国 防 科 技 大 学 学 报
JOURNAL OF NATIONAL UNIVERSITY OF DEFENSE TECHNOLOGY
第 19 卷第 4 期　　1997 年 8 月　　Vol. 19 No. 4

高阶稠密矩阵的并行求逆与应用*

陈　军　王正华　李晓梅

(并行与分布处理国家重点实验室　长沙　410073)

摘　要　采用基于列交换的 Gauss- Jordan 并行算法来解决空气动力学中超音速高阶面元法的稠密矩阵求逆问题, 该方法采取了块循环数据分配方式, 尤其对超立方体结构的并行机系统来说具有通讯优势。在 4 台 SGI 工作站构成的 2×2 网格上进行的实验表明, 对秩为 1000 左右的矩阵可得到 57% ～64% 的效率。

关键词　超音速流, 高阶面元法, 矩阵求逆, Gauss- Jordan 算法, SGI 工作站

分类号　V211. 3, TP312

Parallel High-order Dense Matrix Inversion and Applications

Chen Jun　Wang Zhenghua　Li Xiaomei

(P&DP National Laboratory, Changsha, 410073)

Abstract　In this paper we propose Gauss-Jordan algorithm using column interchanges for computing a high-order dense matrix inverse produced by the high-order panel method in aerodynamics. The algorithm implements the block cyclic data distribution, which especially has communication advantages for the hypercube network. Results are achieved when running this program on the 4 SGI workstations configured as a 2×2 processor grid. It shows that 57% ～64% efficiency can be gained when solving a matrix is about 1000 order.

Key words　supersonic flow, high-order panel method, matrix inversion, Gauss-Jordan algorithm, SGI

1　应用背景描述

在定常、无旋和小扰动假设下, 无粘流动的 Euler 方程退化为 Prandtl-Glauert 方程:

* 国家自然科学基金和 863 计划基金资助项目
1996 年 11 月 14 日收稿

$$(1 - M_{\infty}^2)\Psi_{xx} + \Psi_{yy} + \Psi_{zz} = 0 \tag{1}$$

其中 M_∞ 为来流马赫数，Ψ 为速度位函数，速度 $\boldsymbol{v}$ 与速度位函数之间存在如下关系：

$$\boldsymbol{v} = \nabla \Psi \tag{2}$$

令：

$$S = \mathrm{sign}(1 - M_{\infty}^2) \tag{3}$$

将坐标作适当的缩放

$$\bar{x} = x, \bar{y} = \beta y, \bar{z} = \beta z \tag{4}$$

则方程 (1) 可以写成

$$S\Psi_{\bar{x}\bar{x}} + \Psi_{\bar{y}\bar{y}} + \Psi_{\bar{z}\bar{z}} = 0 \tag{5}$$

当 $M_\infty = 0$ 或 $M_\infty = \sqrt{2}$ 时，(5) 和 (1) 同一。当 $M_\infty < 1$ 时，流动为亚音速流动，此时 $S = 1$，(5) 为拉普拉斯方程；当 $M_\infty > 1$ 时，流动为超音速流动，此时 $S = -1$，(5) 为波动方程，应用格林定理，流动中任意一点的速度位函数可以表示成沿物面的积分

$$\Psi(p) = \frac{1}{k}\iint_S \left[\frac{\sigma}{R} - \mu \bar{n} \cdot \nabla \left(\frac{1}{R}\right)\right] \mathrm{d}S \tag{6}$$

其中 σ 为源密度，μ 是偶极子密度，$\boldsymbol{n}$ 是物面的单位外法线矢量，R 是 P 点到曲面 S 上的积分动点的距离，K 是常数，超音速时取 -2π，亚音速时取 -4π。

将复杂飞行器的表面分成若干个网格，每个网络由 $M \times N$ 个网格点组成，形成 $(M-1) \times (N-1)$ 个面元，将网格中心及各边中点互连构成 8 个三角形子面元，然后布置线性源和二次偶极子，考虑边界条件和偶极子匹配条件，(6) 式的计算可以转化为一个高阶稠密线性方程组的求解，即下式 $\boldsymbol{AIC}$ 的求逆问题[1]。

$$\boldsymbol{AIC}\lambda = \boldsymbol{b} \tag{7}$$

当边界条件变化时，$\boldsymbol{b}$ 改变但 $\boldsymbol{AIC}$ 不变，因而一次求解可以带多个右端项，求解多组状态；当矩阵 $\boldsymbol{AIC}$ 的阶数达到几千时，在单机上求解，一是容易超内存，二是求逆的时间太长，满足不了工程实践的需要。因此必须寻求该问题的并行算法。基于列交换的 Gauss-Jordan 并行求逆算法能满足此需要，该方法采取的块循环数据分配方式、远程通讯及负载平衡性好，对超立方体结构的并行机尤其适用。

2　基于列交换的 Gauss-Jordan 矩阵求逆的并行算法

文献 [3] 中分析了基于列交换的 Gauss-Jordan 算法的稳定性，认为它具有很好的向量特性及负载平衡性。以下为了方便起见，设矩阵 $\boldsymbol{A}$ 的阶数为 N，有一个 P 台处理机相连的多机系统，并设 P_i 为第 i 台处理机。

2.1　数据分配

在高阶稠密矩阵并行求逆过程中，由于数据相关性强，将哪些数据分到哪台处理机上影响着通讯量、负载平衡性以及算法的性能和可扩放性。算法的实现与结构紧密相关，因此好的数据分配方案，与具体应用的体系结构有关。该算法是面向超立方体结构的并行机系统的，采用了块循环数据分配方式[2,4]，即将 $N \times N$ 矩阵分成大小为 $nr1 \times nr2$ 的块，行列方向固定距离的元素被分给同一个处理机，行列方向的步距分别为 $r1$ 和 $r2$，其中 $r1 \times r2 = P$，此组成了 $r1 \times r2$ 处理网格。$nr1$、$nr2$ 和 $r1$、$r2$ 的关系分别为：

$$nr1 = N / r1, nr2 = N / r2$$

将矩阵 A 中行号为 $(k-1)/r2+1+r1\times(i-1)$　　　$i=1,\cdots,nr1$

列号为 $\mathrm{mod}(k-1,r2)+1+r2\times(j-1)$　　　$j=1,\cdots,nr2$

的元素分给处理机 P_k，作为 P_k 的元素 $mma(i,j)$，其中 $k=1,\cdots,P$

2.2　算法

设 $rp1$ 是 2 的幂中不小于 $r1$ 的最小值，$rp2$ 是 2 的幂中不小于 $r2$ 的最小值。$d1=\log_2(rp1)$，$d2=\log_2(rp2)$。处理机网格中与处理机 P_{me} 处于同一行或同一列的处理机分别记录在 rowp 和 colp 中

```
for k= 1 to N do
    if Pme 上有第 K 行元素 then
        这些元素构成局部主元行 pr (i), i= 1, …, nr1; 将 pr 广播给 p∈rowp.
    else 接收消息到 pr (i), i= 1, …, nr1
    end if
        取 pr 中绝对值最大的元素为 pe, 相应列号为 col, 且记录 Pme 上有关第 col
        列元素段到 pc (i), i= 1, …, nr2。
    if Pme 上有第 K 列元素 and col≠K then
        交换这两列元素
    end if                    (可能有交换)
         在 colp 中的 nr2 个处理机上比较 pe. 最终, 每个处理机得到全局主元 pe
         和相应 pc. 用 pc、pr 和 pe 修改本地数据 mma (i, j), i= 1, …, nr1, j
         = 1, …, nr2
end for
```

经过上述算法，得到的 mma 元素值即是原矩阵 $\boldsymbol{A}$ 经扰动后的逆矩阵的相应值。

设一个整型数的消息长度为 INT，一个浮点数的消息长度为 FLOAT。该算法的第 k 次循环要做:

(1) 在处理机网格的列中广播具有 $\boldsymbol{A}$ 的第 k 行元素的主元消息 $pr(j)$，$j=1,\cdots,nr2$. 然后每个处理机自行决定局部主元 $pe1$、该主元所在的列 $pc(i)$，$i=1,\cdots,nr1$，以及相应的列号 col. 这里 $r2$ 个处理机具有主元行消息，分别向 $r1-1$ 个处理机广播该消息，消息长度为 $nr2\times$FLOAT. 这 $r2$ 个处理机是并行广播的。

(2) 在处理机网格的同一行中的 $r2$ 个处理机比较各自的 pel，得到最大值即全局主元 pe、相应的主元列 $pc(i)$，$i=1,\cdots,nr1$. 以及列号 col. 每个处理机 P_{me} 依次与 P_j 相互比较，其中 $(j-1)$ 与 $(me-1)$ 的二进制中某一位互为相反，这样的 j 至多有 $d2$ 个，即每个处理机至多要与 $d2$ 个处理进行点对点通讯，消息长度为 $1\times$INT $+(1+nr1)\times$FLOAT.

(3) 用 pe、pr 和 pc 修改本地数据 $mma(i,j)$，$i=1,\cdots,nr1$. $j=1,\cdots,nr2$. 该步不与其它处理机通讯，计算量主要集中在此，为 $o(nr1\times nr2)$.

3　数值实验

实验是在 4 台 SGI 工作站组成的多机系统上完成的。每台 SGI 算作一个结点，其主

频为100MHz,具有32MB 内存的容量。4 个结点构成 2×2 处理机网格。该系统运行PVM 形成虚拟机, 通讯带宽为10Mbps。所有的时间结果均是在该系统上运行一个对随机产生的矩阵求逆的 Fortran 程序而得到的。首先给出衡量算法性能的标准。

3.1 相对加速度

$$Sp = T_{ser}/T_{par}$$

其中 T_{ser} 是并行程序在1台处理机上求解 I 的时间, T_{par} 是并行程序在 P 台处理机上求解 I 的时间。其中 I 是问题实例, P 是处理机台数。

3.2 相对效率 $\boldsymbol{E = S_p/P}$.

所得到的结果见附表:

附表

N	T_{ser} (s)	T_{par} (s)	S_p	E
500	81.9	52.9	1.55	38.7%
800	281.9	123.0	2.28	57%
1000	523.7	204.5	2.56	64%

表中 $N=500$ 时, S_p 和 E 不理想。原因在于 N 不甚大时, 通讯量与计算量相比所占比重较大, 加速比不高。但从表中可知, 当阶 N 增大到 1000 左右时, 加速比与效率提高了。

由于处理机资源有限, 未能对 $r1$、$r2$ 进一步增大的情况作一番测试。但文献 [2] 中的实验结果表明, 对维数为 1000 的矩阵, 当 P 增大到 16 台, $r1$、$r2$ 均是 4 时, 效率可达到 97%.

4 结论

本文利用了面向超立方体的基于列交换的 Guass-Jordan 矩阵求逆的并行思想, 采取块循环数据分配方式, 改进了文献 [2] 的算法; 采用广播方式代替循环点对点方式, 消去了处理机网格 $r1 \times r2$ 中对 $r1$、$r2$ 的限制, 解决了空气动力学中超音速高阶面元法的稠密矩阵求逆问题。但计算效率还不很高, 有望于通过计算与通讯重叠来提高该算法的性能; 同时对于不同的体系结构, 如mesh、ring 等, 寻求适合的数据分配方式和相应的机间通讯方式, 研究与之对应的高效的并行算法。

参考文献

1 Medan RT, Magnus AE PAN AIR- A Computer Program for Predicting Subsonic or Super-sonic Linear Potential Flows About Arbitrary Configurations Using a High Order Panel Methd, Vol. 3- Case Manual. NASA- 3253, 1981

2 Eleanor Chu, Alan George and Darcy Quesnel. Parallel matrix inversion on a subcube- grid. Parallel Comput. 1993, 19: 243

3 Johnsson SL. Communication efficient basic liner algebra computations on hypercube architectures. J. ParallelDis-trib. Comput. 1987, 4: 133

4 孙家昶等. 网格并行计算与分布式编程环境. 北京: 科学出版社, 1996

(责任编辑　石少平)