algorithm for a class of nonlinear equations applicable to MIMD systems, Math. Numer. Sinica, 3(1991), 297-306(in Chinese).

[25] Hu Jiagan: Convergence of a generalized iterative matrix, Math. Numer.Sinica, 2(1984), 174-181(in Chinese).

[26] Frommer,A., Mayer,G.: Parallel interval multisplittings, Numer. Math., 56(1989), 255-267.

[27] Rheinboldt,W.C.: On M-functions and their application to nonlinear Gauss-Seidel iterations and to network flows, J.Math.Anal.Appl.,32(1970), 274-307.

[28] White,R.E.: A nonlinear parallel algorithm with applications to the Stefan problems, SIAM J.Numer.Anal., 23(1986), 639-652.

[29] White,R.E.: Parallel algorithms for nonlinear problems, SIAM J.Alg.Disc.Methods, 7(1986), 137-149.

[30] Schechter,M.: Iteration methods for nonlinear problems,

Trans.Amer.Math.Soc., 104(1962), 179-189.

[31] Young,D.M.: Iterative Solution of Large Linear Systems, New York:Academic Press, 1971.

[32] Hageman,L.A., Young,D.M.: Applied Iterative Methods, New York: Academic Press, 1981.

[33] Axelsson,O., Barker,V.A.: Finite Element Solution of Boundary Value Problems, New York:Academic Press, 1984.

[34] Bank,R.E., Dupont,T.F.: An optimal order process for solving finite elementMath.Comp., 36(1981), 35-51.

[35] Brass,D.: The contraction number of a multigrid method for solving Poisson Numer.Math., 37(1981), 387-404.

[36] Hachbush,W.: Multigrid Methods and Applications, Berlin Heidelberg New York: Springer, 1985.

[37] Young,D.M.: Convergence properties of the symmetric and unsymmetric successive overrelaxation method and related methods, Math. Comp., 24(1970), 793-807.

[38] Yserentant,H: On the multi-level splitting of finite element spaces, Numer.Math., 49(1986), 379-412.

[39] Yserentant,H.: On the multi-level splitting of finite element spaces for indefinite elliptic boundary value problems, SIAM J.Numer.Anal., 23(1986), 581-595.

[40] Baudet,G.M.: Asynchronous iterative methods for multi-processor, J.Assoc.Comp.Mach., 25(1978), 226-244.

[41] Bertsekas,D.P.: Distributed asynchronous computation of fixed points, Math.Programming, 27(1983), 107-120.

[42] Lubachevsky,B., Mitra,D.: A chaotic asynchronous algorithm for computing the fixed point of a nonnegative matrix of unit spectral raduis, J.Assoc.Comp.Mach., 33(1986), 130-150.

[43] Golub,G.H., Van Loan,C.F.: Matrix Computations, Baltimore:The Johns Hopkins University Press,1983.

[44] Householder,A.S.: The Theory of Matrices in Numerical Analysis, Walthm, Mass.:Ginn/Blaisdell, 1964.

[45] Wilkinson,J.H.: The Algebraic Eigenvalue Problem, New York:Oxford University Press, 1965.

[46] Hockney,R.W., Jesshope,C.R.: Parallel Computers: Architecture,

Programming and Algorithms, Adam Hilger Ltd, Bristol, 1981.

[47] Berman,A., Plemmons,R.J.:Nonnegative Matrices in the Mathematical Science, New York:Academic Press, 1979.

[48] Elsner,L., Neumann,M., Vemmer,B.: The effect of the number of processors onblock Jacobi method, Linear Algebra Appl., 154-156 (1991), 311-330.

[49] Evans,D.J.,Wang Deren: An asynchronous parallel algorithm for solving a class of nonlinear simultaneous equations, Parallel Comput.,17(1991), 165-180.

[50] El-Tarazi,M.N.: Some convergence results for asynchronous algorithms, Numer.Math., 39(1982), 325-340.

[51] Bertsekas,D.P., Tsitsiklis,J.N.: Parallel and Distributed Computation-Numerical Methods, Prentice-Hall, Englewood Cliffs, N.J., 1989.

[52] Ortega,J.M.: Introduction to Parallel and Vector Solution of Linear Systems, New York and London:Plenum Press, 1989.

[53] Hwang,K., Briggs,F.A.: Computer Architecture and Parallel Processing, New York:McGraw-Hill, 1984.

[54] Dennis,J.E.Jr., Schnabel,R.B.: Numerical Methods for Constrained Optimization and Nonlinear Equations, Prentice-Hall, Englewood Cliffs, N.J., 1983.

[55] Quinn,M.J.: Designing Efficient Algorithms for Parallel Computers, New York:McGraw-Hill, 1987.

[56] Gantmacher,F.R.: The Theory of Matrices, New York:Chelsea, 1959.

[57] Hadjidimos,A., Psimarni,A., Yejios,A.: On the convergence of some generalized iterative methods, Linear Algebra Appl., 75(1986), 117-132.

[58] Hadjidimos,A.: Accelerated overrelaxation method, Math.Comp., 32 (1978), 149-157.

[59] Hadjidimos,A.: On the generalization of the basic iterative methods for the solution of linear systems, Intern.J.Computer Math., 14(1983), 335-319.

[60] Albrecht,P., Klein,M.P.: Extrapolated iterative methods for linear systems, SIAM J.Numer.Anal., 21(1984), 192-201.

[61] Papatheodorous,T.S.: Block AOR iteration for nonsymmetric

matrices, Math.Comp., 41(1983), 511-525.

[62] Hoyer,W., Schmidt,T.W.: Newton-type decomposition methods for equations arising in network analysis, ZAMM, 64(1984), 397-405.

[63] Wang Deren, Bai Zhongzhi: Relaxed parallel multisplitting methods for large-scale linear least-square problems, Numer.Math.A J.of Chinese Univ., 1(1993), 1-14(in Chinese).

[64] Evans,D.J., Wang Deren, Bai Zhongzhi: Matrix multi-splitting multi-parameter relaxation methods, Intern.J.Computer Math., 43(1992), 173-188.

[65] Bai Zhongzhi, Wang Deren: Generalized matrix multispliting relaxation methods and their convergence, Numer.Math.A J.of Chinese Univ.(Seri.B), 1992.

[66] Wang Deren, Bai Zhongzhi: Convergence analysis for a class of matrix factorization update Quasi-Newton methods, Comm.on Appl.Math.andComput., 1(1991), 50-60(in Chinese).

[67] Bai Zhongzhi, Wang Deren: On the convergence of factorization update algorithm, J.Comp.Math., 1(1993).

[68] Wang Deren, Bai Zhongzhi: On the monotone convergence of matrix multisplitting iteration methods, Proc. of the 92'Shanghai Intern. andthe 6-th National Conference on Numer.Algebra, Shanghai, 1992.

[69] Wang Deren, Bai Zhongzhi, Evans,D.J.: A class of asynchronous nonlinear multisplitting relaxation methods, to appear(1993).

[70] Feng Guochen: Iterative Methods of System of Nonlinear Equations, Shanghai Press of Science and Technology, 1989(in Chinese).

[71] Ortega,J.M., Voigt,R.G.: Solution of partial differential equations on vector and parallel computers, SIAM Review, 27(1985), 149-240.

[72] Hwang,K., Cheng,Y.H.: Partitioned algorithm for VLSI arithmetic systems, IEEE Trans.Computers, 215-224(1982), December.

[73] Stout,Q.: "Divide and Conquer" algorithms for parallel image processing, J. Parallel and Distributed Computing, 2(1987)

Parallel Iterative Methods
for Large-scale Systems of Algebraic Equations

Ph.D. student: Bai Zhongzhi
Supervisors: Guo Benyu and Wang Deren

Abstract

This dissertation is constructed by two parts.

The first part is emphasized on the designs of parallel algorithms for solving the large-scale system of linear equations

$Ax=b,\ A\in L(R^n)$ nonsingular, $x,\ b\in R^n$.

By making use of the concept of matrix multisplitting and applying the relaxation acceleration techniques of the matrix multisplitting, we set up general framework of parallel matrix multisplitting relaxed algorithms; and based on the principle of sufficiently using the delayed information, we design a class of asynchronous parallel matrix multisplitting accelerated relaxation algorithms, which can greatly execute the efficiency of the practical computations of the corresponding MIMD-systems. When the coefficient matrix $A\in L(R^n)$ is symmetric positive and is of block type, stimulated by the multilevel iteration idea, we construct a class of parallel multilevel block iteration algorithms. The convergence theories of these algorithms are researched in depth. Meanwhile, their convergence rates are estimated in detail.

In the second part, we mainly study how to design parallel algorithms to get the solution of the large-scale system of nonlinear equations

$F(x)=0,\quad F:\ \mathfrak{D}\subseteq R^n \longrightarrow R^n.$

In the sense of the nonlinear multisplitting , by making use of the relaxation acceleration techniques of the nonlinear multisplitting, we propose a class of parallel nonlinear multisplitting accelerated relaxation algorithms; which cover many practical and efficient nonlinear multisplitting relaxed type of algorithms, such as the multisplitting AOR-Newton, -Chord, and -Steffensen algorithms,etc.. Under proper conditions, we study the local convergence, the monotone convergence as well as the global convergence of these new algorithms, and deep theory results are therefore obtained. A lot of numerical tests show that our algorithms are feasible and efficient, all the numerical results are thorothly coincide with the new theories. Again, based on the principle of sufficiently using the delayed information, we design a class of asynchronous parallel nonlinear multisplitting accelerated relaxation algorithms, which are suitable to the corresponding MIMD-systems. These asynchronous algorithms also cover many high-valuable and practical asynchronous parallel nonlinear multisplitting relaxed type of algorithms, such as the asynchronous multisplitting AOR-Newton, -Chord, and -Steffensen algorithms, etc. .We research the convergence theories of these algorithms in depth and estimate their convergence rates in detail, too.

Key words: system of linear equations, matrix multisplitting, L-matrix, M-matrix, H-matrix, symmetric positive matrix, multilevel iteration, preconditioning matrix, condition number, system of nonlinear equations, nonlinear multisplitting, synchronous parallel, asynchronous parallel, accelerated relaxation, local convergence, monotone convergence, global convergence, convergence rates.