*Appl. Math. Lett.* Vol. 4, No. 2, pp. 75–79, 1991
Printed in Great Britain. 
0893-9659/91 $3.00 + 0.00


# A Note on a Block Preconditioner

MIRON TISMENETSKY
Institute of Supercomputing and Applied Mathematics
IBM Scientific Center, Heidelberg, Germany

*(Received July 1990)*

**Abstract.** The well known block preconditioning technique is modified to suit certain ill-conditioned linear systems arising in reservoir modelling, semiconductor simulation and other fields of applications. Incorporated into the conjugate or biconjugate gradients algorithm, the proposed preconditioner gives a significant improvement in the condition number and the resulting convergence rate.

## INTRODUCTION

Sparse linear systems

$$A\mathbf{x} = \mathbf{b} \tag{1}$$

with a block tridiagonal coefficient matrix $A = [A_{i,j}]_{i,j=1}^{m}$, where $A_{i,j} \in \mathbf{R}^{n\times n}$, $A_{i,j} = 0, (|i-j| > 1), i, j = 1, 2, \ldots, m$ and sparse nonzero blocks appear frequently in practice. A powerful and popular class of methods for solving (1) is based on a block incomplete $LDU$ factorization of the matrix (so-called preconditioning) incorporated into the conjugate or biconjugate gradients algorithm.

The subject of this report is a modification of the basic block preconditioning algorithm for solving (1) proposed and treated in [2–4,7,8,13–16].

In presenting this algorithm (without row sum compensation of errors in spirit of [9]), we use the notation $\tilde{S}$ for the matrix obtained from $S$ by replacing its entries outside the preassigned (usually, the original) support by zeros. An approximate inverse of $S$ is denoted by $\mathrm{inv}(S)$.

**A block incomplete preconditioning algorithm.** Set $\tilde{U}_1 = U_1 = A_{11}$ and compute for $j = 1, 2, \ldots, m-1$:

$$U_{j+1} = A_{j+1,j+1} - A_{j+1,j}\mathrm{inv}(\tilde{U}_j)A_{j,j+1}, \tag{2}$$

Replace the matrix $U_{j+1}$ by $\tilde{U}_{j+1}$.

Let $b_j$ denote the $j$-th component of $\mathbf{b}$. The preconditioner for (1) is then found in the following way: set $y_1 = b_1$ and compute $y_{j+1} = b_{j+1} - A_{j+1,j}\tilde{U}_j^{-1}y_j$ for $j = 1, 2, \ldots, m-1$. The components $x_j$ of the solution vector $\mathbf{x}$ are now computed recursively by the formulas

$$x_m = \tilde{U}_m^{-1}y_m, \quad x_j = \tilde{U}_j^{-1}(y_j - A_{j,j+1}y_{j+1}) \quad (j = m-1, m-2, \ldots, 1) \qquad \blacksquare$$

The objective of this report is an improved computation of the approximate inverse $\mathrm{inv}(U_j)$ of $U_j$ (we omit the notation 'tilde') which in the standard way is obtained by the formula

$$\mathrm{inv}(U_j) = (2I - D_j^{-1}U_j)D_j^{-1}, \tag{3}$$

---

On leave from IBM Haifa Scientific Center, Technion City, Haifa, Israel.
The author wishes to thank Dr. R. Janssen for encouraging this work.

where $I$ stands for the $n \times n$ identity matrix and $D_j$ denotes the diagonal part of $U_j$. As shown in [4], the use of (3) is satisfactory when the matrix $A$ is a nonsingular M-matrix.

However, matrices appearing in a wide range of applications usually do not obey this condition and are extremely ill-conditioned. The spectral radius of the matrices $I - D_j^{-1}U_j$ in (3) may be much greater than one; and, therefore, the idea of replacing the true inverse by a part of the (divergent in this case) Neumann series cannot lead to satisfactory results.

In the symmetric positive definite case this obstacle can be overcome (see [16]) by introducing a scalar $\theta_j = 2 \parallel D_j^{-1}U_j \parallel^{-1}$, where, for instance, the row sums norm is used, and replacing (3) by

$$\mathrm{inv}(U_j) = (2I - \theta_j D_j^{-1}U_j)(\theta_j D_j^{-1})$$

It is easily verified that the spectral radius of the matrix $I - \theta_j D_j^{-1}U_j$ is less than 1 and hence the resulting approximate inverse is satisfactory in many cases.

In this report, we suggest a technique for computing approximate inverse of $U_j$ in cases when the matrices $U_j$ are not necessarily definite, and the large magnitudes of the spectral radius are caused by a few relatively large eigenvalues, which are separated from the rest of the spectrum. The approach adopted here is closely related to the idea of 'inserting points' proposed by V. Amdursky [1]. Note that a preliminary version of this work is reported in [16].

## 1. A Description of the Approach

The approach proposed in this work relies on the notion of the *Schur complement* $M/M_{11}$ of the square matrix $M = [M_{i,j}]_{i,j=1}^2$ defined, when $M_{11}$ is nonsingular, by the expression

$$M/M_{11} \doteq M_{22} - M_{21}M_{11}^{-1}M_{12}$$

Now observe that the equation in (2) shows, assuming that $\mathrm{inv}(U_j)$ stands for the exact inverse of $U_j$, that

$$U_{j+1} = M_j/U_j, \text{ where } M_j = \begin{bmatrix} U_j & A_{j+1,j} \\ A_{j,j+1} & A_{j+1,j+1} \end{bmatrix}.$$

Let $E_j$ stand for an arbitrary nonsingular principal submatrix of $U_j$. Using the well known *quotient formula* (see [5,6]), we obtain

$$U_{j+1} = (M_j/E_j)/(U_j/E_j) \tag{1.1}$$

The choice of $E_j$ is at our disposal and will be discussed in the next section.

To bring (1.1) into a form appropriate for a practical use, we need some additional notations. Given an $n \times n$ matrix $S$ and a set of indices $\mathbf{k} = \{k_1, k_2, \ldots, k_r\}, 1 \le k_1 < k_2 < \ldots < k_r \le n$, we denote by $S^{[k;]}$ (resp., $S^{[;k]}$) the submatrix of $S$ containing the rows (resp., columns) whose indices are in $\mathbf{k}$. The submatrices of $S$ obtained by deleting the rows and columns in $\mathbf{k}$ are denoted by $S^{(k;)}$ and $S^{(;k)}$, resp. The notation $S^{(k;k)}$ stands for the $(n-r) \times (n-r)$ submatrix of $S$ obtained by deleting the rows and columns with indices in $\mathbf{k}$. Finally, the submatrix of $S$ containing rows (resp., columns) in $\mathbf{k}$ with excluded columns (resp., rows) $\mathbf{k}$ is denoted by $\bar{S}^{[k;]}$ (resp., $\bar{S}^{[;k]}$).

Using this notation, we easily derive the formula $M^j/E^j = [F_{ij}]_{i,j=1}^2$, where

$$F_{11} = U_j^{(k;k)} - \bar{U}_j^{[;k]}E_j^{-1}\bar{U}_j^{[k;]} = U_j/E_j,$$

$$F_{12} = A_{j,j+1}^{(k;)} - \bar{U}_j^{[;k]}E_j^{-1}A_{j,j+1}^{[k;]}, \quad F_{21} = A_{j+1,j}^{(;k)} - A_{j+1,j}^{[;k]}E_j^{-1}\bar{U}_j^{[k;]},$$

$$F_{22} = A_{j+1,j+1} - A_{j+1,j}^{[;k]}E_j^{-1}A_{j,j+1}^{[k;]}$$

and $E_j$ stands for the submatrix of $U_j$ located on the intersection of row and columns in $\mathbf{k}$.

Formula (1.1) can be thus rewritten in the form

$$U_{j+1} = F_{22} - F_{21}F_{11}^{-1}F_{12}$$
$$(F_{11} \in \mathbf{R}^{(n-r)\times(n-r)},\ F_{12} \in \mathbf{R}^{(n-r)\times n},\ F_{21} \in \mathbf{R}^{n\times(n-r)},\ F_{22} \in \mathbf{R}^{n\times n}) \tag{1.2}$$

The original structure (often banded) of the matrices involved in the computation of $F_{11}, F_{12}$, $F_{21}$ is perturbed. To preserve the structure of these matrices (desirable for implementation purposes), we proceed as follows.

The matrix $U_j^{(k;k)}$ is replaced by the $n \times n$ matrix $\hat{U}_j$ obtained from $U_j$ by placing zeros in the rows and columns $\mathbf{k}$ except at their intersection which is set to be a $r \times r$ identity matrix (up to permutations). The matrix $\bar{U}_j^{[k;]}$ (resp., $\bar{U}_j^{[;k]}$) is replaced by the matrix $\hat{U}_j^{[k;]}$ (resp., $\hat{U}_j^{[;k]}$) derived from $U_j^{[k;]}$ (resp., $U_j^{[;k]}$) by setting the elements of columns (resp., rows) $\mathbf{k}$ to be zeros.

The matrix $\hat{F}_{11} \doteq \hat{U}_j - \hat{U}_j^{[;k]}E_j^{-1}\hat{U}_j^{[k;]}$ thus has the following properties: $\hat{F}_{11}^{[k;k]} = I$, $\hat{F}_{11}^{(k;k)} = F_{11}$ and the elements of the rows and columns with indices in $\mathbf{k}$ (except those lying at their intersection) are zeros.

Denote by $\hat{A}_{j+1,j}^{(;k)}$ (resp., $\hat{A}_{j,j+1}^{(k;)}$) the matrix derived from $A_{j+1,j}$ (resp., $A_{j,j+1}$) on setting the elements of columns (resp., rows) in $\mathbf{k}$ to be zeros. Hence the matrices

$$\hat{F}_{12} = \hat{A}_{j,j+1}^{(k;)} - \hat{U}_j^{[;k]}E_{j-1}\hat{A}_{j,j+1}^{[k;]},\ \hat{F}_{21} = \hat{A}_{j+1,j}^{(;k)} - \hat{A}_{j+1,j}^{[;k]}E_j^{-1}\hat{U}_j^{[k;]}$$

possess the property of having zero rows (resp., columns) numbered by indices in $\mathbf{k}$.

In the notation introduced above, the equation (1.2) can be thus equivalently rewritten in the form

$$U_{j+1} = F_{22} - \hat{F}_{21}\hat{F}_{11}^{-1}\hat{F}_{12},$$

where the original numeration of nonzero patterns is easily preserved.

To reduce the number of arithmetic operations required for computing $U_{j+1}$, the exact inverse of $\hat{F}_{11}$ is replaced by its approximate $\mathrm{inv}(\hat{F}_{11})$. This can be done in several ways (see [2 – 4, 7, 11, 13 – 16]). The most standard approach is to use truncated Neumann series (provided the convergence is assured). A different approach for determining an approximate inverse of the matrix $\hat{F}_{11} \doteq \hat{U}_j - \hat{U}_j^{[;k]}E_j^{-1}\hat{U}_j^{[k;]}$ proposed in this work, is based on the Woodbury-Sherman-Morrison formula (cf. e.g., [10, 12]):

$$(A - BD^{-1}C)^{-1} = A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1}, \tag{1.3}$$

where on the $j$-step we have $A = \hat{U}_j, B = \hat{U}_j^{[;k]}, C = \hat{U}_j^{[k;]}, D = E_j$.

A right choice of the index set $\mathbf{k}$ makes the matrix $\hat{U}_j$ better than $U_j$ in the sense that the approximate formula (3) for $\mathrm{inv}(\hat{U}_j)$ gives better results than for $U_j$. Observe that the matrix $E_j$ is small and hence the inversion of the middle factor on the right in (1.3) can be performed exactly.

## 2. The Choice of the Matrix $E$

The accuracy of the approximate inverse depends on the choice of the matrix $E$ on each stage of the algorithm.

Consider first the computation of the approximate inverse by formula (3). In this case the optimal choice of $E$ is reduced to determining indices $1 \le k_1 < k_2 \ldots < k_r \le n$ such that the spectral radius of $I - F = I - U/E$ is as small as possible.

Let the matrix $U$ be symmetric positive definite with eigenvalues $0 \le \lambda_1 \le \lambda_2 \le \ldots \le \lambda_n$. The matrix $E$ defined by the index set $\mathbf{k}$ is a principal submatrix of $U$ and hence the Schur complement $U/E$ is also symmetric and positive definite (cf. e.g., [5, 12]) having eigenvalues denoted by $0 \le \tilde{\lambda}_1 \le \tilde{\lambda}_2 \le \ldots \le \tilde{\lambda}_r$. Furthermore, it is well known that the inverse of the Schur complement $U/E$ coincides with the submatrix of $U^{-1}$ obtained by

deleting the rows and columns in $\mathbf{k}$. Thus, using interlacing property (cf. e.g., [12]), we have $\lambda_j \leq \tilde{\lambda}_j \leq \lambda_{j+r}$ for each $j = 1, 2, \ldots, n-r$. Hence $|1 - \tilde{\lambda}_{n-r}| \leq \max(|1 - \lambda_{n-r}|, |1 - \lambda_n|)$, $|1 - \tilde{\lambda}_1| \leq \max(|1 - \lambda_1|, |1 - \lambda_{1+r}|)$ and, consequently, the spectral radii of the matrices $I - U/E$ and $I - U$ are related as follows:

$$\rho(I - U/E) = \max(|1 - \tilde{\lambda}_1|, |1 - \tilde{\lambda}_{n-r}|) \leq \max(|1 - \lambda_1|, |1 - \lambda_n|) = \rho(I - U)$$

We thus arrived at the following result.

PROPOSITION. *For any choice of the principal submatrix $E$ of the symmetric positive definite matrix $U$, the spectral radius of $I - U/E$ cannot exceed the spectral radius of $I - U$.*

On the other hand, the interlacing relations show that the effect of using an $r \times r$ matrix $E$ is restricted (even in the ideal case) to 'eliminating' the $r$ largest eigenvalues of $U$. Indeed, let, for instance, $\lambda_1 + \lambda_{n-r} \geq 2$. Then, it is easily verified that $\rho(I - U/E) = \hat{\lambda}_{n-r} - 1 \geq \lambda_{n-r} - 1$ and the optimal choice of $E$ is such that $\hat{\lambda}_{n-r}$ is as close to $\hat{\lambda}_{n-r}$ as possible.

The general problem of determining the index set $\mathbf{k}$ of the preassigned dimension $r$ can be formulated in the following way.

PROBLEM. *Given a nonsingular matrix $U$ and an integer $r$. Find a set $\mathbf{k}$ of $r$ indices determining a principal submatrix $E$ of $U$ such that the spectral radius of $I - U/E$ is minimal among all $r \times r$ principal submatrices of $U$.*

A similar (in fact, a dual) problem arises in the case when the computation of an approximate inverse is performed by using the Woodbury-Sherman-Morrison formula. In this case a set $\mathbf{k}$ is to be found to achieve the minimum of the spectral radius of the matrix $I - \hat{U}$, where $\hat{U}$ is derived from $U$ by replacing its off-diagonal elements in rows and columns $\mathbf{k}$ by zeros and the corresponding diagonal entries—by ones. Due to the relation between the Schur complement and the appropriate submatrix of the inverse, the first problem is, in fact, equivalent to the second applied to the inverse matrix. Note that, as in the first case, the spectral radius of $I - \hat{U}$ cannot exceed the spectral radius of $I - U$ for any choice of $\mathbf{k}$.

A right choice of the matrix $E$ is often known or can be easily found from physical or other considerations (see [14]). When such information is absent, the following heuristics can be suggested. Numerical experiments show that in the simplest case $r = 1$ (that is, dim $E = 1$) a good choice of the integer $k$ for a symmetric matrix $U$ is the number of its row (or column) at which the row sums norm $\| U \|$ is achieved. In case of several such rows the integer which is most close to the half of the matrix size gives best results.

In the nonsymmetric case we suggest to choose $k$ as the number of row such that the quotient of the $l_1$-norm of column $k$ to the same norm of row $k$ is maximal. When several such rows occur, the number of th row with the maximal $l_1$-norm is taken.

An investigation of these problems wiil be carried out in a subsequent work.

## REFERENCES

1. V. Amdursky, private communication, (1984).
2. O. Axelsson, A survey of preconditioned iterative methods for linear systems of algebraic equations, *BIT* **25**, 166–187 (1985).
3. O. Axelsson, *Solution of Linear Systems of Equations: Iterative Methods, Sparse Matrix Techniques*, V.A. Barker, Ed., Springer-Verlag, pp. 1–51, (1977).
4. P. Concus, G.H. Golub and G. Meurant, Block preconditioning for the conjugate gradient methods, *SIAM J. Sci. Stat. Comp.* **6** (1), 120–252 (1985).
5. R. Cottle, Manifestations of the Schur complement, *Lin. Alg. and Appl.* **8**, 189–211 (1974).
6. D.E. Crabtree and E.V. Haynsworth, An identity for the Schur complement of a matrix, *Proc. Amer. Math. Soc.* **22**, 36-366, (1969).
7. D.J. Evans, Editor, *Preconditioning Methods, Theory and Applications*, Gordon and Breach, New York, (1983).
8. I. Efrat and M. Tismenetsky, Parallel iterative linear solvers for oil reservoir models, *IBM J. Res. and Develop.* **30** (2), 184–192 (1986).
9. I. Gustafsson, A class of first order factorizations, *BIT* **18**, 142–156 (1978).

10. H.V. Hederson and S.R. Searle, On deriving inverse of sum of matrices, *SIAM Review* **23**, 53–60 (1981).
11. O.G. Johnson, C.A. Micchelli and G. Paul, Polynomial preconditioning for conjugate gradient calculations, *SIAM J. Num. Anal.* **20**, 362–376 (1983).
12. P. Lancaster and M. Tismenetsky, *The theory of matrices*, Academic Press, Orlando, (1985).
13. G. Meurant, The block preconditional conjugate method on vector computers, *BIT* **24**, 263–273 (1985).
14. S. Schlafman and M. Tismenetsky, On a preconditioned conjugate gradient technique for singular structural analysis problems, *IBM Haifa Sc. Report* **274** (1989).
15. M. Tismenetsky and I. Efrat, An efficient preconditioning algorithm and its analysis, *IBM Haifa Sc. Report* **172** (1985).
16. M. Tismenetsky, A modification of the MBID preconditioning algorithms, *IBM Haifa Sc. Report* **213** (1987).

Institute of Supercomputing and Applied Mathematics, IBM Scientific Center, D-6900 Heidelberg, West Germany