

解高阶 Hermitian 矩阵特征值问题的 并行块消去迭代法^{*1)}

孙家昶 邓健新 曹建文

(中国科学院软件研究所)

THE PARALLEL BLOCK ELIMINATION ALGORITHM FOR SOLVING HERMITIAN MATRIX LARGE EIGENVALUE PROBLEMS

Sun Jiachang Deng Jianxin Cao Jianwen
(*Institute of Software, Chinese Academy Sciences*)

Abstract

A parallel block elimination algorithm for solving Hermitian matrix large eigenvalue problems was provided in this paper. The algorithm possess crude grain parallel properties. The high-quality black-bos for solving matrix eigenvalue problems, multi-processors and it's local memory can be use effectively in this algorithm. It can be implemented on high-performance distributed memory parallel computer.

The convergence, error analysis of the algorithm, and parallel design are presented. A part of the numerical results are listed in this paper.

§ 1. 引 言

在现代计算物理、计算化学、计算生物以及许多科学研究与工程计算中都涉及很高阶的 Hermitian 矩阵特征值问题。对这个问题已有许多的研究成果和计算方法，但是现有的方法或者是不能并行执行或者是仅能计算少量的特征值。对于很高阶的矩阵，要求不是太少量的特征值的问题还没有一个好的并行算法。本文工作的目的是给出一个粗粒度的并行算法，能充分利用高性能大规模分布式存贮的并行计算机巨大潜力解决实际应用的计算问题。

本文给出的并行分块消去迭代法把在单处理器上串行求解矩阵特征值问题的解算器看作“黑盒子”来使用，解算器的方法可以是任意的有效方法，例如 QR 算法。

* 1996年3月26日收到。

1) 国家自然科学基金资助项目，国家基础性攀登项目 B 支持。

在 [6] 中曾提出过块 Jacobi 算法, 它是 Jacobi 算法的推广. 由于算法在本质上并没有改变, 因而仍然停留在一般的并行 Jacobi 算法的水平, 计算量与数据传输量比值小, 实用效率不高. 关于用分块逐次消去迭代的算法在 [6] 中提出过用来解奇异值问题, 但对于算法的收敛性和误差分析以及并行设计等等问题都还没有系统研究. 本文对于并行块消去迭代算法, 给出了总体收敛性和渐近二次收敛性的证明, 也给出算法的误差分析、并行实现与分析. 最后在国家智能计算机研究开发中心研制的大规模并行计算机曙光 1000 上进行数值实验, 文中附有数值结果.

§ 2. 并行块消去迭代算法

假设矩阵特征值问题是

$$Ax = \lambda x, \quad (1)$$

其中 A 是 $n \times n$ Hermitian 矩阵, x, λ 分别是求解的特征向量与对应的特征值. 首先划分 A 为块矩阵

$$A = (A_{ij}), \quad i, j = 1, 2, \dots, s. \quad (2)$$

为叙述方便, 假定 n 可被 s 整除, A_{ij} 都是 n/s 阶的方阵. 并行分块消去迭代法可表示如下:

并行分块消去迭代法. 步 1. 按照某种并行消去迭代次序, 以分块数 $s = 8$, 处理器个数 $p = 4$ 为例, 常用的并行计算的 7 次并行变换为

$$\begin{array}{l} 1 \quad (1, 2) \quad (3, 4) \quad (5, 6) \quad (7, 8) \\ 2 \quad (1, 3) \quad (2, 4) \quad (5, 7) \quad (6, 8) \\ 3 \quad (1, 4) \quad (2, 3) \quad (5, 8) \quad (6, 7) \\ 4 \quad (1, 5) \quad (2, 6) \quad (3, 7) \quad (4, 8) \\ 5 \quad (1, 6) \quad (2, 5) \quad (3, 8) \quad (4, 7) \\ 6 \quad (1, 7) \quad (2, 8) \quad (3, 5) \quad (4, 6) \\ 7 \quad (1, 8) \quad (2, 7) \quad (3, 6) \quad (4, 5) \end{array} \quad (3)$$

其中 (i, j) 表示用“黑盒子”解子问题, 子问题的矩阵由 $A_{ii}, A_{ij}, A_{ji}, A_{jj}$ 组成. 计算的结果是 A_{ij}, A_{ji} 被化为零, A_{ii}, A_{jj} 化为对角形矩阵. 上述 7 次并行变换称为一次扫描 (每一个非对角块都被消去迭代过一次). 以变换 (1,2) 为例, 消去迭代过程的变换可记为

$$A^{(1)} = \begin{pmatrix} D_1 & 0 & & \\ & D_2 & & \\ & & A_{ij}^{(1)} & \\ & & & A_{ij} \end{pmatrix} = \begin{pmatrix} Q_1^H & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} & & \\ & A_{22} & & \\ & & A_{ij} & \\ & & & A_{ij} \end{pmatrix} \begin{pmatrix} Q_1 & 0 \\ 0 & I \end{pmatrix}, \quad (4)$$

其中 A_{ij}, A_{ji} 表示该位置上的矩阵的当前值, 并且

$$\begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix} = Q_1^H \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} Q_1, \quad A_{ij}^{(1)} = Q_1^H A_{ij}, \\ A_{ji}^{(1)} = A_{ij}^H Q_1, \quad Q^{(1)} = Q_0 Q_1. \quad (5)$$

步 2. 计算

$$\text{off}(A) = \sum_{i \neq j} \|A_{ij}\|^2, \quad (6)$$

其中矩阵范数 $\|A\|$ 定义为 $\|A\|^2 = \sum_{i,j} |a_{ij}|^2$. 若 $\text{off}(A) \leq \delta$, 则停止; 否则, 继续迭代 (执行步 1 和步 2), 其中 δ 是给定的精度控制参数.

算法的工作量. 以用 QR 黑盒子解子问题为例, 由 (5), (6) 可知, 每次变换的工作量如下: 解子问题 $-5(n/k)^3$, 矩阵乘积 $-(k-1)(n/k)^3$, 特征向量累积 $-k \cdot n(n/k)^2$. 一次扫描有 $k(2k-1)$ 次变换, 其中 $2k = s$. 因此一次扫描的工作量为 $(k(2k-1)(5 + (k-1) + k^2)/k^3)n^3$ (FLOP). 通常用 4 次扫描可以使计算解达到很高的精度, 需要的计算量为 $(8k + 8 + 56/k - 16/k^2)n^3$ (FLOP). 以 $k = 2, 3, 4, 5$ (即 $s = 4, 9, 16, 25$) 为例, 算法的工作量分别相当于 QR 的 9.6, 9.8, 10.6, 11.7 倍. 由此可见, 该算法不是一个好的串行算法, 但它容易并行实现, 算是一个好的并行算法 (见 §5).

§ 3. 算法的收敛性

并行块消去迭代算法的收敛性证明基于矩阵范数压缩的原理, 基本思想如下:

(i) 对于任给 $n \times n$ 阶的 Hermitian 矩阵 $A = (a_{ij})$, $\|A\|^2 = \sum_{i=1}^n |\lambda_i|^2$, 其中 λ_i 是 A 的特征值;

(ii) 矩阵范数酉变换不变, 即若 $P^H P = I$, 则

$$\|P^H A P\| = \|P^H A\| = \|A P\| = \|A\|. \quad (7)$$

若分块消去迭代变换使得 $\text{off}(A)$ 逐渐压缩, 最终 A 化为对角形, 则其对角元必是 A 的特征值.

3.1. 并行块消去迭代法的总体收敛性. 由式 (5) 可知, 子问题 (i, j) 的对角化使两个非对角子块被化零, 因此矩阵的非对角元范数减少了 $2\|A_{ij}\|$, 即

$$\text{off}(A) - \text{off}(A^{(1)}) = 2\|A_{ij}\|^2 \geq 0. \quad (8)$$

例如算法使得每一个子问题的 $\|A_{ij}\|$ 不少于所有非对角子块范数的平均值, 那么

$$\text{off}(A^{(k)}) = \left(1 - \frac{1}{N}\right)^k \text{off}(A), \quad (9)$$

其中 $N = s(s-1)/2$. 若用 p 个处理器并行迭代, 加速接近 p , 实际计算表明, 并行执行上述并行块消去迭代法所用的通信时间不多.

3.2. 算法的渐近收敛性.

引理. 对于形如下式的矩阵等式:

$$\begin{pmatrix} D_1 & W \\ W^H & D_2 \end{pmatrix} \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} \begin{pmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{pmatrix}, \quad (10)$$

$$\begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} = I, \quad (11)$$

若

$$\|W\| = O(\varepsilon), \quad D_1 = \text{diag}(d_1, d_2, \dots, d_m),$$

$$D_2 = \text{diag}(d_{m+1}, d_{m+2}, \dots, d_{2m}), \quad \Lambda_1 = \text{diag}(\mu_1, \mu_2, \dots, \mu_m),$$

$$\Lambda_2 = \text{diag}(\mu_{m+1}, \mu_{m+2}, \dots, \mu_{2m}),$$

则

$$\|X_{12}\| = O(\varepsilon), \quad \|X_{21}\| = O(\varepsilon),$$

$$X_{11}^H X_{11} = I + O(\varepsilon), \quad X_{22}^H X_{22} = I + O(\varepsilon). \quad (12)$$

证明. 设向量 $\begin{pmatrix} l \\ f \end{pmatrix}$ 是 (10) 式中 $\begin{pmatrix} x_{11} \\ x_{12} \end{pmatrix}$ 的第 i 列, 由 (10) 式, 有

$$\begin{pmatrix} D_1 & W \\ W^H & D_2 \end{pmatrix} \begin{pmatrix} l \\ f \end{pmatrix} = \mu_i \begin{pmatrix} l \\ f \end{pmatrix}, \quad 1 \leq i \leq m, \quad (13)$$

于是

$$W^H l + D_2 f = \mu_i f, \quad (D_i - \mu_i I) f = -W^H l,$$

$$f = -(D_2 - \mu_i I)^{-1} W^H l, \quad (14)$$

$$\|f\| \leq c \cdot \|W\| \|l\|, \quad (15)$$

其中 c 是一个相对定数, 由下式确定

$$c = \frac{1}{\min_{m+1 \leq j \leq 2m} (d_j - \mu_i)}$$

由引理假定可得

$$\|f\| \leq O(\varepsilon) \|l\|. \quad (16)$$

注意到

$$\begin{pmatrix} l \\ f \end{pmatrix}^H \begin{pmatrix} l \\ f \end{pmatrix} = \|l\|^2 + \|f\|^2 = 1,$$

从而可推得

$$\|X_{21}\| = O(\varepsilon).$$

类似地可证明

$$\|X_{12}\| = O(\varepsilon).$$

由 (11) 式可得出

$$X_{11}^H X_{11} + X_{21}^H X_{21} = I, \quad X_{12}^H X_{12} + X_{22}^H X_{22} = I, \quad (17)$$

从而

$$X_{11}^H X_{11} = I + O(\varepsilon), \quad X_{22}^H X_{22} = I + O(\varepsilon). \quad (18)$$

定理 1. 解 Hermitian 矩阵特征值问题 (1) 的并行块消去迭代算法 (3)—(7) 是渐近“二次收敛”的.

证明. 假设并行块消去迭代算法迭代到这样的阶段, $\|A_{ij}\| = O(\varepsilon)$ ($i, j = 1, 2, \dots, l$, $i \neq j$), 为了叙述简明, 我们以 $l = 4$ 的例子来表示. 假若一次扫描的次序为

$$(1, 2), \quad (3, 4), \quad (1, 4), \quad (2, 3), \quad (1, 3), \quad (2, 4),$$

我们考察某一非对角块, 例如 A_{14} 的变化历程.

对于 (1, 2) 的变换, A_{14}, A_{24} 的改变如下:

$$\begin{pmatrix} A_{14}^{(1)} \\ A_{24}^{(1)} \end{pmatrix} = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} \begin{pmatrix} A_{14} \\ A_{24} \end{pmatrix},$$

$$A_{14}^{(1)} = Q_{11}A_{14} + Q_{12}A_{24}, \quad A_{24}^{(1)} = Q_{21}A_{14} + Q_{22}A_{24}. \quad (19)$$

由引理可得

$$\|A_{14}^{(1)}\| = \|A_{14}\| + O(\varepsilon^2), \quad \|A_{24}^{(1)}\| = \|A_{24}\| + O(\varepsilon^2).$$

类似地对于所有非零化 A_{14} 的变换, A_{14} 的影响都是 ε^2 数量级的. 我们现在观察 A_{14} 被化零之后重新变为非零的情况, 例如 (1, 4) 变换之后 $A_{14} = 0$ 接着 (2, 4) 变换使 A_{14} 恢复为非零.

$$\begin{pmatrix} A_{12}^H \\ A_{14} \end{pmatrix} = \begin{pmatrix} Q_{22} & Q_{24} \\ Q_{42} & Q_{44} \end{pmatrix} \begin{pmatrix} A_{12}^H \\ 0 \end{pmatrix}, \quad (20)$$

于是

$$A_{14}^{H(1)} = Q_{42}A_{12}^H, \quad \|A_{14}^{H(1)}\| \leq \|Q_{42}\| \cdot \|A_{12}^H\|.$$

由引理有

$$\|A_{14}^{H(1)}\| \leq O(\varepsilon^2). \quad (21)$$

容易看出继续非零化 A_{14} 的变换, A_{14} 将保持 ε^2 量级.

对于一般情况的证明, 仅仅是符号的复杂与书写的冗繁, 按习惯我们对于非对角块范数为 $O(\varepsilon)$ 经一次扫描之后变为 $O(\varepsilon^2)$ 的过程称为渐近“二次收敛”.

§ 4. 算法误差分析

4.1. 特征值的误差

定理 2. 假设执行并行块消去迭代法的 k 次变换之后, 记

$$A^{(k)} = D + E, \quad (22)$$

其中 $D = \text{diag}(d_1, d_2, \dots, d_n)$, $E = (e_{ij})$, $e_{ii} = 0$, $e_{ij}^H = e_{ji}$. 若 $\|E\| \leq \delta$, 则 $|d_i - \lambda_i| \leq \delta$.

证明. 按照 Weyl-Ludaski 定理^[7], 在 D 作适当的置换后必有

$$|\lambda_i(A^{(k)}) - d_i| \leq \|E\|. \quad (23)$$

计算的特征值的误差界为

$$|d_i - \lambda_i| \leq \delta. \quad (24)$$

若 $A^{(k)}$ 的非对角元达到 $O(\varepsilon)$ 量级, 那么上述的界表明计算的特征值误差界也是 $O(\varepsilon)$ 量级, 当假定 $d_i \neq d_j$ ($i \neq j, i, j = 1, 2, \dots, n$) 时还可证明误差的界达到 $O(\varepsilon^2)$ 量级. 不失一般性, 设 $d_1 = \lambda + d\varepsilon^2$, 有

$$\det(A^{(k)} - \lambda I) = 0, \quad (25)$$

$$\det \begin{pmatrix} d & \alpha_{12} & \alpha_{13} & \alpha_{14} & \cdots & \alpha_{1n} \\ \alpha_{21} & d_2 - \lambda & \alpha_{23}\varepsilon & \alpha_{24}\varepsilon & \cdots & \alpha_{2n}\varepsilon \\ \alpha_{31} & \alpha_{32}\varepsilon & d_3 - \lambda & \alpha_{34}\varepsilon & \cdots & \alpha_{3n}\varepsilon \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \alpha_{n1} & \alpha_{n2}\varepsilon & \alpha_{n3}\varepsilon & \cdots & \cdots & d_{n-\lambda} \end{pmatrix} = 0. \quad (26)$$

用高斯消去法消去对角线以上的元素, 可得

$$d = c + O(\varepsilon), \quad (27)$$

c 是一个与 d_i, λ 和 α_{ij} ($i = 1, 2, \dots, n, j = 1, 2, \dots, n$) 有关的相对常数. 因此, 计算的特征值 d_1 的误差是 $O(\varepsilon^2)$ 精度的.

4.2. 特征向量的正交性. 由 (22) 式, 我们有

$$D = Q(A + G)Q^H, \quad (28)$$

其中 $G = Q^{-1}EQ^{-H}$, $Q = Q_k Q_{k-1} \cdots Q_1$. 我们认为 Q 及 D 的对角元是 A 的小扰动矩阵 $A + G$ 的特征向量与特征值, 这里只需证明 Q 的正交性. 假设计算的酉变换为

$$Q_i = Q'_i + F_i, \quad (29)$$

其中 $Q_i^H Q_i' = I$, F_i 是计算误差. 在 Wilkinson 的 [8] 中, 已证明了下述结果:

$$\|Q_k Q_{k-1} \cdots Q_1 - Q_k' Q_{k-1}' \cdots Q_1'\| \leq \sum_{i=1}^k \|F_i\|. \quad (30)$$

实际计算的 F_i 可以达到计算机精度, 例如

$$\|F_i\| \leq O(2^{-t}), \quad (31)$$

其中 t 是计算机尾数字长, 因此用新算法计算的特征向量矩阵的酉性质为

$$Q^H Q = I + O(2^{-t}). \quad (32)$$

只要计算的误差控制参数 δ 给得合理, 计算的 Q 的酉性质是相当满意的.

§ 5. 并行设计与分析

根据国家智能计算机研究开发中心研制的大规模并行计算机曙光 1000 的处理器 2D-mesh 结构, 我们作了算法并行设计与分析, 假定用 $p = k \times k$ 个处理器 P_{ij} . 划分矩阵 A 为 $k \times k$ 块矩阵 $A = (A_{ij})$, $i, j = 1, 2, \dots, k$, 把块矩阵 A_{ij} 对应地存贮于 P_{ij} 的局部存贮器中. 每个块矩阵又标记为 2×2 小块矩阵, 即 A 被划分为 $s \times s$ 小块矩阵 ($2k = s$). 用 P_{ii} 作特征向量的累积和解 A_{ii} 的特征值问题, 使得 A_{ii} 中的非对角小块被化为零. 用 $P_{ij} (i \neq j)$ 执行块矩阵的变换与块矩阵变换, 把要消去迭代的小块矩阵送到 P_{ii} 中. 实现一次扫描把所有 A 的非对角小块矩阵均被化零一次. 以 $P = 4 \times 4$ 为例, 块消去迭代的一次并行变换可表为

$$A^{(1)} = V^H A V, \quad A^{(1)} = (A_{ij}^{(1)}), \quad i, j = 1, 2, 3, 4, \quad (33)$$

$$A_{ii}^{(1)} = \text{diag}, \quad (34)$$

$$A_{ij}^{(1)} = V_i^H A_{ij} V_j, \quad A_{ji}^{(1)} = A_{ij}^{(1)H}, \quad (35)$$

$$V^{(1)} := V V^{(1)}. \quad (36)$$

由上式可知, 一次并行变换的工作量包括: 解子问题 $-5(n/k)^3$, 块矩阵变换 $-2(n/k)^3$, 特征向量累积 $-n(n/k)^2$. 一次扫描由 $(2k-1)$ 次并行变换组成, 可知一次扫描的并行计算步为 $(2/k + 13/k^2 - 7/k^3)n^3$ (FLOP). 通常用 4 次扫描可以使计算解达到很高精度, 即并行计算量为

$$(\delta/k + 52/k^2 - 28/k^3)n^3. \quad (37)$$

实际计算表明, 对于 $n > 500$ 的情况, 通信量相对较小, 我们可以得出计算量和加速比如下:

$P = k \times k$ 处理机个数	串行分块 并行步数 (n^3)	串行分块 并行步 (n^3)	串行块 / QR	串行块 / 并行块	QR / 并行块
2×2	48	13.5	9.6	3.5	0.3
3×3	48.9	7.5	9.8	6.5	0.6
4×4	53	4.5	10.6	11.7	1.1
5×5	58.6	3	11.7	19.5	1.6

由 (37) 可看出, 随着 k 的增加, 工作量相对增量减少, 因此并行分块消去迭代法是很好的. 上表显示, 并行分块消去迭代与串行算法相比, 加速比超过了 \sqrt{p} ; 与 QR 算法相比, 当 $p > 4 \times 4$ 时, 并行分块消去迭代法的计算机时小于 QR 算法. 此外, 加速比 \sqrt{p} 不是算法的本质限制, 所以设计较复杂的并行算法提高加速比是完全可能的.

§ 6. 数值结果

我们在国家智能计算机研究开发中心研制的大规模分布式并行计算机曙光 1000 上, 针对并行分块消去迭代法, 利用一些实验矩阵和计算物理应用问题矩阵的特征值问题为背景进行了计算, 计算结果表明, 迭代过程开始收敛慢, 后期收敛很快. 如果是普通的矩阵特征值问题, 通常需要扫描 3 至 4 次才能得到预期的收敛结果 (特征值精确至 10^{-3} , 残差 $\text{off}(A)$ 精确至 10^{-4}). 主要优点是程序容易编写, 可以用已有的解特征值问题的串程序在短期内完成计算工作, 有实用价值. 这里列出一个典型的试验矩阵计算例子:

$$A = (a_{ij}), \quad a_{ii} = i + 0.5,$$

$$a_{ij} = 0.5 + 0.02\sqrt{-1}, \quad i > j, \quad a_{ij} = 0.5 - 0.02\sqrt{-1}, \quad i < j.$$

图 1—3 表示, 对于 $n = 1024$, PE= 4, 16, 25 的收敛过程.

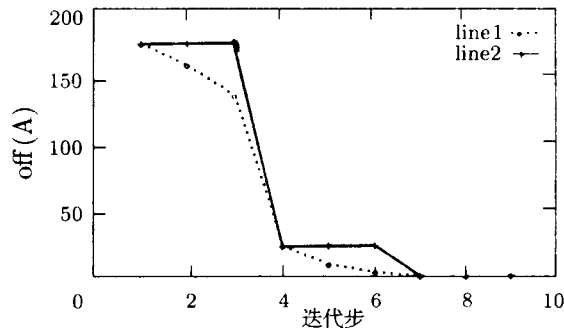


图 1. Off(A) 的变化 (4 个处理器情形)

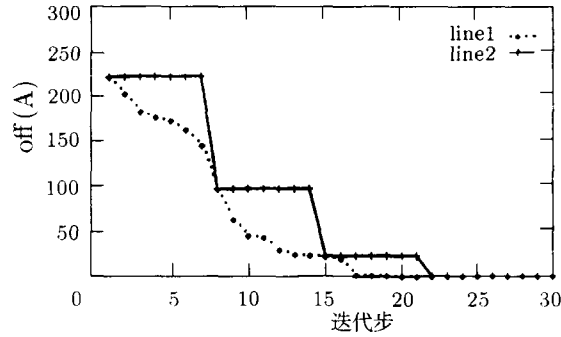


图 2. Off(A) 的变化 (16 个处理器情形)

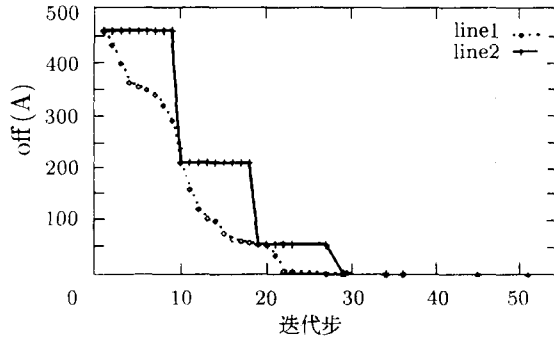


图 3. Off(A) 的变化 (25 个处理器情形)

PE=16

n	黑盒子	矩阵乘	通信	扫描次数	机时
256	62	268	498	4	848
512	518	688	522	4	1753
1024	4338	4110	544	4	9023

1024 阶迭代一步计算机时

PE	QR 子问题	矩阵乘	通信	迭代步数	扫描次数	总机时
16	156	145	19.7	28	4	9023
25	81.6	82.9	22.5	37	4	6770

§ 7. 讨论与结论

(1) 一般来说, QR 算法是解矩阵特征值问题的最有效算法, 但是 QR 算法不能大并行度实现, 因此, 我们提出了适用于分布式大规模并行机的并行分块消去迭代算法. 对于处理器个数大于 16 的情况, 计算速度将快于单处理机的串行 QR 算法. 新算法的另一个特点是编程十分简单、通信量小, 可以利用已有的任何最有效的串

行特征值问题解算器和分布处理器及其局部存贮器, 充分发挥高性能分布式大规模并行机的潜力, 解决大规模问题.

(2) 我们对新算法已给出了总体收敛性和渐近二次收敛性的证明, 也给出了算法的误差分析及并行设计和分析与数值实验结果一致.

(3) 我们用新算法已计算了几个计算物理中的高阶矩阵特征值问题和一些数值实验, 表明了理论分析的正确性并有实际的使用价值.

最后, 我们衷心感谢王鼎盛研究员、张文清博士和国家智能计算机研究开发中心对本文工作的帮助.

参 考 文 献

- [1] Deng Jianxin, On parallel farst Jacobi algorithm for the eigenproblem of real symmetric matrices, *Journal of Computational Mathematics*, **7** (1989), 412-417.
- [2] G. Golub, *Matrix Computations* (中译本, 大连理工大学出版社), 1988.
- [3] W.F. Mascarenhas, On the convergence of the Jacobi method for arbitrary orderings, *SIAM J. Matrix Anal. Appl.*, **16** (1995), 1197-1209.
- [4] D.N. Parlett, *The Symmetric Eigenvalue Problem*, Prontice-Hall Inc, 1980.
- [5] G.M. Shroff, A parallel algorithm for the eigenvalues and eigenvectors of a general complex matrix, *Numerishe Mathematik*, **58** (1991), 779-805.
- [6] G.M. Shroff, R. Schreiber, On the convergence of the cyclic Jacobi method for parallel block orderings, *SIAM J. Matrix Anal. Appl.*, **10** (1989), 326-346.
- [7] 孙继广, 矩阵扰动分析, 科学出版社, 1987.
- [8] J.H. Wilkinson, *The Algebraic Eigenvalue Problem* (中译本, 科学出版社), 1987.