

利用并行方法解 $AX + XB = C$ 型线性矩阵方程

慕德俊 戴冠中

(西北工业大学自动控制系 西安 710072)

摘 要

提出了一种新的递推算法用于求解 $AX + XB = C$ 型线性矩阵方程, 这种算法可以用脉动阵列结构并行实现, 该算法和结构还可求解其它几种类似的线性矩阵方程, 特殊情况下求解方程的阵列结构可进一步简化. 仿真结果表明, 这种并行方法有较高的加速比及效率.

关键词: 矩阵方程, 脉动阵列, 并行算法.

1 引言

线性非齐次矩阵方程 $AX + XB = C$ 及其类似的线性方程在控制理论中起着很重要的作用, 对此类方程的求解, 有多种串行方法^[1-2]. 本文提出了一种新的递推方法, 用于并行求解方程. 该方法首先用并行的 QR 分解方法^[3]对实矩阵 A, B 进行正交相似变换, 将它们变为上 Hessenberg 阵, 然后阵列结构并行求解 X . 对其它类似的矩阵方程 $AXM + X = C$ 和 $AXM + LXB = C$ 的并行计算也进行了简要分析. 这些并行算法都可通过脉动 (Systolic) 阵列结构实现.

2 $AX + XB = C$ 的递推算法

设线性非齐次矩阵方程

$$AX + XB = C, \quad (1)$$

式中 A 和 B 分别是 $R^{m \times m}$ 和 $R^{n \times n}$ 的方阵, C 是 $R^{m \times n}$ 的矩阵, X 是 $R^{m \times n}$ 未知矩阵, 不失一般性设 $m \geq n$.

由实 Schur 分解定理可知, 存在正交矩阵 Q_1, Q_2 使得 A, B 收敛于上 Hessenberg 阵 H 和 R , 即 $H = (h_{ij}) = Q_1^T A Q_1, R = (r_{ij}) = Q_2^T B Q_2$, 且 H 和 R 的主对角为 1×1 或 2×2 的子矩阵. 通过三角阵列结构并行 QR 分解方法是化矩阵为上 Hessenberg 形式的一种有效的方法^[3], 最终使(1)式变为

$$HY + YR = F, \quad (2)$$

式中 $Y = Q_1^T X Q_2, R = Q_2^T B Q_2, F = Q_1^T C Q_2, H = Q_1^T A Q_1$.

Gloub^[1] 用 Hessenberg-Schur 方法串行求解(1)式. 若将(2)式分解成递推形式, 可用并行方法求解, 分解方法如下:

若 $r_{21} = 0$, (2)式可变为

$$H[\mathbf{y}_1 \ Y_2] + [\mathbf{y}_1 \ Y_2] \begin{bmatrix} r_{11} & \mathbf{R}_{12} \\ 0 & \mathbf{R}_{22} \end{bmatrix} = [\mathbf{f}_1 \ F_2]. \quad (3)$$

式中为 \mathbf{y}_1 为 Y 的第一列向量, Y_2 为除第一列外的 $m \times (n-1)$ 阶矩阵, \mathbf{f}_1 为 F 阵的第一列向量, F_2 为 F 余下的 $m \times (n-1)$ 阶矩阵, \mathbf{R}_{12} 为矩阵 R 的第一行除 r_{11} 外的行向量, \mathbf{R}_{22} 为 R 余下的 $(n-1) \times (n-1)$ 阶子矩阵, (3)式可等价表达为

$$(H + r_{11}I)\mathbf{y}_1 = \mathbf{f}_1, \quad (4)$$

$$HY_2 + Y_2\mathbf{R}_{22} = F_2 - \mathbf{y}_1\mathbf{R}_{12}. \quad (5)$$

若 $r_{21} \neq 0$, (2)式表示为

$$H[\mathbf{y}_1 \ \mathbf{y}_2 \ Y_3] + [\mathbf{y}_1 \ \mathbf{y}_2 \ Y_3] \begin{bmatrix} r_{11} & r_{12} & \mathbf{R}_{13} \\ r_{21} & r_{22} & \mathbf{R}_{23} \\ 0 & 0 & \mathbf{R}_{33} \end{bmatrix} = [\mathbf{f}_1 \ \mathbf{f}_2 \ F_3], \quad (6)$$

式中各个量的定义与(3)式类似. 上式可等价于

$$\begin{bmatrix} r_{21}I & H + r_{11}I \\ H + r_{22}I & r_{12}I \end{bmatrix} \begin{bmatrix} \mathbf{y}_2 \\ \mathbf{y}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \end{bmatrix}, \quad (7)$$

$$HY_3 + Y_3\mathbf{R}_{33} = F_3 - \mathbf{y}_1\mathbf{R}_{13} - \mathbf{y}_2\mathbf{R}_{23}. \quad (8)$$

从(4)~(8)式可以看出, 通过递推计算可以将非齐次方程变为线性齐次矩阵方程, 然后逐次求解未知矩阵 Y , 递推算法概括如下:

$i = 1$

$R^{(1)} = R = (r_{ij}); F^{(1)} = F$

For $j = 1$ to $n - k$

If $r_{i+1,i} = 0$ then (computing column i of Y)

$R^{(i)} = \begin{bmatrix} r_{11} & \mathbf{R}_{12}^{(i)} \\ 0 & \mathbf{R}_{22}^{(i)} \end{bmatrix}; F^{(i)} = [\mathbf{f}_1^{(i)} \ F_2^{(i)}]$

$(H + r_{11}I)\mathbf{y}_i = \mathbf{f}_1^{(i)}$

If $i < n$ then (Update $F_2^{(i)}$)

$F_2^{(i)} = F_2^{(i)} - \mathbf{y}_i\mathbf{R}_{12}^{(i)}$

$R^{(i+1)} = \mathbf{R}_{22}^{(i)}; F^{(i+1)} = F_2^{(i)}$

$i = i + 1$ End If

Else (Computing column i and $i + 1$ of Y)

$R^{(i)} = \begin{bmatrix} r_{i,i} & r_{i,i+1} & \mathbf{R}_{i3}^{(i)} \\ r_{i+1,i} & r_{i+1,i+1} & \mathbf{R}_{23}^{(i)} \\ 0 & 0 & \mathbf{R}_{33}^{(i)} \end{bmatrix}; F^{(i)} = [\mathbf{f}_1^{(i)} \ \mathbf{f}_{i+1}^{(i)} \ F_3^{(i)}]$

$$\begin{bmatrix} r_{i+1,i}I & H + r_{i,j}I \\ H + r_{i+1,i+1}I & r_{i,i+1}I \end{bmatrix} \begin{bmatrix} y_{i+1} \\ y_i \end{bmatrix} = \begin{bmatrix} f_i^{(j)} \\ f_{i+1}^{(j)} \end{bmatrix}$$

If $i < n - 1$ Then (Update $F_3^{(j)}$)

$$F_3^{(j)} = F_3^{(j)} - y_i R_{13}^{(j)} - y_{i+1} R_{23}^{(j)}$$

$$R^{(j+1)} = R_{33}^{(j)}; F^{(j+1)} = F_3^{(j)}$$

$$i = i + 2$$

End If; End If; End For.

其中 $F^{(j+1)}$, $R^{(j+1)}$ 是计算 $y_i(y_{i+1})$ 后得到的修正矩阵, 它们的初始值分别为 F, R , k 是 R 阵次对角线非零元素的个数.

3 利用 Systolic 阵列结构并行计算

利用 Systolic 阵列结构并行求解非齐线性矩阵方程(1), 需 $(2m + 3)m$ 个处理单元组成的梯形阵列以及 $n - 1$ 个单元组成的线性阵列. 假设在 A, B 阵均为 3 阶的情况下求解方程(1)所使用的阵列结构见图 1, 为了叙述方便, 将整个结构分为 4 个部分, 并且图中的数据流是第一次递推计算时的情形.

当 $r_{21} = 0$ 时, 只需用(4), (5)两式计算 y_1 及修正 F_2 . 由于(4)式初始阵为稀疏阵, 为了减少计算步, 将 r_{11} 和 H, f_1 的每行元素, 按先下后上的顺序以时滞(skew)方式输入阵列 3, 实现对增广矩阵 $[(H + r_{11}I) f_1]$ 进行正交三角变换, 使(4)式变为 $Ry_1 = e_1$ 型, 式中, R 为上三角阵; 输入 $I_{m \times n}$ 阵和 0 与存在阵列中的上三角阵 R 和 e_1 进行高斯消去, 得到 $R^{-1}e_1$ 即 y_1 , 从阵列的底部输出. 进行(5)式的矩阵修正计算时, 将 y_1 矢量以及 R_{12} 和 F_2 输入阵列 4 进行 $F_2 - y_1 R_{12}$ 运算, 输出修正矩阵 F'_2 , 用来进行下一步递推计算.

当 $r_{21} \neq 0$ 时, 需计算(7), (8)两式, 求出 y_1, y_2 以及修正 F_3 , 由于(7)式的初始矩阵为稀疏矩阵, 为了缩短计算周期, 将 $H, r_{22}, r_{12}I$ 及 f_2 按行先下后上的顺序输入阵列 1, 2, 通过阵列 1, 2, 3 组成的梯形阵列进行正交三角变换. 这一步计算完成后, (7)式变为 $R_1[y_2^T y_1^T]^T = [e_1'^T e_2'^T]^T$, 式中, R_1 是 $2m \times 2m$ 的上三角阵; 输入 $I_{2m \times 2m}, 0$ 至阵列 1, 2, 与存在阵列 1, 2, 3 中的 R_1 阵及 e_1', e_2' 进行高斯消去计算, 得到 y_2, y_1 并从阵列 3 的下方输出. 为进行(8)式的 F_3 阵修正计算, 将 y_2, F_3 与 R_{23} 同进输入阵列 4 执行 $F'_3 = F_3 - y_2 R_{23}$ 的计算, 然后将输出的 F'_3 返回到阵列 4 的输入端再与 y_1 及 R_{13} 一起输至阵列 4, 最终得到 $F''_3 = F_3 - y_2 R_{23} - y_1 R_{13}$, 并作为下次迭代计算的修正矩阵.

在有效地组织数据流输入、输出的情况下, 通过时序分析可知, 求解 Y 的计算共需 $n(m + 3)$ 个计算步.

$n \times n$ 的阵列完成 $X = Q_1 Y Q_2^T$ 的计算. 通过适当的分时功能转换, 可应用阵列 1, 2 完成 Schur 分解及 Q_1, Q_2 的计算. 这样, 阵列 1—4 即可完成整个计算过程.

4 其它几种线性方程的并行计算

(1) 式中的一种特殊情况是当 A, B 的特征值都为实数时, 由 Schur 定理可知, (2)式

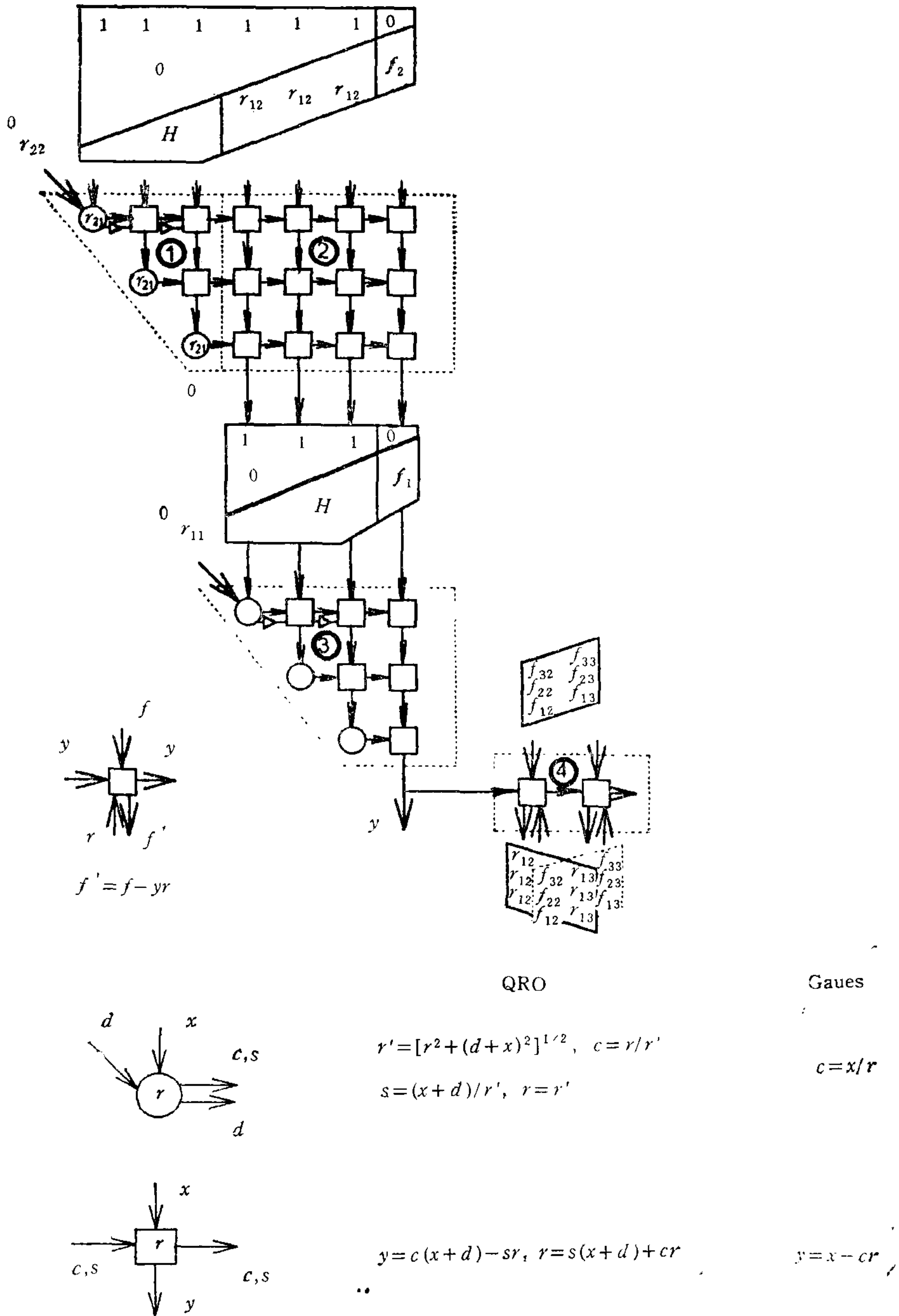


图 1 求解(4),(5),(7)和(8)式所需的阵列结构

中的 H, R 为上三角阵, 递推算法如下:

$$i = 1$$

$$R^{(1)} = R = (r_{ij}); F^{(1)} = F$$

For $j = 1$ to n

$$R^{(j)} = \begin{bmatrix} r_{ii} & R_{22}^{(j)} \\ 0 & R_{22}^{(j)} \end{bmatrix}; F^{(j)} = [f_i^{(j)} \quad F_2^{(j)}]$$

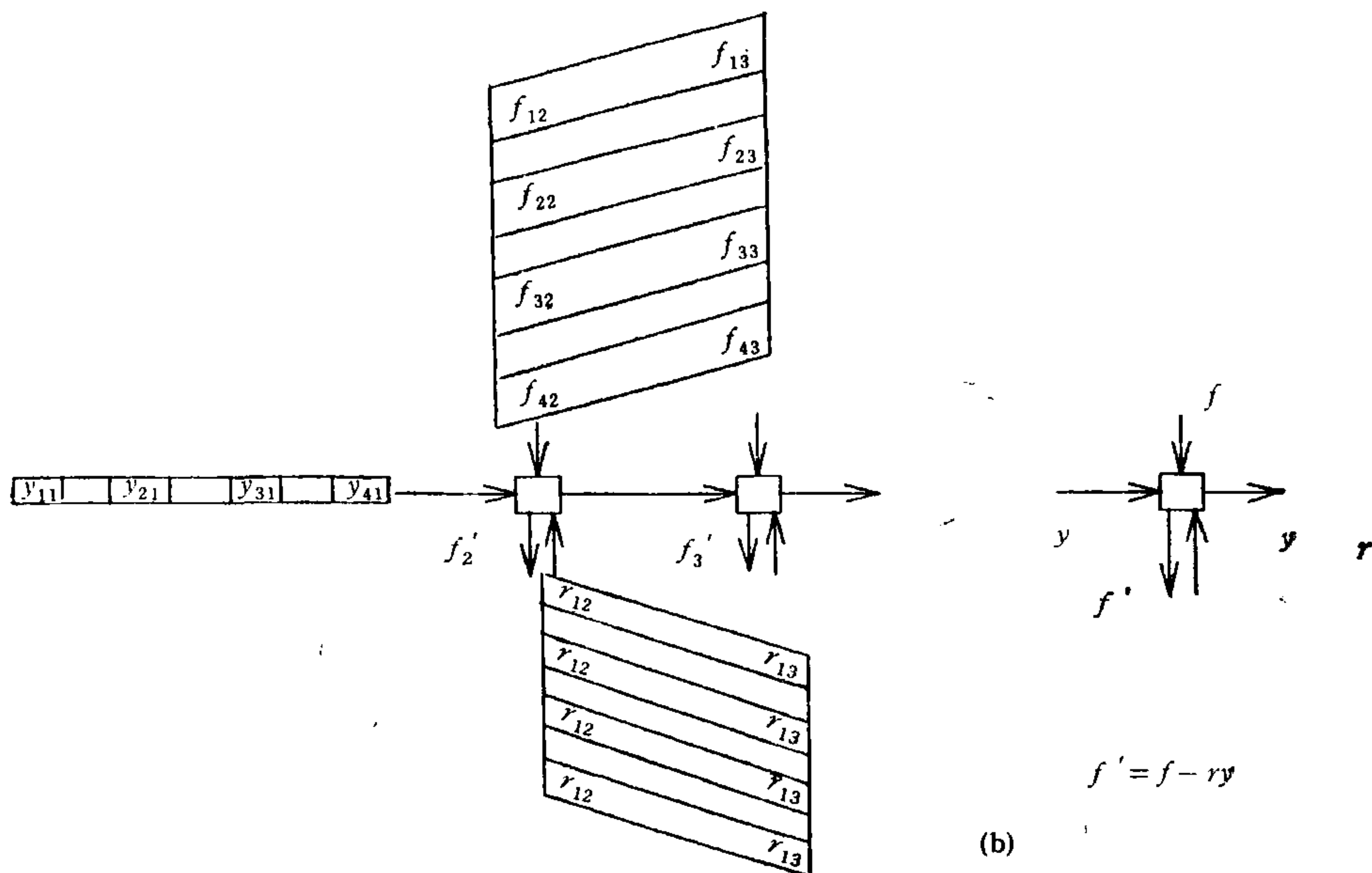
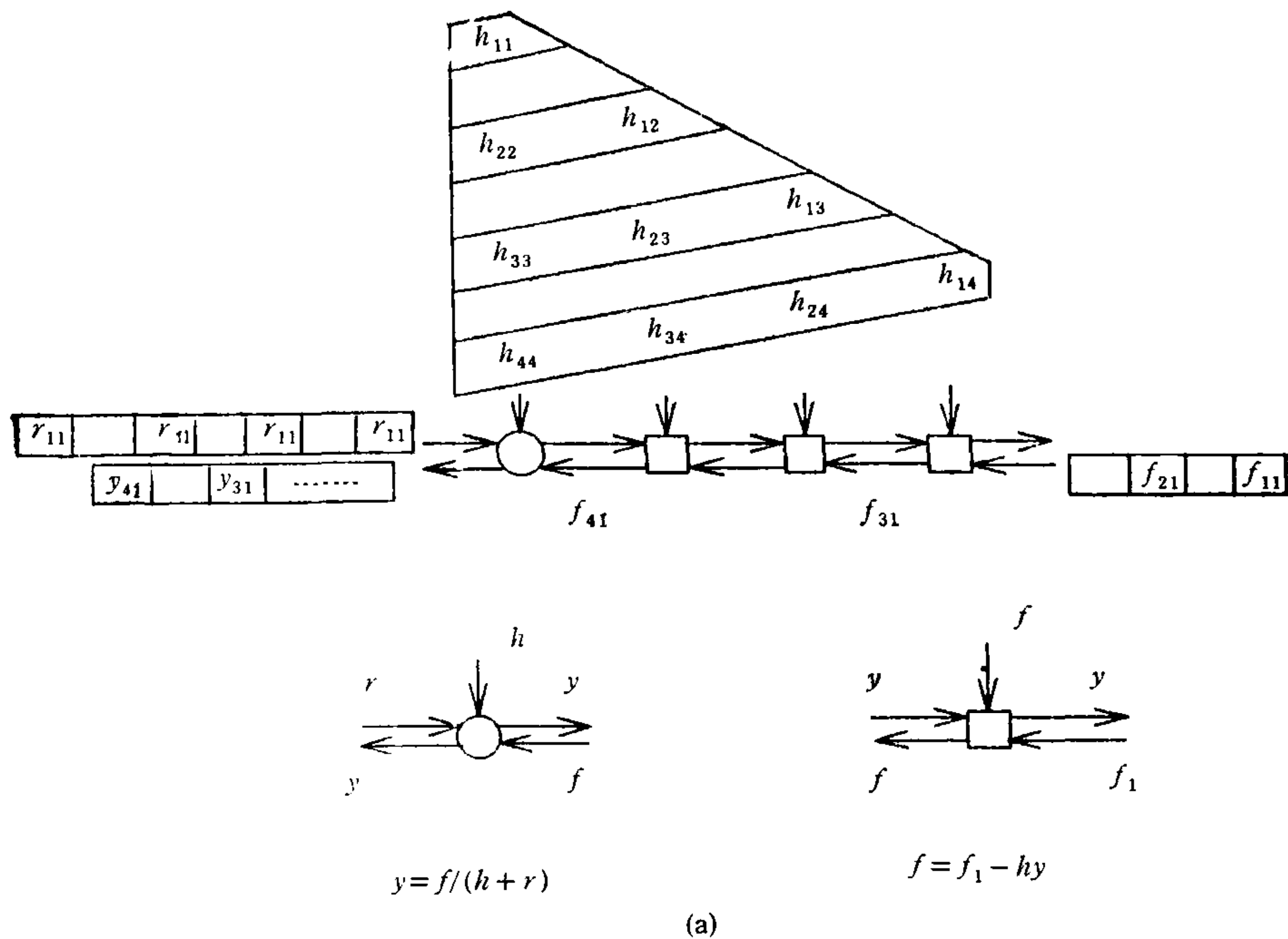


图 2 求解(2)式所需的阵列结构 (A, B 具有实特征值)

$$\begin{aligned} (H + r_{ii}I)y_i &= f_i^{(j)} \\ \text{If } i < n &\text{ then (Update } F_2^{(j)}) \\ F_2^{(j)} &= F_2^{(j)} - y_i R_{12}^{(i)} \\ R^{(j+1)} &= R_{22}^{(j)}; F^{(j+1)} = F_2^{(j)} \\ i &= i + 1 \text{ End If; End For} \end{aligned}$$

由于 $(H + r_{ii}I)$ 为上三角矩阵,应用回代方法能使并行计算的方形阵列^[4]简化为图 2 所示的线性阵列结构,图中设 $m = 4, n = 3$ 。第一次递推计算后, y_1 从图 2(a) 输出,修正后的 f_2, f_3 从图 2(b) 下方输出。通过有效地组织数据流的输入输出,完成整个计算需 $2mn$ 个计算步,由图 2 可以看出,这种结构使用的处理单元比较少,节省硬件开销。

若两个矩阵方程为

$$AXM + X = C, \quad (9)$$

$$AXM + LXB = C, \quad (10)$$

式中 $A, L \in R^{m \times m}; M, B \in R^{n \times n}; C \in R^{m \times n}$ 均为可逆阵,将(9), (10)式转换为 $AX + XB = C$ 的形式

$$AX + X(M^{-1}) = (CM^{-1}), \quad (11)$$

$$(L^{-1}A)X + X(BM^{-1}) = (L^{-1}CM^{-1}). \quad (12)$$

通过上述的算法及阵列结构可并行计算 $M^{-1}, CM^{-1}, L^{-1}A, BM^{-1}, L^{-1}CM^{-1}$, 及方程(11), (12)。

5 仿真结果

在并行计算中,通常用加速比 S 和效率 E 来衡量一个并行算法的有效性。加速比用于度量算法并行性对时间的改进程度;效率用于度量处理能力发挥的程度。它们的定义如下:

$$S = T_1/T_p, \quad E = S/P,$$

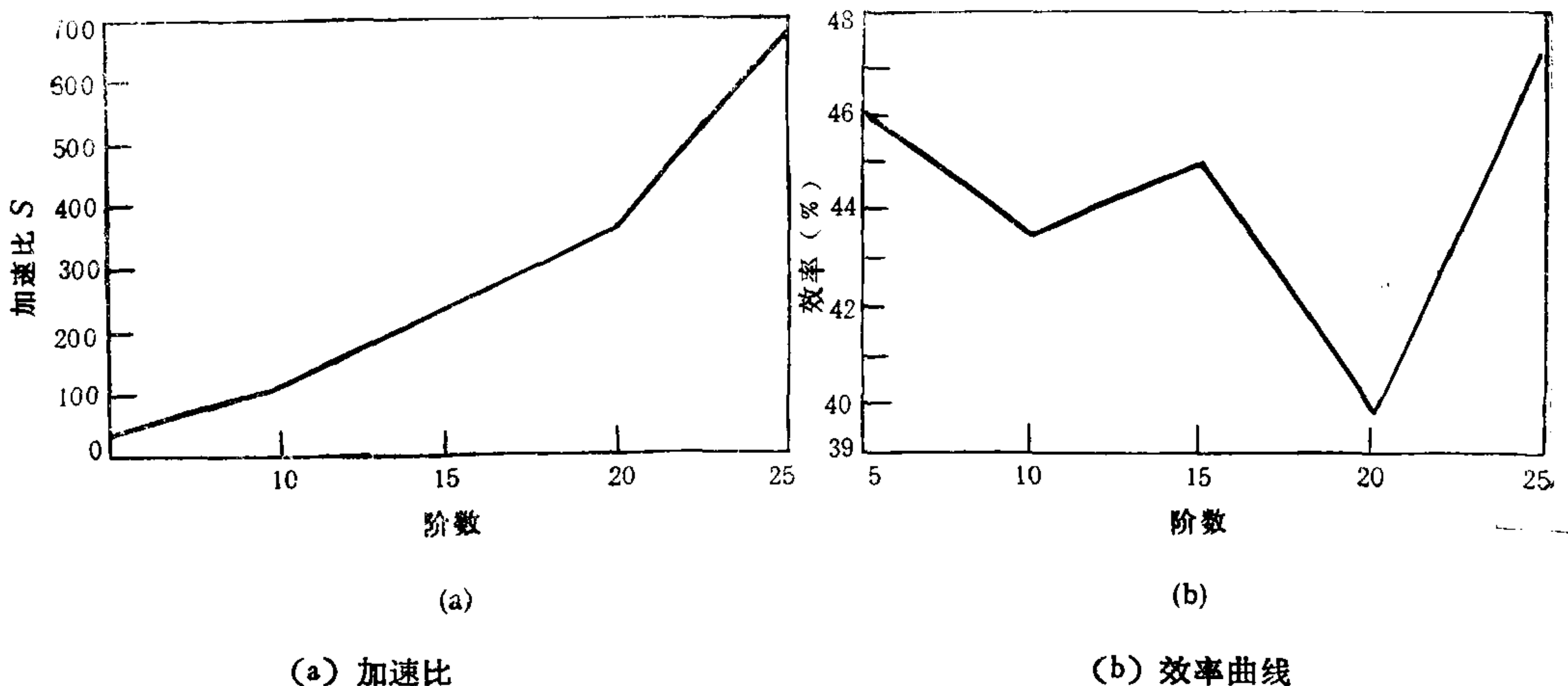


图 3

式中 T_1 为最快的串行算法在单处理机上的运行时间, T_p 为并行算法在 p 个处理单元上的运行时间。

在串行计算中, 矩阵 B 收敛于上 Hessenberg 阵的准则是: $\text{Off}(B_k)$ 小于一个给定值, 这里 $\text{Off}(B_k) = \sum_{i < j} b_{ij}^2(k)/N$ 。但在并行计算中 $\text{Off}(B_k)$ 很难检测到, 按照文献 [4], 可以认为在阶数小于 30 的情况下, 并行递推 8 次后矩阵 B 即可收敛于上 Hessenberg 阵。

假设 $m = n, k = [n/2]$, 通过用 MATLAB 得到上述不同阶系统的串、并行算法的浮点运算次数, 求出的加速比及效率曲线见图 3(a), 3(b)。

从图中可以看出, 该算法具有良好的并行性。随着问题规模的增大, 加速比也随着增加, 但这种增加是以硬件开销为代价。

6 结语

本文系统地分析了并行求解 $AX + XB = C$ 的整个过程, 文中提出的递推方法可有效地映射到 Systolic 阵列上并行计算, 加快了计算速度。对上述方程的一种特殊形式, 文中给出了一种简化的阵列结构, 对 $AXM + LXB = C$ 及 $AXM + X = C$ 型方程也作了简要分析。通过适当的预处理, 可使用求解 $AX + XB = C$ 的递推方法及阵列结构来实现, 仿真结果表明, 这种并行方法有较高的加速比及效率。因此, 本文所提出的算法和阵列结构是一种有效地并行求解非齐次线性矩阵方程的方法。

参 考 文 献

- [1] Golub G H. A Hessenberg-Schur method for the problem $AX + XB = C$. *IEEE Trans.*, 1979, **AC-24**(6):909—913.
- [2] 韩京清, 何关钰, 许可康. 线性系统理论代数基础. 辽宁科学出版社, 1987.
- [3] Kuojuvey R L and kung Y. Multiphase systolic algorithms for spectral decomposition. *IEEE Trans.*, 1992, **SP-40**(1):190—20.
- [4] Hueso J L. A systolic algorithm for the triangular stein equation. *J. VLSI Signal Prodesing*, 1993, **5**(1):49—55.

A SYSTOLIC ALGORITHM FOR MATRIX EQUATION $AX + XB = C$

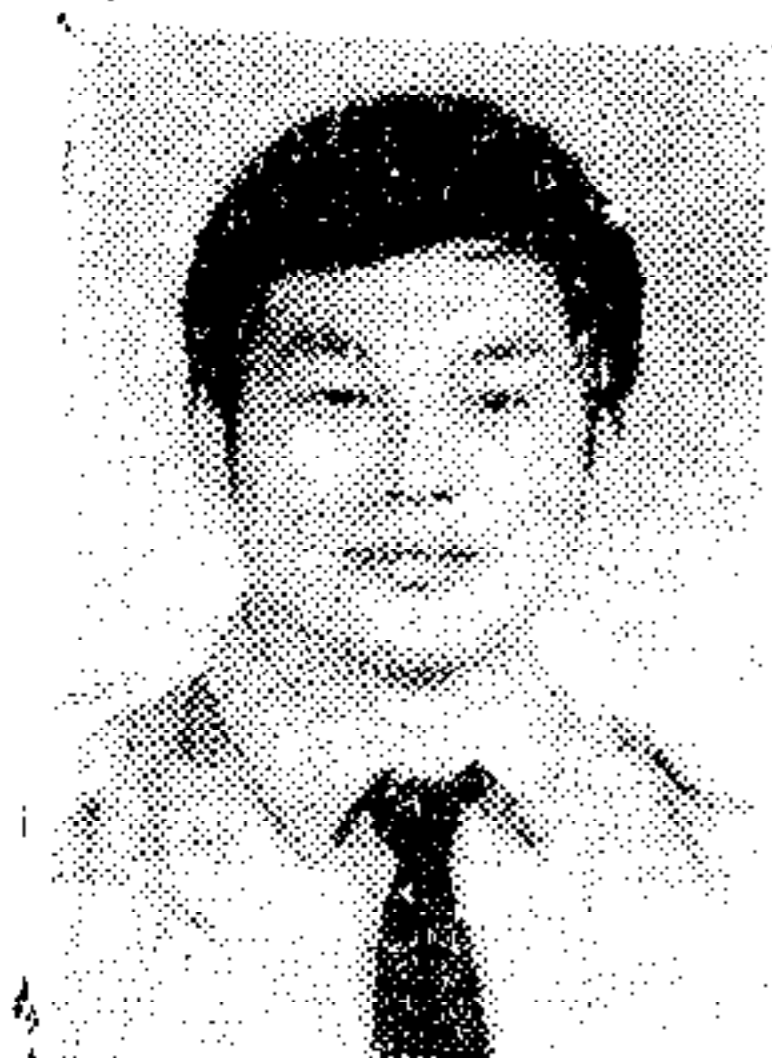
MU DEJUN DAI GUANZHONG

(Dept. of Automatic Control, Northwestern Polytechnical University Xi'an 710072)

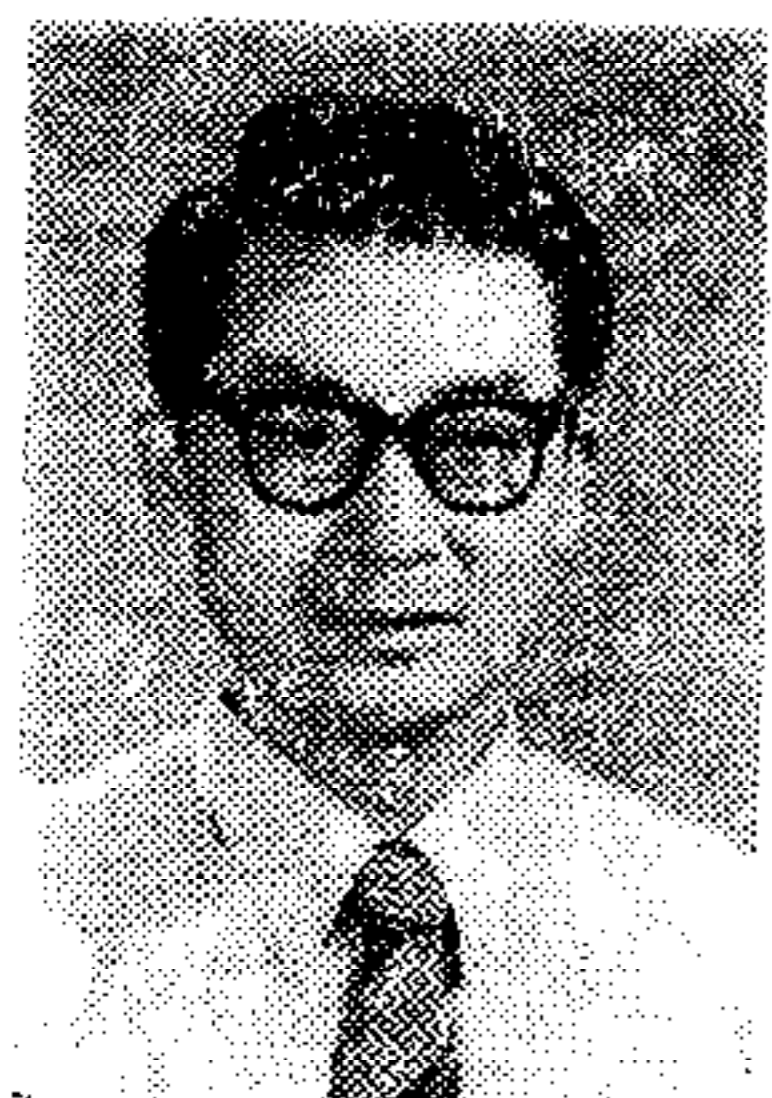
ABSTRACT

In this paper, an iterative algorithm is proposed to solve the matrix equation $AX + XB = C$ based on QR decomposition and Gaussian elimination algorithm. The recursive algorithm can be implemented by a multiphase systolic architecture readily. Details on these multiphase operations of the algorithm are discussed in this paper. Similar matrix equations $AXM + LXB = C$ and $AXM + X = C$ are analyzed simply. Simulation results show that the algorithms have a high speed up.

Key words: Matrix equation, systolic array, systolic algorithm.



慕德俊 1963年生。1983年获军械工程学院导弹专业学士学位,此后在新疆某部从事导弹业务。1990和1994年分别获西北工业大学惯导专业硕士学位和自动控制理论及应用专业博士学位,现为南京航空航天大学自动控制研究所博士后,主要研究领域是控制系统中的并行处理算法、低空突防中的地形跟随及地形回避等。



戴冠中 1937年生。教授,自动控制理论及应用学科博士生导师,现任西北工业大学校长,目前的主要研究领域是大系统估计与控制理论、智能控制、控制系统中的并行处理理论、算法与并行仿真计算机、控制理论在石油和天然气勘探中的应用等等。