Comm. Korean Math. Soc. **11** (1996), No. 2, pp. 495–501

A PARTIAL PROOF OF THE CONVERGENCE OF THE BLOCK-ADI PRECONDITIONER

SANGBACK MA

ABSTRACT. There is currently a regain of interest in ADI (Alternating Direction Implicit) method as a preconditioner for iterative Method for solving large sparse linear systems, because of its suitability for parallel computation. However the classical ADI is not applicable to FE(Finite Element) matrices. In this paper we propose a Block-ADI method, which is applicable to Finite Element matrices. The new approach is a combination of classical ADI method and domain decompositi on. Also, we provide a partial proof of the convergence based on the results from the regular splittings, in case the discretization matrix is symmetric positive definite.

1. Introduction

Finite difference or finite element discretizations of the following partial differential equation(PDE)

$$-(K_1(x,y)u_x)_x - (K_2(x,y)u_y)_y + f(x,y)u = g(x,y) \quad (1)$$
$$\Omega = (0,1) \times (0,1)$$
$$u = 0 \ \text{ on } \ \partial\Omega$$

with meshsize $h = 1/(n+1)$ give rise to a linear system

$$Au = b \quad (2)$$

of order $N = n \times n$, where the the matrix A is a sparse matrix. The matrix A is nonsymmetric due to the presence of the terms of first order derivatives. In particular, we call the problem of the Laplace equation

Received March 21, 1995. Revised January 12, 1996.
1991 AMS Subject Classification: 65F10.
Key words and phrases: Preconditioning; ADI; Block ADI; Domain decomposition.

on a square with Diri chlet boundary condition the model problem. For throughout this paper we assume that the underlying PDE is positive. In case of finite difference method(FDM) with standard central difference for the first-order derivatives, we can split $A = H + V + \Sigma$, where H and V comes from the discretizations in x and y directions, respectively, and Σ comes from the term f in Eq. (1). Let A be the matrix obtained by the standard 5-point finite difference operator with the unknowns ordered in the *natural ordering.* Then, a *block tridiagonal* matrix is obtained, H is a tridiagonal matrix, and V is a block tridiagonal matrix.

For Eq. (2) we decompose A as$A = H_0 + V_0 + \Sigma$, where Σ comes from the fu component, and H_0, V_0 are the contributions from the x, and y directional derivatives, respectively. With $H = H_0 + \frac{1}{2}\Sigma$ and $V = V_0 + \frac{1}{2}\Sigma$, PR2-ADI(Peaceman-Rachford ADI) method is defined by

$$(3) \qquad (H + \rho_i I)u_{i+1/2} = -(V - \rho_i I)u_i + b$$

$$(4) \qquad (V + \rho_i I)u_{i+1} = -(H - \rho_i I)u_{i+1/2} + b, \quad i \geq 0,$$

where u_0 is an arbitrary initial vector approximation of the solution of Eq. (1), and $\{\rho_i,\ i \geq 0\}$ are positive constants called *acceleration parameters*, which are chosen to speedup the convergence of this process. Each of Eq. (3) and (4) form n sets of linear system of order n where the n linear systems are completely *decoupled.* Furthermore, the matrices H and V could be made *tridiagonal* with proper reordering. For example, under natural ordering in x direction H is tridiagonal, and with natural ordering in y direction V could be made tridiagonal. This ensures a minimum degree of parallelism of n, which makes PR2-ADI attractive in parallel computations. Also we note that Gaussian elimination method for the tridiagonal linear systems is very effective in terms of costs.

We combine Eq. (3) and Eq. (4) into the form

$$(5) \qquad u_{i+1} = T_{\rho_i}\ u_i + v \ \ i \geq 0,$$

where

$$(6) \qquad T_{\rho_i} \equiv (V + \rho_i I)^{-1}(\rho_i - H)(H + \rho_i I)^{-1}(\rho_i I - V),$$

and

$$(7) \qquad v = (V + \rho_i I)^{-1}\{(\rho_i - H)(H + \rho_i I)^{-1} + I\}b.$$

We call T_{ρ_i} the *Peaceman-Rachford matrix.* If $\varepsilon_i = u_i - u$ is the error at the i-th iteration, then $\varepsilon_{i+1} = T_{\rho_i}\varepsilon_i$, and in general

$$\varepsilon_i = \left(\prod_{j=1}^{j=i} T_{\rho_j}\right)\varepsilon_0, \quad i \geq 1, \tag{8}$$

where

$$\left(\prod_{j=1}^{j=i} T_{\rho_j}\right) \equiv T_{\rho_i}T_{\rho_{i-1}}\dots T_{\rho_1} \tag{9}$$

As for the convergence of Peaceman-Rachford iteration, we first consider the ***stationary*** case, where all the constants ρ_i are equal. Then we have the following theorem[Va62].

THEOREM 1.1. *Let H and V be $N \times N$ Hermitian non-negative definite matrices, where at least one of the matrices H and V is positive definite. Then, for any $\rho > 0$, the stationary PR2-ADI iteration is convergent.*

The above result still holds true without the assumption that H and V commute, i.e, $HV = VH$. This is very crucial, since the condition that H and V commute is a very str ingent one. Actually, it dictates that the underlying PDE which gave rise to the matrix is *separable*, i.e,

$$K_1(x,y) = K_1(x,y_0), \forall y, \quad K_2(x,y) = K_2(x_0,y), \forall x \tag{10}$$

and

$$d(x,y) = d(x,y_0), \forall y, \quad e(x,y) = e(x_0,y), \forall x \tag{11}$$

2. Block versions of ADI

Since H and V depend on the finite difference discretizations of original PDEs, the classical ADI is not defined for FE matrices. For example, the piecewise linear shape functions on triangles give rise to 7-point matrices, for which there is no natural splitting of A into the sum of two matrices H and V that are both tridiagonal, or defined discretizations of one-dimensional operators. The question then arises as to how to generalize the classical ADI for Finite Elements applications. There are several options available. Here we propose a technique which is based on recasting the Peaceman Rachford ADI in the framework of *Domain Decomposition* methods.

2.1. The classical ADI and Domain Decomposition

In the classical ADI of Eq. (3) and (4), H is the discretization matrix of x-directional derivatives. In terms of domain decomposition the domain is decomposed into *horizontal* lines. Then H is obtained by applying the original PDE on the subdomains, while imposing the Neumann boundary conditions between the *vertical* lines. Similarly for V. After H and V are found we could write A as

$$A = H + (A - H) = V + (A - V).$$

These two splittings of A are used in each of two stages of the iteration (3). A parameter ρ_i was added to the diagonals of H and V as a relaxation parameter. In other words ADI can be viewed as an extreme case of domain decomposition in the plane, where the subdomain consists of nonoverlapping horizontal rectangles consisting of one line each. We can also view ADI as a means of using a domain decomposition strategy to reduce two-dimensional domains into one-dimensional subdomains. By alternating between the x and y directions we can achieve the overlapping between the domains that is desirable in domain decomposition. As we noted earlier in domain decomposition the convergence deteriorates if the number of subdomains increases and there is no overlap between the subdomains. By the *alternation* we hope to achieve the equivalent effect of overlapping subdomains.

2.2. A Block ADI Algorithm

We have seen in the previous discussions that the two stages of the classical ADI are characterized by the way in which the matrix A is split in two additive components. It is natural to think of considering the subdomains of horizontal/vertical *strips* consisting of a few, say k, lines, instead of just one line. The same procedure as in the classical ADI can then be defined. Let us call ADI(k) this variant of ADI, and let $H^{(k)}$ and $V^{(k)}$ denote the matrices obtained by applying the original PDE on this decomposition of the domains. In essence, for each of the two domain partitionings, these matrices are obtained from the original matrix by neglecting the interactions between grid points across interfaces, or rather replacing them with Neumann boundary conditions. Then A is split as

$$A = H^{(k)} + (A - H^{(k)}) = V^{(k)} + (A - V^{(k)}).$$

From this we can define our block ADI procedure, denoted by ADI(k).

ALGORITHM 2.1. **ADI(k)**

$$(H^{(k)} + \rho_i I)u_{i+1/2} = -(A - H^{(k)} - \rho_i I)u_i + b \tag{12}$$

$$(V^{(k)} + \rho_i I)u_{i+1} = -(A - V^{(k)} - \rho_i I)u_{i+1/2} + b \tag{13}$$

In the case of model problem with $N = 4 \times 4$ grid $H^{(2)}$ is as follows.

$$H^{(2)} = \begin{pmatrix} C^{(2)} & 0 \\ 0 & C^{(2)} \end{pmatrix}$$

where

$$C^{(2)} = \begin{bmatrix} 3 & -1 & 0 & 0 & -1 & 0 & 0 & 0 \\ -1 & 3 & -1 & 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 3 & -1 & 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 3 & 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 & 3 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 & -1 & 3 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & -1 & 3 & -1 \\ 0 & 0 & 0 & -1 & 0 & 0 & -1 & 3 \end{bmatrix}.$$

But for $k > 1$ $H^{(k)}$ and $V^{(k)}$ no longer commute even for the Laplace equation, and no longer $A = H^{(k)} + V^{(k)}$ is true. These make the analysis very di fficult, since we can neither write out the spectral radius in terms of the eigenvalues. However, as k increases $H^{(k)}$ and $V^{(k)}$ get *closer* to A. In terms of the Frobenius norm

$$\|A - H^{(k)}\|_F < \|A - H\|_F \,, \|A - V^{(k)}\|_F < \|A - V\|_F \,, k > 1 \tag{15}$$

where $\|A\|_F$ denotes the Frobenius norm of the matrix defined by

$$\|A\|_F{}^2 = \sum a_{i,j}{}^2 \tag{16}$$

This leads us to expect that the Block ADI(k) iteration will also converge. In fact when ρ is sufficiently large and $H^{(k)}$ and $V^{(k)}$ are symmetric for all k, we have the following result. First, we recall that a matrix A is called a *Stieltjes* matrix if it is symmetric positive definite and $a_{i,j} \leq 0$ for $i \neq j$. We denote by Sp(A) the spectral radius of A. Then, we have the following theorem[Va62].

THEOREM 2.1. *Let $A = M_1 - N_1 = M_2 - N_2$ be two regular splittings of A, where $A^{-1} > 0$. If $N_2 \geq N_1 \geq 0,\ N_1 > 0, N_1 \neq N_2$, then*

$$0 < Sp(M_2{}^{-1}N_2) < Sp(M_1{}^{-1}N_1) < 1. \tag{17}$$

PROPOSITION 2.1. *Let A^{-1} ¿ 0 and suppose ρ is large enough so that $(\rho_i I - H) > 0$, $(\rho_i I - V) > 0$, and $H^{(k)}$ and $V^{(k)}$ are Stieltjes matrices. Then, we have*

$$\begin{aligned} 0 &< Sp((H^{(k)} + \rho_i I)^{-1}(\rho_i I - (A - H^{(k)}))) \\ &\leq Sp((H + \rho_i I)^{-1}(\rho_i I - V)) < 1\,, \\ 0 &< Sp((V^{(k)} + \rho_i I)^{-1}(\rho_i I - (A - V^{(k)}))) \\ &\leq Sp((V + \rho_i I)^{-1}(\rho_i I - H)) < 1\,. \end{aligned} \tag{18}$$

PROOF. If $H^{(k)}$ and $V^{(k)}$ are Stieltjes matrices then $H^{(k)} + \rho_i I$ and $V^{(k)} + \rho_i I$ are M-matrices ([Va62]). Also, $(\rho_i I - (A - H^{(k)})) > (\rho_i I - (A - H)) = (\rho_i I - V) > 0$. Then, $A = (H + \rho_i I) - (\rho_i I - (A - H)) = (H^{(k)} + \rho_i I) - (\rho_i I - (A - H^{(k)}))$ are **regular splittings**. Then, with the theorem 2.1 the result foll ows. Similarly, for $V^{(k)}$.

For $k > 1$ we have,

COROLLARY 2.1. *Assume that the matrix A resulting from the discretization of the original PDE is Stieltjes matrix. If ρ is sufficiently large then the stationary ADI(k) iteration will con verge.*

PROOF. The hypothesis implies that $A, H^{(k)}$ and $V^{(k)}$ are symmetric. We need to show the spectral radius of

$$T_\rho^{(k)} = (V^{(k)}+\rho_i I)^{-1}(A-V^{(k)}-\rho_i I)(H^{(k)}+\rho_i I)^{-1}(A-H^{(k)}-\rho_i I) \tag{19}$$

is less than 1. Using the symmetry of $A, H^{(k)}$ and $V^{(k)}$,

$$\begin{aligned}
&Sp(T_\rho^{(k)})\\
\leq&\|(V^{(k)} + \rho_i I)^{-1}(A - V^{(k)} - \rho_i I)(H^{(k)} + \rho_i I)^{-1}(A - H^{(k)} - \rho I)\|_2\\
\leq&\|(V^{(k)} + \rho_i I)^{-1}(A - V^{(k)} - \rho_i I)\|_2\|(H^{(k)} + \rho_i I)^{-1}(A - H^{(k)} - \rho_i I)\|_2\\
=&Sp((V^{(k)} + \rho_i I)^{-1}(A - V^{(k)} - \rho_i I))\\
&Sp((H^{(k)} + \rho_i I)^{-1}(A - H^{(k)} - \rho_i I))\\
<&1, \quad \text{by proposition 2.1} \qquad \square
\end{aligned}$$

If $H^{(k)}$ and $V^{(k)}$ come from FEM discretizations on the subdomains consiting of k horizontal/vertical lines, $H^{(k)}$ and $V^{(k)}$ are Stieltjes matrices. Also, if the matrix A is obtained by standard FEM discretizations with piecewise lin ear functions on triangle elements, A is irreducible and Stieltjes matrix, hence $A^{-1} > 0$, by corollary 3 in page 85, of [3]. Then, with the model problem if $\rho > 1$ the hypothesis of the proposition 2.1 is satisfied, hence the ADI(k) method for the FEM will converge.

References

1. G. Birkhoff, R. Varga, S. R., and D. Young, *Alternating direction implicit methods*, in Advances in Computers, pp. 189-273, Academic Press, New York, 1962.
2. D. Peaceman and H. Rachford, *The numerical solution of elliptic and parabolic differential equations*, Journal of SIAM. **3** (1955), 28-41.
3. R. Varga, *Matrix Iterative Analysis*, Prentice-Hall, New York, 1962.

Dept of Computer Science
Hanyang University
Ansan, Kyungki Do