Numerical Solution of Block Tridiagonal Form

by Katsuhiko SANADA

1. Introduction

We study the numerical solution of the linear systems arising from certain implicit finite difference approximation to the linear differential equations. Especially we consider those schemes which lead to matrices of block tridiagonal form. When only one equation is involved, this scheme is tridiagonal form.

For example, the heat equation

$$u_t = u_{xx} \qquad 0 \leqq x \leqq 1, \qquad 0 \leqq t \leqq T$$
$$u(x, 0) = f(x), \qquad u(0, t) = g_1(t), \qquad u(1, t) = g_2(t)$$

by the Crank-Nicolson scheme

$$\left(I - \frac{k}{2} D_+ D_-\right) u_j^{m+1} = \left(I + \frac{k}{2} D_+ D_-\right) u_j^m$$

leads to a tridiagonal matrix equation, for which the factorization method is often used as direct solution. [see (1)].

$$u_j^m = u(j\Delta x,\ m\Delta t)$$
$$k = \Delta t = \text{time step}$$
$$h = \Delta t = \text{space step}$$

However, when we consider a system of differential equations, the structure of the matrix is more complicated.

For example, the system of the parabolic equations

$$\boldsymbol{u}_t = P(x, t) \boldsymbol{u}_{xx}, \qquad \boldsymbol{u} = (u_1, \ldots\ldots, u_p)$$
$$\left.\begin{aligned} Q_1 \boldsymbol{u}_x + Q_0 \boldsymbol{u} &= g_1(t) \\ Q_2 \boldsymbol{u} &= g_2(t) \end{aligned}\right\} \text{ at } \ x = 0$$
$$\left.\begin{aligned} R_1 \boldsymbol{u}_x + R_0 \boldsymbol{u} &= g_3(t) \\ R_2 \boldsymbol{u} &= g_4(t) \end{aligned}\right\} \text{ at } \ x = 1$$

the Crank-Nicolson scheme

$$\left(I - \frac{k}{2} D_+ P\left(x_j - \frac{h}{2},\ t_{m+1}\right) D_-\right) u_j^{m+1} = \left(I + \frac{k}{2} D_+ P\left(x_j - \frac{k}{2}, t_m\right) D_-\right) u_j^m$$

with any kind of discrete boundary conditions can be expressed in the block tridiagonal form. [5]

2. Block Tridiagonal Form

Block tridiagonal matrix is

$$A=\begin{pmatrix} A_1 & C_1 & & \\ B_2 & A_2 & \ddots & \\ & \ddots & \ddots & C_{n-1} \\ & & B_n & A_n \end{pmatrix} \tag{1}$$

where each of A_i represents a square matrix of order m_i, and each of the B_i and C_i are rectanguler matrices. That is, B_i must have m_i rows and m_{i-1} columns, and C_i must have m_i rows and m_{i+1} columns. And if all $m_i=m$, then all the sub-matrices are square and of order m.

Thus, let the system be

$$Ax=f \tag{2}$$

where $$x=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad f=\begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} \tag{2)'$$

and each x_i and f_i are m_i-component column vectors.

The system (2) may be solved by a procedure analogous to the factorization method of a tridiagonal form. [1]

Assume matrix (1) can be decomposed as

$$A=LU=\begin{pmatrix} P_1 & & & \\ B_2 & P_2 & & \\ & \ddots & \ddots & \\ & & B_n & P_n \end{pmatrix}\begin{pmatrix} I_1 & Q_1 & & \\ & I_2 & \ddots & \\ & & \ddots & Q_{n-1} \\ & & & I_n \end{pmatrix} \tag{3}$$

where I_j are identity matrices.

We can find that

$$\left.\begin{array}{l} P_1=A_1 \quad , \quad Q_1=A_1^{-1}C_1 \\ P_i=A_i-B_iQ_{i-1} \\ Q_i=P_i^{-1}C_i \end{array}\right\} \quad i=2\,,\,\ldots\ldots\,,\,n \tag{4}$$

For numerical stability, we need that P_i and Q_i $(i=1, \ldots\ldots, n)$ are bounded for some norm, i. e. $\max\,(\|P_i\|, \|Q_i\|)\leqq M$.

Definition 1.

The matrix A is block diagonally dominant with respect to the matrix norm if

$$\|B_i\|+\|C_i\|\leqq\|A_i\| \quad , \quad i=1\,,\,\ldots\ldots\,,\,n \tag{5}$$

[2]

Here we assume that A_i is nonsingular and $C_i\neq0$.

THEOREM 2.1

If A is block diagonally dominant, then

$$(a)\quad \|Q_i\|<1 \quad , \quad (b)\quad \|P_i\|\leqq\|A_i\|+\|B_i\|$$

so that the factorization is numerically stable.

Proof. We show (a) by induction. For $i=1$, from (5)

$$\|Q_1\|=\|P_1^{-1}\|\|C_1\|=\|A_1^{-1}\|\|C_1\|\leqq\|C_1\|(\|B_1\|+\|C_1\|)^{-1}<1$$

Assume $\|Q_i\|<1$ for $j-1$. Then from (4)

$$Q_i=P_i^{-1}C_i=\frac{C_i}{A_i-B_iQ_{i-1}}=\frac{A_i^{-1}C_i}{I-A_i^{-1}B_iQ_{i-1}}$$

Thus,

$$\|Q_i\|\leqq\frac{\|A_i^{-1}\|\ \|C_i\|}{1-\|A_i^{-1}\|\ \|B_i\|\ \|Q_{i-1}\|}$$

$$\leqq\frac{\|A_i^{-1}\|\ \|C_i\|}{1-\|A_i^{-1}\|\|B_i\|}$$

by the inductive assumption. Finally, by using (5),

$$\|Q_i\|<1$$

Using this result and (4), it follows that

$$\|P_i\|\leqq\|A_i\|+\|B_i\|$$

3. Algorithms of this System

The system (2) is equivalent to

$$Ly=f,\qquad Ux=y \tag{6}$$

where y also has the compound form indicated in (2)′.
Thus, we can obtain the following algorithm from (3) in (4),

$$\left.\begin{aligned}&y_1=P_1^{-1}f_1\\&y_i=P_i^{-1}(f_i-B_i),\qquad i=2,3,\ldots\ldots,n\end{aligned}\right\} \tag{7}$$

and

$$\left.\begin{aligned}&x_n=y_n\\&x_i=y_i-Q_ix_{i+1},\qquad i=n-1,n-2,\ldots\ldots,1\end{aligned}\right\} \tag{8}$$

Here, in the case of $m_i=m$, we consider to estimate the total number of operations used. [1]
We require nm^3 ops. (operations) for all P_i
and $2(3n-2)m^3$ ops. for all $P_i^{-1}C_i$ and B_iQ_{i-1}.
Thus, the evaluation of (4) involves not more than $(3n-2)m^3$ ops.
The evaluation of (7) and (8) involves

$(2n-1)m^2$ ops. for (7)

and $(n-1)m^2$ ops. for (8).
The total evaluation is thus

$(3n-2)(m^3+m^2)$ ops.

to solve the system (2) with coefficient matrix (1).

Furthermore, when we wish to improve upon the accuracy of the solution of (6), we can apply the iteration method to (7). At first, we will apply to the

first expression of (7) and, for simplicity, write as $\boldsymbol{y}=P^{-1}\boldsymbol{f}$.
Let the approximating inverse matrix of P_1 be $\bar{P}^{-1}$.

$$\boldsymbol{y}^{(1)}=\bar{P}^{-1}\boldsymbol{f}$$

$$\triangle\boldsymbol{y}^{(1)}=\bar{P}^{-1}(\boldsymbol{f}-P\boldsymbol{y}^{(1)})=\bar{P}^{-1}\boldsymbol{r}^{(1)}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\boldsymbol{y}^{(k)}=\boldsymbol{y}^{(1)}+\triangle\boldsymbol{y}^{(1)}+\triangle\boldsymbol{y}^{(2)}+\cdots\cdots+\triangle\boldsymbol{y}^{(k)}$$

$$\triangle\boldsymbol{y}^{(k)}=\bar{P}^{-1}(\boldsymbol{f}-P\boldsymbol{y}^{(k)})=\bar{P}^{-1}\boldsymbol{r}^{(k)}$$

Here, if $\{\triangle\boldsymbol{y}^{(k)}\}$ is convergent, $\{\boldsymbol{y}^{(k)}\}$ is also convergent.

$$\begin{aligned}
r^{(1)}&=\boldsymbol{f}-P\boldsymbol{y}^{(1)}=\boldsymbol{f}-P\bar{P}^{-1}\boldsymbol{f}=(I-P\bar{P}^{-1})\boldsymbol{f}\\
r^{(2)}&=\boldsymbol{f}-P(\boldsymbol{y}^{(1)}+\triangle\boldsymbol{y}^{(1)})\\
&=(I-P\bar{P}^{-1})\boldsymbol{f}-P\triangle\boldsymbol{y}^{(1)}\\
&=(I-P\bar{P}^{-1})\boldsymbol{f}-P\bar{P}^{-1}\boldsymbol{y}^{(1)}\\
&=(I-P\bar{P}^{-1})\boldsymbol{f}-P\bar{P}^{-1}(I-P\bar{P}^{-1})\boldsymbol{f}\\
&=(I-P\bar{P}^{-1})^2\boldsymbol{f}
\end{aligned}$$

In general, $\boldsymbol{r}^{(k)}=(I-P\bar{P}^{-1})^k\boldsymbol{f}$

Therefore, $\boldsymbol{y}^{(k)}$ is convergent if and only if the spectral radius $\rho(I-P\bar{P}^{-1})<1$. Thus we can apply the obtained value $\boldsymbol{y}_1^{(k)}$ to the second expression of (7) and continue this process.

Reference

1. E. B'. Keller, 'Analysis of Numerical Methods', Wiley, 1966.
2. R. S. Varga, 'Matrix Iterative Analysis', Prentice-Hall, 1962.
3. A. R. Mitchell, 'Computational Methods in Partial Differential Equations', John Wiley, 1969.
4. J. M. Varah, 'On the Solution of Block-Tridiagonal Systems Arising from Certain Finite-Difference Equations', Math. Comp. vol. 26, 1972, pp. 856-868.
5. H. B. Keller, 'Accurate Difference Methods for Linear Ordinary Differential Systems Subject to Linear Constraints', SIAM J. Numer. Anal. Vol. 6, No. 1, pp. 8-30.
6. K. Sanada, 'On Stability Criteria of Explicit Difference Schemes for Heat Equations', Bull. Fac. Educ. Kagoshima Univ., Vol. 23, pp. 6-11.