应用数学和力学, 第 20 卷 第 2 期(1999 年 2 月)
Applied Mathematics and Mechanics

应用数学和力学编委会编
重 庆 出 版 社 出 版

论文编号: 1000_0887(1999)02_0169_174

并行区间矩阵多分裂 AOR 算法及其收敛性*

白中治

中国科学院计算数学与科学工程计算研究所, 北京 100080

(张汝清推荐)

摘要: 本文提出了一类求解大型区间线性方程组的并行区间矩阵多分裂松弛算法, 并在系数矩阵是区间 H_矩阵的条件下, 建立了这类算法的收敛理论·

关键词: 并行算法; 区间矩阵多分裂; 松弛; 收敛性
分类号: O241.6 文献标识码: A

引 言

用 IR^n 和 $IR^{n\times n}$ 分别表示实 n 维区间向量和实 $n\times n$ 区间矩阵· 给定 $A\in IR^{n\times n}$ 非奇异 (即所有的 $A\in A$ 均非奇异), $b\in IR^n$, 为了确定集合

$$S=\left\{x\in R^n \mid Ax=b, A\in A,\ b\in b\right\},$$

Frommer 和 Mayer[1] 于 1989 年率先运用区间矩阵多分裂的概念, 建立了一类并行区间矩阵多分裂算法, 并在一定条件下, 证明了这类算法的收敛性·

受上述思想的启发, 本文构造了一类确定包含解集合 S 的区间向量 $x^*\in IR^n$ 的并行区间矩阵多分裂AOR算法· 这类新算法以并行区间矩阵多分裂SOR, Gauss_Seidel 及 Jacobi 等许多有效而实用的算法为特例· 通过进一步引进松弛参数, 我们给出了上述算法的外插形式· 在对松弛参数和区间矩阵多重分裂的适当限制下, 当 A 是区间 $H\_$ 矩阵时, 我们讨论了这两类算法的收敛性质· 本文结果是[1]的进一步深入和发展, 亦是[2]对于大型区间线性方程组的推广·

1 区间矩阵多分裂与松弛算法

设 $A=(A_{ij})\in IR^{n\times n}$ 非奇异, $D=\mathrm{diag}(A)$· 对 $k=1,2,\cdots,\alpha$ ($\alpha\leqslant n$ 为一正整数), 分别取 $L_k=(L_{ij}^k)\in IR^{n\times n}$ 为严格下三角矩阵, $U_k=(U_{ij}^k)\in IR^{n\times n}$ 为零对角矩阵而 $E_k\in L(R^n)$ 为非负对角矩阵, 如果

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

* 来稿日期: 1997_02_17
基金来源: 国家自然科学基金资助项目(19601036)
作者简介: 白中治(1965~)男, 副研究员

b) D 非奇异;

c) $\sum_{k=1}^{\alpha} E_k = I \quad (I \in L(R^n)$ 为单位矩阵),

则称 $(D - L_k, U_k, E_k)\ (k = 1, 2, \cdots, \alpha)$ 是区间矩阵 A 的一个多分裂·

基于此概念,我们可建立如下的包含解集 S 的并行区间矩阵多分裂 AOR 算法·

算法 I 给定初始向量 $x^0 \in IR^n$,对 $m = 0, 1, 2, \cdots$ 计算

$$x^{m+1} = \sum_{k=1}^{\alpha} E_k x^{m,k},$$

其中对 $k = 1, 2, \cdots, \alpha$,

$$x^{m,k} = (D - rL_k)^F[(1 - \omega)D + (\omega - r)L_k + \omega U_k]x^m + \omega(D - rL_k)^F b,$$

这里,$r \geqslant 0$ 为松弛参数, $\omega > 0$ 为加速参数;对每个 $k \in \{1, 2, \cdots, \alpha\}$ 和任意 $v = (v_1, v_2, \cdots, v_n)^{\mathrm{T}} \in IR^n$,区间向量 $u = (u_1, u_2, \cdots, u_n)^{\mathrm{T}} = (D - rL_k)^F v$ 定义为

$$u_i = \left(v_i - r\sum_{j=1}^{i-1} L_{ij}^k u_j\right) / A_{ii}, \qquad (i = 1, 2, \cdots, n)·$$

由[3]知,对任意 $D \in D, L_k \in L_k, v \in v$,恒有 $(D - rL_k)^{-1} \in (D - rL_k)^F\ (k = 1, 2, \cdots, \alpha)$·

显然,对 $k = 1, 2, \cdots, \alpha$, $x^{m,k}$ 的计算相互独立,且相应于 E_k 的对角元素为零的 $x^{m,k}$ 的分量可不被计算,因此,该算法不仅具有很强的并行计算功能,而且还可大大节省计算工作量· 由于该算法含有两个任意参数,这不仅使得它在具体实施中更加灵活与实用,而且通过参数的不同选取,还可产生许多有效的并行区间矩阵多分裂松弛算法;特别是,当选取松弛参数对 (r, ω) 分别为 $(\omega, \omega), (1, \omega), (1, 1), (0, \omega)$ 和 $(0, 1)$ 时,即可相应地得到区间矩阵多分裂 SOR,外插 Gauss_Seidel, Gauss_Seidel,外插 Jacobi 及 Jacobi 算法·

另外,若 $\alpha = 1$, $-L_1$, $-U_1 \in IR^{n\times n}$ 分别为 A 的严格下、上三角部分,则算法 I 即化为已知的区间 AOR 算法[4]·

现分别引进区间矩阵

$$\mathscr{L}_{\mathrm{IMAOR}}(r, \omega) \overset{1}{=} \sum_{k=1}^{\alpha} E_k(D - rL_k)^F[(1 - \omega)D + (\omega - r)L_k + \omega U_k] \tag{1.1}$$

和区间向量

$$b_{\mathrm{IMAOR}}(r, \omega) = \omega\sum_{k=1}^{\alpha} E_k(D - rL_k)^F b,$$

则算法 I 可等价地表示成

$$x^{m+1} = \mathscr{L}_{\mathrm{IMAOR}}(r, \omega)x^m + b_{\mathrm{IMAOR}}(r, \omega), \qquad (m = 0, 1, 2, \cdots)· \tag{1.2}$$

通过进一步引进松弛因子 $\beta > 0$,即可得到算法 I 的外插形式如下:

算法 II 给定初始向量 $x^0 \in IR^n$,对 $m = 0, 1, 2, \cdots$,计算

$$x^{m+1} = \sum_{k=1}^{\alpha} E_k x^{m,k},$$

其中对 $k = 1, 2, \cdots, \alpha$,

$$x^{m,k} = \{\beta(D - rL_k)^F[(1 - \omega)D + (\omega - r)L_k + \omega U_k] + (1 - \beta)I\}x^m + \beta(D - rL_k)^F(\omega b),$$

这里,$r \geqslant 0$ 为松弛参数, $\omega, \beta > 0$ 为加速参数·

同样地,若分别定义区间矩阵

$$\mathscr{L}_{\mathrm{IMAORR}}(r, \omega, \beta) = \beta\sum_{k=1}^{\alpha}E_k(D - rL_k)^F[(1-\omega)D + (\omega - r)L_k + \omega U_k] + (1-\beta)I \tag{1.3}$$

和区间向量

$$b_{\mathrm{IMAORR}}(r, \omega, \beta) = \beta\sum_{k=1}^{\alpha}E_k(D - rL_k)^F(\omega b),$$

则算法Ⅱ也可写成如下等价形式:

$$x^{m+1} = \mathscr{L}_{\mathrm{IMAORR}}(r, \omega, \beta)x^m + b_{\mathrm{IMAORR}}(r, \omega, \beta), \qquad (m = 0, 1, 2, \cdots). \tag{1.4}$$

由(1.1)和(1.3)知,算法Ⅰ与算法Ⅱ的迭代矩阵之间成立如下关系:

$$\mathscr{L}_{\mathrm{IMAORR}}(r, \omega, \beta) = \beta\mathscr{L}_{\mathrm{IMAOR}}(r, \omega) + (1-\beta)I. \tag{1.5}$$

2 记 号 与 引 理

在下面的讨论中,总认为退化的区间向量(矩阵)为 $R^n(L(R^n))$ 中的元素,为简便计,下文中的区间向量(矩阵)或实向量(矩阵)用向量(矩阵)代替,其含义依上下文确定. 另外,为简化定义与叙述,视 $n\times 1$ 矩阵为向量而 1×1 矩阵为实数(区间).

设 $A = (A_{ij}), B = (B_{ij}) \in IR^{n\times n}$. 若对所有的 $A \in A, B \in B$, 均有 $A \geqslant B(>)$, 则称 $A \geqslant B(>)$. A 的绝对值定义为 $|A| = \sup\{|A| \mid A \in A\}$, 显然它是 $L(R^n)$ 中的非负矩阵, 对所有的 $A \in A$, $|A| \leqslant |A|$, 且对任何 $B \in IR^{n\times n}$, 成立 $|AB| \leqslant |A||B|$. 与之有关的性质可详见文[3]和[4]. 用 $\langle A\rangle = (\alpha_{ij}) \in L(R^n)$ 表示区间矩阵 A 的比较矩阵, 其中

$$\alpha_{ij} = \begin{cases} \inf\{|A_{ii}| \mid A_{ii} \in A_{ii}\}, & i = j, \\ -|A_{ij}|, & i \neq j, \end{cases} \quad i, j = 1(1)n,$$

而

定义 $\tau = \max\limits_{1\leqslant i\leqslant n}\{|A_{ii}|/\alpha_{ii} \mid \alpha_{ii} \neq 0\}$, 则显然 $\tau \geqslant 1$. 称 A 是区间 $H_$矩阵, 若 $\langle A\rangle$ 是 $M_$矩阵, 即 $\langle A\rangle^{-1} \geqslant 0$.

上述概念可自然地诱导出 $IR^n, L(R^n)$ 和 R^n 中相应的概念.

我们现引进若干有用的引理.

引理1[3] 每个 $H_$矩阵 $A \in IR^{n\times n}$ 均是非奇异的.

引理2 设 $(D - L_k, U_k, E_k)$ $(k = 1, 2, \cdots, \alpha)$ 是 $H_$矩阵 $A \in IR^{n\times n}$ 的一个区间多分裂, 则对每个 $k \in \{1, 2, \cdots, \alpha\}$, 由算法Ⅰ定义的算子 $(D - L_k)^F$ 满足

$$|(D - L_k)^F| \leqslant (\langle D\rangle - |L_k|)^{-1}.$$

引理3[5] 设 $A \in IR^{n\times n}$. 若 $|A|$ 的谱半径小于1, 即 $\rho(|A|) < 1$, 则以下结论为真:

a) $x = Ax + d(d \in IR^n)$ 有唯一解 $x^* \in IR^n$;

b) 以 $x^0 \in IR^n$ 为初始的迭代 $x^{m+1} = Ax^m + d(m = 0, 1, \cdots)$ 收敛到 x^*;

c) 若 $x^0 \subseteq x^1$, 则 $x^0 \subseteq x^*$.

3 算法的收敛性

定理1 设 $A \in IR^{n\times n}$ 为区间 $H_$矩阵, $(D - L_k, U_k, E_k)(k = 1, 2, \cdots, \alpha)$ 为 A 的一个区间多分裂且满足

$$\langle A\rangle = \langle D\rangle - |L_k| - |U_k| \equiv \mathscr{D} - \mathscr{B}(\mathscr{D} = \langle D\rangle), \qquad (k = 1, 2, \cdots, \alpha). \tag{3.1}$$

则当松弛参数对 (r,ω) 满足条件(3.2) ~ (3.4) 之一, 即

$$0\leqslant r\leqslant\omega,\ (\tau-1)/(\tau-\rho(\mathscr{D}^{-1}\mathscr{B}))<\omega<(\tau+1)/(\tau+\rho(\mathscr{D}^{-1}\mathscr{B}));\tag{3.2}$$

$$\begin{cases}(\tau-1)/(\tau+\rho(\mathscr{D}^{-1}\mathscr{B}))<\omega<(\tau+1)/(\tau-\rho(\mathscr{D}^{-1}\mathscr{B}))\\ \omega\leqslant r<(1+\omega\rho(\mathscr{D}^{-1}\mathscr{B})-|1-\omega|\tau)/(2\rho(\mathscr{D}^{-1}\mathscr{B}))\end{cases},\quad \rho(\mathscr{D}^{-1}\mathscr{B})\neq 0;\tag{3.3}$$

$$\omega\leqslant r,\ 1-\frac{1}{\tau}<\omega<1+\frac{1}{\tau},\ \rho(\mathscr{D}^{-1}\mathscr{B})=0\tag{3.4}$$

之一时, 从 $x^0\in IR^n$ 出发由算法I产生的区间序列$\{x^m\}$收敛到某包含 S 的区间向量 $x^*\in IR^n$•

证明 在定理的条件下, 我们首先验证

$$\rho(\mathscr{L}_{\mathrm{IMAOR}}(r,\omega))<1,\tag{3.5}$$

这里, $\mathscr{L}_{\mathrm{IMAOR}}(r,\omega)$ 由(1.1) 定义•

利用引理2和(3.1)得

$$\begin{aligned}|\mathscr{L}_{\mathrm{IMAOR}}(r,\omega)|&\leqslant\sum_{k=1}^{\alpha}E_k|(D-rL_k)^F|[|1-\omega||D|+|\omega-r||L_k|+\omega|U_k|]\leqslant\\&I+\sum_{k=1}^{\alpha}E_k(\langle D\rangle-r|L_k|)^{-1}\langle D\rangle[|1-\omega||D|\langle D\rangle^{-1}-\\&I+(|\omega-r|+r)\mathscr{D}^{-1}\mathscr{B}]\leqslant\\&I+\sum_{k=1}^{\alpha}E_k(\langle D\rangle-r|L_k|)^{-1}\langle D\rangle[(|1-\omega|\tau-1)I+\\&(|\omega-r|+r)(\mathscr{D}^{-1}\mathscr{B}+\varepsilon ee^{\mathrm{T}})],\end{aligned}\tag{3.6}$$

其中 ε 是任意的正实数, 而 $e=(1,1,\cdots,1)^{\mathrm{T}}\in R^n$•

简记 $J=\mathscr{D}^{-1}\mathscr{B}$• 由于 A 是区间H_矩阵, 故 $\rho(J)<1$• 注意到松弛参数的取值范围(3.2) ~ (3.4) 等价于

$$|1-\omega|\tau+(|\omega-r|+r)\rho(J)<1,$$

利用谱半径的连续性, 可取 ε 充分小, 使得

$$\rho_\varepsilon=\rho(J+\varepsilon ee^{\mathrm{T}})<1,\tag{3.7}$$

并且

$$|1-\omega|\tau+(|\omega-r|+r)\rho_\varepsilon<1•\tag{3.8}$$

根据非负矩阵理论中的Perron_Frobenius[6]定理, 存在正向量 $x_\varepsilon\in R^n$ 满足

$$(J+\varepsilon ee^{\mathrm{T}})x_\varepsilon=\rho_\varepsilon x_\varepsilon•\tag{3.9}$$

另外, 对 $k=1,2,\cdots,\alpha$, 因 $\langle D\rangle-r|L_k|\leqslant\langle D\rangle$, 故 $(\langle D\rangle-r|L_k|)^{-1}\langle D\rangle\geqslant I$• 由(3.6) ~ (3.9) 得

$$\begin{aligned}|\mathscr{L}_{\mathrm{IMAOR}}(r,\omega)|x_\varepsilon&\leqslant[1+(|1-\omega|\tau-1)+(|\omega-r|+r)\rho_\varepsilon]x_\varepsilon=\\&[|1-\omega|\tau+(|\omega-r|+r)\rho_\varepsilon]x_\varepsilon<x_\varepsilon•\end{aligned}$$

从而, [6, p.47] 练习2保证(3.5)成立•

由引理3a)、b)知, $x=\mathscr{L}_{\mathrm{IMAOR}}(r,\omega)x+b_{\mathrm{IMAOR}}(r,\omega)$ 有唯一解 $x^*\in IR^n$, 并且由算法I从 $x^0\in IR^n$ 出发产生的区间序列$\{x^m\}$收敛到此解 x^*•

任取 $x\in S$, 则存在 $A\in A$, $b\in b$, 使得 $Ax=b$• 现对于 $k=1,2,\cdots,\alpha$, 分裂 $A=D-L_k-U_k$, 其中 $D\in D$, $L_k\in L_k$, $U_k\in U_k$, 则有

$$x=\sum_{k=1}^{\alpha}E_k(D-rL_k)^{-1}[(1-\omega)D+(\omega-r)L_k+\omega U_k]x+\sum_{k=1}^{\alpha}E_k(D-rL_k)^{-1}(\omega b)\in$$

$$\sum_{k=1}^{\alpha} E_k(D - rL_k)^F[(1-\omega)D + (\omega - r)L_k + \omega U_k]x + \sum_{k=1}^{\alpha} E_k(D - rL_k)^F(\omega b) = \mathscr{L}_{IMAOR}(r,\omega)x + b_{IMAOR}(r,\omega).$$

选取 $x^0 = x$, 上式表明 $x^0 \in x^1 = \mathscr{L}_{IMAOR}(r,\omega)x + b_{IMAOR}(r,\omega)$, 进而由引理3c) 知, $x \in x^*$. 因此, $S \subseteq x^*$. □

定理 2 设 $A \in IR^{n\times n}$ 为区间H_矩阵, $(D - L_k, U_k, E_k)(k = 1, 2, \dots, \alpha)$ 为 A 的一个区间多分裂且满足(3.1), 则当松弛参数 r, ω, β 满足条件(3.10) ~ (3.12) 之一, 即

$$\begin{cases} 0 \leqslant r \leqslant \omega, \ (\tau - 1)/(\tau - \rho(\mathscr{D}^{-1}\mathscr{B})) < \omega < (\tau + 1)/(\tau + \rho(\mathscr{D}^{-1}\mathscr{B})), \\ 0 < \beta < 2/(1 + \rho(\mathscr{H}(\omega))), \ \mathscr{H}(\omega) = |1-\omega|\tau I + \omega\mathscr{D}^{-1}\mathscr{B}; \end{cases} \tag{3.10}$$

$$\begin{cases} \omega \leqslant r, \ 1 - \frac{1}{\tau} < \omega < 1 + \frac{1}{\tau}, \quad \rho(\mathscr{D}^{-1}\mathscr{B}) = 0; \\ 0 < \beta < 2/(1 + |1-\omega|\tau) \end{cases} \tag{3.11}$$

$$\begin{cases} (\tau - 1)/(\tau + \rho(\mathscr{D}^{-1}\mathscr{B})) < \omega < (\tau + 1)/(\tau - \rho(\mathscr{D}^{-1}\mathscr{B})), \quad \rho(\mathscr{D}^{-1}\mathscr{B}) \neq 0, \\ \omega \leqslant r < (1 + \omega\rho(\mathscr{D}^{-1}\mathscr{B}) - |1-\omega|\tau)/(2\rho(\mathscr{D}^{-1}\mathscr{B})), \\ 0 < \beta < 2/(1 + \rho(\mathscr{H}(r,\omega))), \ \mathscr{H}(r,\omega) = |1-\omega|\tau I + (2r - \omega)\mathscr{D}^{-1}\mathscr{B} \end{cases} \tag{3.12}$$

之一时, 从 $x^0 \in IR^n$ 出发由算法II产生的区间序列 $\{x^m\}$ 收敛到某包含 S 的区间向量 $x^* \in IR^n$.

证明 由(1.5)和(3.6)可得

$$\begin{aligned} |\mathscr{L}_{IMAORR}(r,\omega,\beta)| &\leqslant \beta|\mathscr{L}_{IMAOR}(r,\omega)| + |1-\beta| I \leqslant \\ &(\beta + |1-\beta| I) + \beta\sum_{k=1}^{\alpha} E_k(\langle D\rangle - r|L_k|)^{-1}\langle D\rangle \cdot \\ &[(\mathscr{H}(r,\omega) + \varepsilon ee^T) - I], \end{aligned} \tag{3.13}$$

其中 $\mathscr{H}(r,\omega) = |1-\omega|\tau I + (|\omega - r| + r)\mathscr{D}^{-1}\mathscr{B}$, ε 是任意的正实数.

由松弛参数的取值范围(3.10) ~ (3.12)知, $\rho(\mathscr{H}(r,\omega)) < 1$ 并且 $|1-\beta| + \beta\rho(\mathscr{H}(r,\omega)) < 1$. 利用谱半径的连续性, 可取 ε 充分小, 使得

$$\rho_\varepsilon = \rho(\mathscr{H}(r,\omega) + \varepsilon ee^T) < 1, \ |1-\beta| + \beta\rho_\varepsilon < 1. \tag{3.14}$$

根据非负矩阵理论中的Perron_Frobenius定理, 存在正向量 $x_\varepsilon \in R^n$, 满足

$$(\mathscr{H}(r,\omega) + \varepsilon ee^T)x_\varepsilon = \rho_\varepsilon x_\varepsilon. \tag{3.15}$$

现依(3.13) ~ (3.15)即得

$$|\mathscr{L}_{IMAORR}(r,\omega,\beta)|x_\varepsilon \leqslant (|1-\beta| + \beta\rho_\varepsilon)x_\varepsilon < x_\varepsilon.$$

再利用[6, p.47]练习2可得 $\rho(|\mathscr{L}_{IMAORR}(r,\omega,\beta)|) < 1$.

由引理3a) ~ b)知, $x = \mathscr{L}_{IMAORR}(r,\omega,\beta)x + b_{IMAORR}(r,\omega,\beta)$ 有唯一解 $x^* \in IR^n$, 并且由算法II从 $x^0 \in IR^n$ 出发产生的区间序列 $\{x^m\}$ 收敛到 x^*. 同样地, 由引理3c)我们也易证得 $S \subseteq x^*$.

致谢: 作者感谢审稿推荐人提出的宝贵意见.

参 考 文 献

[1] Frommer A, Mayer G. Paralled interval multisplitting[J]. Numer Math, 1989, **56**: 255~ 267

[2] Wang D. On the convergence of the parallel multisplitting AOR algorithm[J]. Linear Algebra Appl, 1991, **154/156**: 473~ 486

[3] Neumaier A. New techniques for the analysis of linear interval equations[J]. Linear Algebra Appl, 1984, **58**: 273~ 325

[4] Alefeld G, Herzberger J. Introduction to Interval Equations[M]. New York: Academic Press, 1983

[5] Zhou R. The convergence of interval AOR method[J]. Math Numer Sinica, 1992, **14**(1): 49~ 52

[6] Varga R S. Matrix Iterative Analysis[M]. Englewood Cliffs, N J: Prentice_Hall, INC, 1962

Parallel Interval Matrix Multisplitting AOR Methods and Their Convergence

Bai Zhongzhi

State Key Laboratory of Scientific, Engineering Computing

Institute of Computational Mathematics and Scientific/ Engineering Computing

Chinese Academy of Sciences, P O Box 2719, Beijing 100080, P R China

Abstract: This paper proposes a class of parallel interval matrix multisplitting AOR methods for solving systems of interval linear equations and discusses their convergence properties under the conditions that the coefficient matrices are interval H _matrices.

Key words: parallel method; interval matrix multisplitting; relaxation; convergence