SOCIETY OF PETROLEUM ENGINEERS OF AIME PAPER NUMBER SPE 5729

Orthomin, an Iterative Method for Solving Sparse Sets of Simultaneous Linear Equations

By

P. K. W. Vinsome, Member SPE-AIME, Brunei Shell Petroleum Co. Ltd.

THIS PAPER IS SUBJECT TO CORRECTION
This paper was prepared for the Fourth Symposium of Numerical Simulation of Reservoir Performance of the Society of Petroleum Engineers of AIME to be held in Los Angeles, Calif., Feb. 19-20, 1976. Permission to copy is restricted to an abstract of not more than 300 words. Illustrations may not be copied. The abstract should contain conspicuous acknowledgment of where and by whom the paper is presented. Publication elsewhere after publication in the JOURNAL OF PETROLEUM TECHNOLOGY or the SOCIETY OF PETROLEUM ENGINEERS JOURNAL is usually granted upon request to the Editor of the appropriate journal, provided agreement to give proper credit is made.
Discussion of this paper is invited. Three copies of any discussion should be sent to the Society of Petroleum Engineers office. Such discussion may be presented at the above meeting and, with the paper, may be considered for publication in one of the two SPE magazines.

ABSTRACT

An iterative method is proposed for the inversion of sparse band-structured matrices of the type that are common in numerical reservoir simulators. The method utilizes orthogonalizations and minimizations to achieve a fast convergence rate. Tests have shown that the method is highly competitive compared with other iterative techniques. The rate of convergence is insensitive to the use of iteration parameters, non-symmetry of the matrix, and ratios of off-diagonal bands in symmetrical matrices.

1. INTRODUCTION

A time-consuming part of the calculation in numerical reservoir simulators is the inversion of a large set of simultaneous linear equations.

$$Ax = b \qquad \ldots\ldots \qquad (1)$$

For example when effecting a solution by the Implicit Potential - Explicit Saturation method x would represent the pressure or potential field, whilst A would be a matrix of transmissibility coefficients describing the interflow of fluids between grid blocks. In the common finite difference approaches, A is a sparse banded matrix, as depicted in figure 1 for a five-point finite difference representation of a two dimensional system. Often, the diagonal element of A is approximately equal to the negative sum of the off-diagonal elements in a given row.

Because of the sparseness of A, iterative methods appear at first sight attractive compared with direct methods. However, specialized direct routines with increased efficiency have become available in recent years[1]. Of the iterative methods SIP[2] (strongly implicit procedure) is possibly the most widely used. SIP has the disadvantage that its convergence rate depends strongly on a set of iteration parameters whose optimal values can be difficult to find.

A new highly competitive iterative method for symmetric matrices, a minimization process conceptually based on the conjugate-gradient technique[3], has lately been developed.[4] The numerical examples and proofs quoted in reference 4 indicate that minimization processes can be expected to offer the following advantages:-

a) Convergence more readily guaranteed.

b) No need for iteration parameters.

c) Insensitivity to number of equations.

d) Insensitivity to transmissibility ratios (the ratios of the off-diagonal bands in figure 1).

Hence the problem is to find a competitive minimization technique applicable to non-symmetric sparse matrices. The method that has been developed is ORTHOMIN, so called because it uses both orthogonalisations, and minimizations to achieve a high rate of convergence. ORTHOMIN is similar to other iterative techniques in as much as it uses an LDU approximation to the original matrix A. Where it differs radically is in the measures taken to increase the convergence rate i.e. orthogonalizations and minimizations versus parameter estimation.

2. THE LDU DECOMPOSITION

During given stages of the solution process it is desired to invert an equation of the form (1), or equivalently, the residual equation.

$$A\,\delta x = r^n \qquad \ldots\ldots \qquad (2)$$

where r is the remaining residual on the n'th iteration, and δx is the shift vector.

The matrix A is decomposed into an approximate LDU form as shown in figure 2. The Dupont et al[5] decomposition was adopted because of its simplicity, although any LDU decomposition would suffice. In this case in two dimensions the i'th diagonal element of D is given by

$$d_i = 1/(\gamma_i - \alpha_{1i} d_{i'} \alpha_{2i'} + \alpha_{3i} d_{i''} \alpha_{4i''}) \qquad \ldots\ldots \qquad (3)$$

with an obvious extension to three dimensions by including a term $\alpha_{5i} d_{i'''} \alpha_{6i'''}$ in the denominator (see figure 2 for a definition of the symbols).

If the diagonal D is stored, it is only necessary to do the decomposition once, and not during every iteration. The other elements of L and U need not be stored separately as they are the same as those in the original matrix A.

Instead of solving (2), the equation

$$LDU\,\delta x^n = r^n \qquad \ldots\ldots \qquad (4)$$

is then solved for an approximate shift vector δx^n, by a forward and backward substitution, which is an efficient process.

$$z = L^{-1} r^n \qquad \ldots\ldots \qquad (5)$$

$$y = D^{-1} z \qquad \ldots\ldots \qquad (6)$$

$$\delta x^n = U^{-1} y \qquad \ldots\ldots \qquad (7)$$

3. THE ORTHOMIN ITERATION

On the n'th iteration a shift vector has been produced by the LDU inversion, and from the previous iterations a set of vectors $\delta x^{n-1}, \delta x^{n-2}, \ldots\ldots \delta x^1$ has been obtained. Suppose that from the shift vectors a set of orthogonal vectors Aq^{n-1}, Aq^{n-2}, ... Aq^1 has been constructed. Using the latest shift vector δx^n, a new vector q^n is produced so that Aq^n is orthogonal to all the previous Aq's.

$$Aq^n = A\delta x^n - \sum_{i=1}^{n-1} a_i\, Aq^i \qquad \ldots\ldots \qquad (8)$$

where the a_i's are orthogonality coefficients.

For orthogonality,

$$(Aq^n/Aq^i) = 0 \qquad \ldots\ldots \qquad (9)$$

Therefore,

$$a_i = \frac{(A\delta x^n/Aq^i)}{(Aq^i/Aq^i)} \qquad \ldots\ldots \qquad (10)$$

Next, it is desired to minimise $\|r^n - wAq^n\|_2$ where w is a minimization parameter. In the second norm (a least squares fit)

$$w = \frac{(Aq^n/r^n)}{(Aq^n/Aq^n)} \qquad \ldots\ldots \qquad (11)$$

Clearly, in principle a direct method has been constructed because after N iterations (N = number of equations), there are N linearly independent vectors to completely span the space. However in practice, serious numerical rounding-off errors occur if too many orthogonalizations are performed. Furthermore the work required per iteration increases because there become more orthogonality coefficients to evaluate in (10).

The way the iteration is made more efficient is by noting that all the orthogonalizations are not necessary for quick convergence. Numerical tests have demonstrated that to orthogonalize to the previous 3, 4 or 5 vectors is optimal. Doing only 1 or 2 orthogonalizations results in too slow a rate of convergence, while after 5 orthogonalizations no appreciable increase in convergence rate with number of iterations was ever observed.

4. THE SIGNIFICANCE OF MINIMIZATION AND ORTHOGONALISATION

The significance of minimization is fairly obvious. A new vector q^n was produced, and its length adjusted by a factor w, to minimise the new residual $r^n - wAq^n$. The effect of minimization is not small. In typical problems w ranged in magnitude from 0.1 to 300, and could be either positive or negative. Because a minimization has been performed during each iteration, the new residual cannot be larger than the preceding one. Hence the method cannot diverge, but a definite proof of convergence has not been found. The method would not converge if ever a vector q^n were produced such that Aq^n were exactly orthogonal to r^n, in which case w would be zero, and the residual would stay constant during all subsequent iterations. In practice a matrix representative of reservoir engineering problems has not yet been discovered where the method does not converge.

The effect of the orthogonalizations is not so straight forward. The decomposition LDU is not exactly the same as A. The LDU inversion has a tendency to produce vectors δx^i in certain directions and not in others. Hence without orthogonalization, convergence is extremely slow because these missing subspaces do not contribute. Orthogonalization subtracts vectors lying in the directions which the LDU inversion has already found, increasing the effect of the directions which the LDU inversion is poor at supplying, and increasing the rate of convergence.

5. NUMERICAL EXAMPLES

In numerical reservoir simulators, it is common to have to solve partial differential equations of the form:

$$\frac{\partial^2 U(x, y)}{\partial x^2} + \frac{\partial^2 U(x, y)}{\partial y^2} = F(x, y) \qquad \ldots\ldots (12)$$

A five point discrete finite difference approximation to this equation produces a matrix with the structure shown in figure 1. The first two numerical examples are taken from reference 4, and are notable for their slow convergence rate.

5.1 Numerical Example 1

The system chosen was a 30 x 30 mesh (900 equations). The right hand side b satisfies the equation $Au = b$, where $u_i = 0$ for $i \neq 654$, and $u_{654} = 1$. The main diagonal of A has a value -4, and the off-diagonals if not zero are equal to 1. The results for 5 orthogonalizations in ORTHOMIN are compared with SIP and two variants of conjugate gradient (CG) in figure 3. The work was estimated by counting the number of multiplications and divisions needed for each method. For ORTHOMIN it was assumed that each iteration required 5 orthogonalizations, which is not the case during the first 5 iterations, so the work for ORTHOMIN is slightly over-estimated.

It can be seen from figure 3 that ORTHOMIN is highly competitive with the fastest known methods on symmetric matrices (two variants of conjugate gradient CG (0) and CG (3)), and easily outstrips the most commonly used method SIP. For SIP, 18 two by two equal parameters were used as recommended by Stone[2].

5.2 Numerical Example 2

For this case the off-diagonal elements if not zero, are equal to 1, while the diagonal elements are the negative sum of the off-diagonals. As this system is singular, 1 is subtracted from the first diagonal. Again 5 orthogonalizations were used for ORTHOMIN. It can be observed from figure 4 that ORTHOMIN has a smoother convergence than any of the other methods. The reason is that ORTHOMIN minimizes directly in the $\|A\tilde{x}-b\|_2$ norm, while the CG variants minimize $(\tilde{x}/A\tilde{x}-b)$. SIP does not minimize anything.

5.3 Numerical Example 3 - Effect of Number of Orthogonalizations

The last two examples, while interesting from a mathematical point of view because of their slow convergence[4], were not very representative of typical reservoir engineering problems. For the third example a 10 x 10 mesh was chosen (100 equations). The off-diagonal elements are 1, if not zero, and the diagonals are the negative sum of the off-diagonals. γ_{12} is set equal to -10, b_{88} has the value 10, and the rest of the elements of b were 0. In physical terms, this is equivalent to the solution of a pressure field obeying the Poisson equation, with a rate constrained injection well in block 88, and a pressure constrained production well in block 12.

The results are shown in figure 5. In this, as in all other examples m = 3, 4 and 5 have been tried and work satisfactorily (m = number of orthogonalizations). To convert the abscissa of figure 5 to a common work scale it is necessary to know that the number of multiplications per iteration is 15 + 3 m.

In general m = 4 is the optimal value on a work scale, but m = 5 is sometimes optimal for very large sets of linear equations which also have slow convergence. $m < 3$ usually has too slow a convergence rate, while it seems pointless going to $m > 5$ because no appreciable gain in convergence rate is obtained. In typical reservoir engineering problems the use of m = 4 is recommended.

5.4 Effect of Iteration Parameter

As the ORTHOMIN iteration may be used with any sort of decomposition the effect of an iteration parameter has also been investigated. The diagonal element of the 2 dimensional Dupont et al[5] decomposition is ~~then~~ given by

$$d_i = 1/\left[\alpha_{1i} d_{i'}(\alpha_{2i'} + R\alpha_{4i'}) \overset{\gamma_i -}{+} \alpha_{3i} d_{i''}(\alpha_{4i''} + R\alpha_{2i''})\right] \qquad \ldots\ldots (13)$$

where R is the iteration parameter. Normally R has a maximum value slightly less than unity[2].

Figure 6 shows the effect of varying the iteration parameter for problem 3. An automatic procedure for estimating the iteration parameter for SIP[2] generates $R = 0.975$. It can be seen from the figure that the convergence rate is not very sensitive to an iteration parameter. Nor is there any systematic trend in convergence rate with iteration parameter. The use of $R = 0$ i.e. no iteration parameter is recommended as it provides a safer decomposition. Notice that in all the examples herein presented, the equations have been solved to high accuracies; much more than is normally required. At lower accuracies the convergence rate is almost completely insensitive to the use of an iteration parameter.

5.5 Transmissibility Ratios and Non-Symmetric Matrices

Tests have also been carried out on symmetric matrices where the ratios of the transmissibility coefficients (ratios α_1/α_3, α_2/α_4) were varied between 10^3 and 10^{-3}. In addition strongly non-symmetric matrices have been inverted where the ratios α_1/α_2, α_3/α_4 ranged between 10^3 and 10^{-3}. Convergence was just as quick as in the preceeding examples, frequently even faster.

As a final example, solution of a strongly non-symmetric matrix is given. A 10 x 10 mesh was chosen. The off-diagonal bands, if not zero have the following values.

$$\alpha_1 = 0.1; \quad \alpha_2 = 10.0; \quad \alpha_3 = 10.0; \quad \alpha_4 = 0.1.$$

The diagonal element is equal to the negative sum of the off-diagonals. As in example 3, b_{88} and γ_{12} were given the values 10 and -10 respectively. $m = 4$ was used for the number of orthogonalizations. The convergence rate is shown in figure 7, and is extremely rapid. Notice that because $|\gamma_{12}|$ is less than the sum of the off-diagonal bands, this matrix is not even diagonally dominant.

7. CONCLUSIONS

A competitive iterative method, ORTHOMIN has been developed for the inversion of sparse banded matrices of the type that arise in numerical reservoir simulators. The method uses minimizations and orthogonalizations to achieve a fast convergence rate.

The advantages of the method are:

a) It is not difficult to code.

b) The LDU decomposition need only be done once, and not during every iteration as is the case with SIP.

c) The convergence rate is insensitive to iteration parameter, transmissibility ratios, and non-symmetry of the matrix.

d) The method is applicable to matrices with any number of bands.

The disadvantages of the method are:-

a) ORTHOMIN takes more computer storage than CG (0), but this is offset by the fact that it can be used on non-symmetric matrices.

b) The optimal value of m, the number of orthogonalisations cannot be a priority determined, although in practice $m = 4$ has always worked well and is frequently the optimal value.

c) Convergence has not yet been proven, although it can be demonstrated that divergence is impossible. In practice a matrix typical of reservoir engineering problems has not been found where the method does not converge.

When using ORTHOMIN it is recommended that the decomposition be done without iteration parameter, and that $m = 4$ is used unless the set of equations is extremely large (> 500), when $m = 5$ should suffice.

8. NOMENCLATURE

a	Orthogonality coefficient.
A	Matrix with sparse banded structure.
b	Right hand side of a system of linear equations.
CG	Conjugate gradient.
d	Diagonal element of matrix D.
D	Diagonal matrix.
h	Inverse of d; diagonal element of L, U.
L	Lower triangular matrix with sparse banded structure.
m	Number of orthogonalisations.
n	Number of iterations.
N	Number of equations.
q	Vector.
r	Residual.
R	Iteration parameter.
u	Vector.
U	Upper triangular matrix with sparse banded structure.

w Minimization parameter.

x Solution vector.

$\tilde{x}$ Approximation to x.

y, z Vectors.

α Off-diagonal element of matrices A, L, U.

γ Diagonal element of matrix A.

δx Shift vector.

9. REFERENCES

1. Price, H.S., and Coats, K.H.: "Direct Methods in Reservoir Simulation," Trans., AIME (1974) Vol. 257, II 295 - 308.
2. Stone, H.L.: "Iterative Solution of Implicit Approximations of Multidimensional Partial Differential Equations," J. Numer. Anal., SIAM (September 1968) Vol. 5, No. 3, 530 - 538.
3. Daniel, J.W.: "The Conjugate-Gradient Method for Linear and Non-Linear Operator Equations," J. Numer. Anal., SIAM (1967) Vol. 4, No. 1, 10 - 26.
4. Meijerink, J.A., and van der Vorst, H.A.: "Iterative Solution Method for Linear Systems of which the Coefficient Matrix is a Symmetric M-Matrix," J. Numer. Anal. SIAM (publication pending).
5. Dupont, T., Kendall, R.P., and Rachford, M.M.: "An Approximate Factorization Procedure for Solving Self-Adjoint Elliptic Difference Equations," J. Numer. Anal., SIAM (1968) Vol. 5, No. 3, 559 - 573.

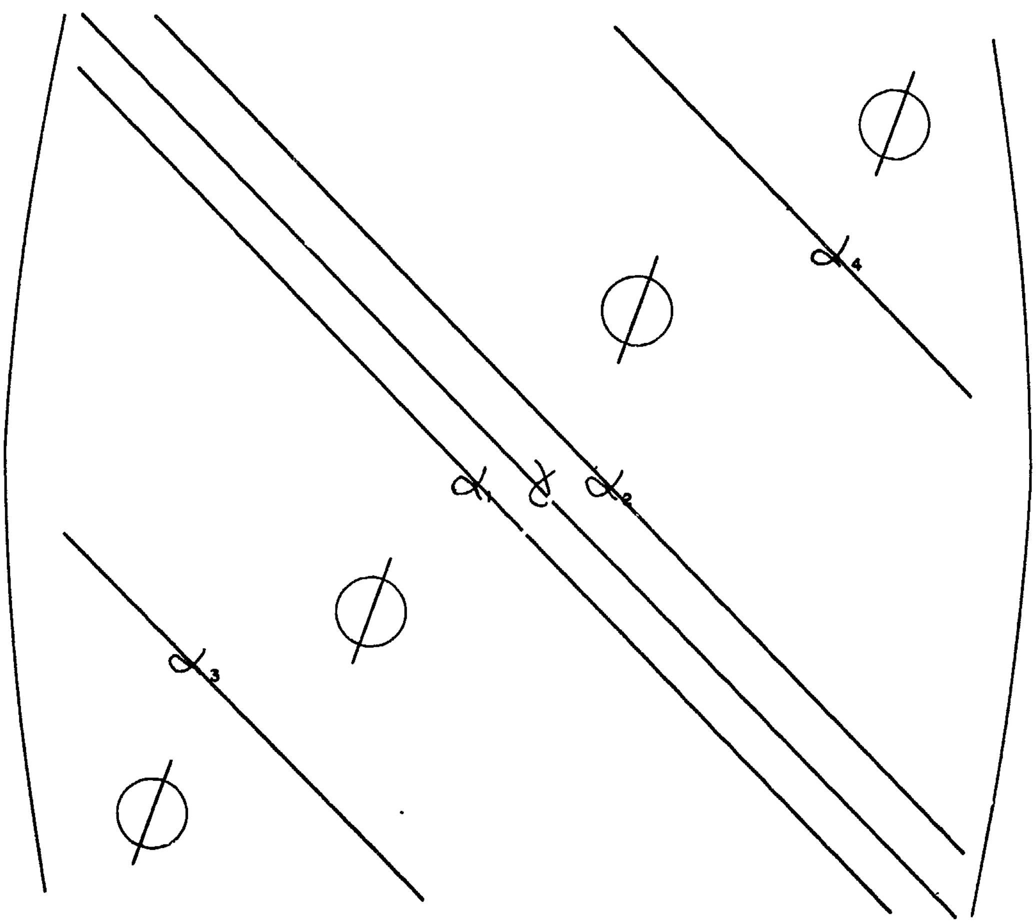
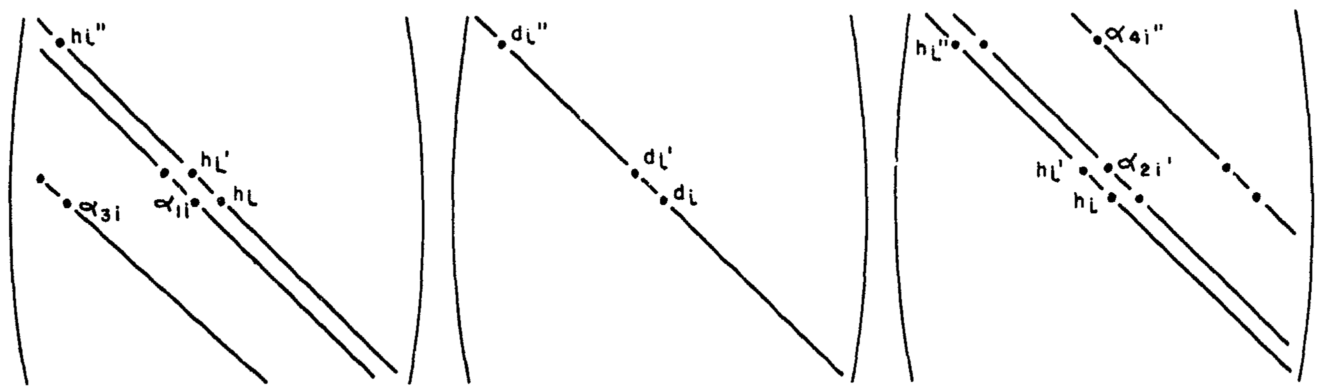


Fig. 1 - Form of matrix A.



$h_i = {}^1/d_i$ $\quad$ $h_i' = {}^1/d_i'$ $\quad$ $h_i'' = {}^1/d_i''$

i' is the equation preceding i

i'' is the equation with its diagonal directly above α_{3i}

For 3 dimensions there are an extra two bands, α_5 and α_6

i''' is the equation with its diagonal directly above α_{5i}

Fig. 2 - Form of LDU decomposition.

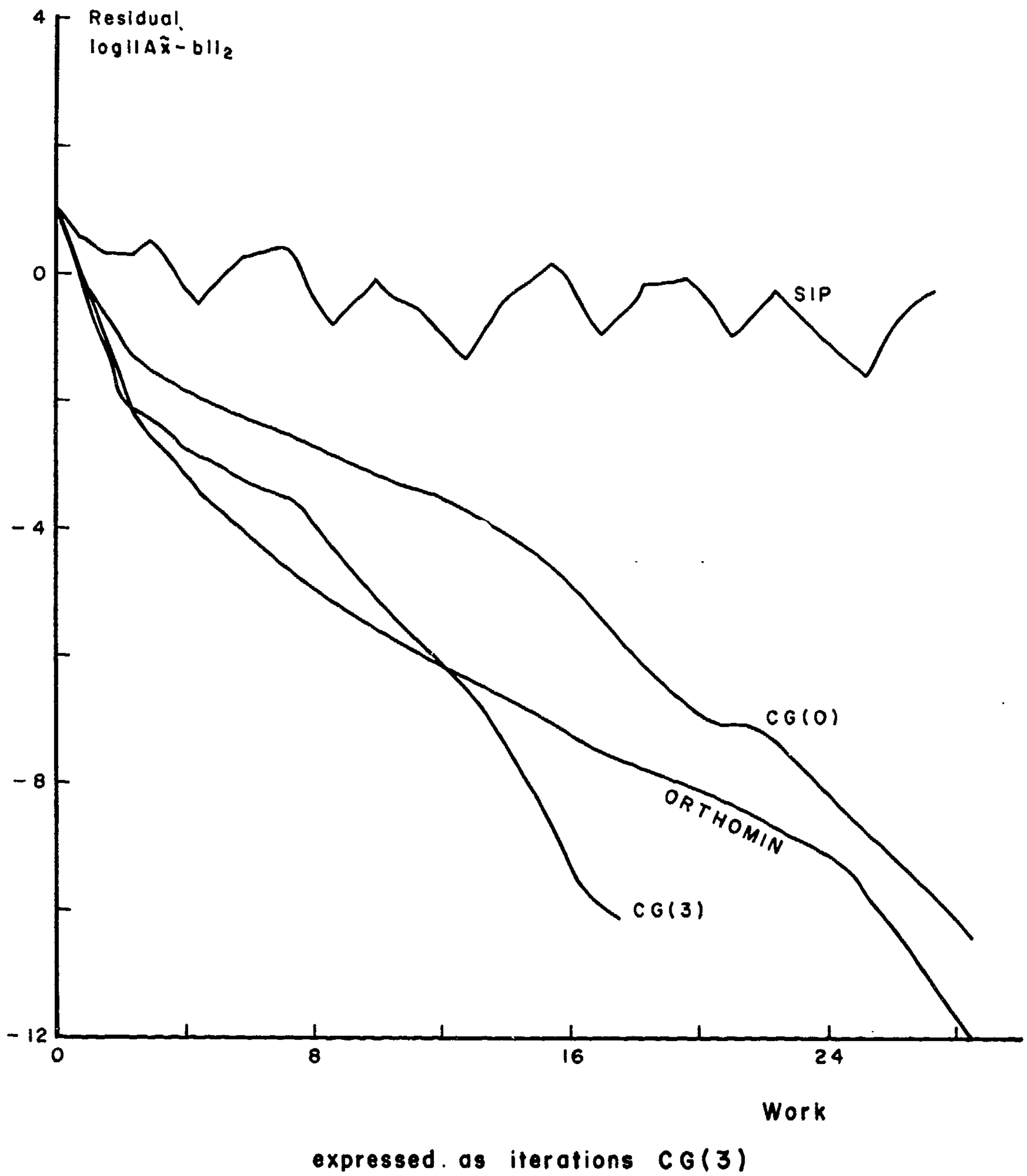


Fig. 3 - Solution of problem 1.

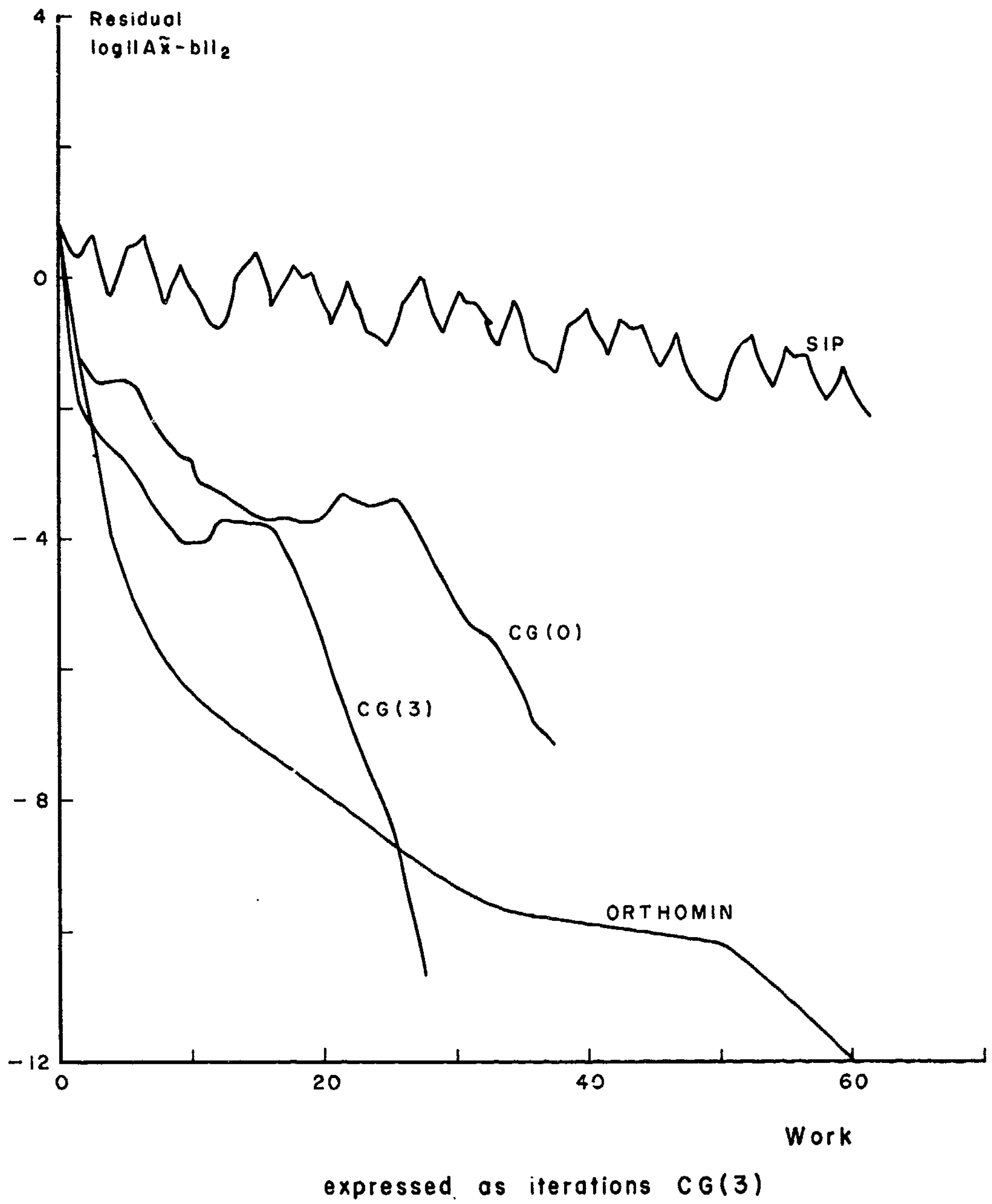


Fig. 4 - Solution of problem 2.

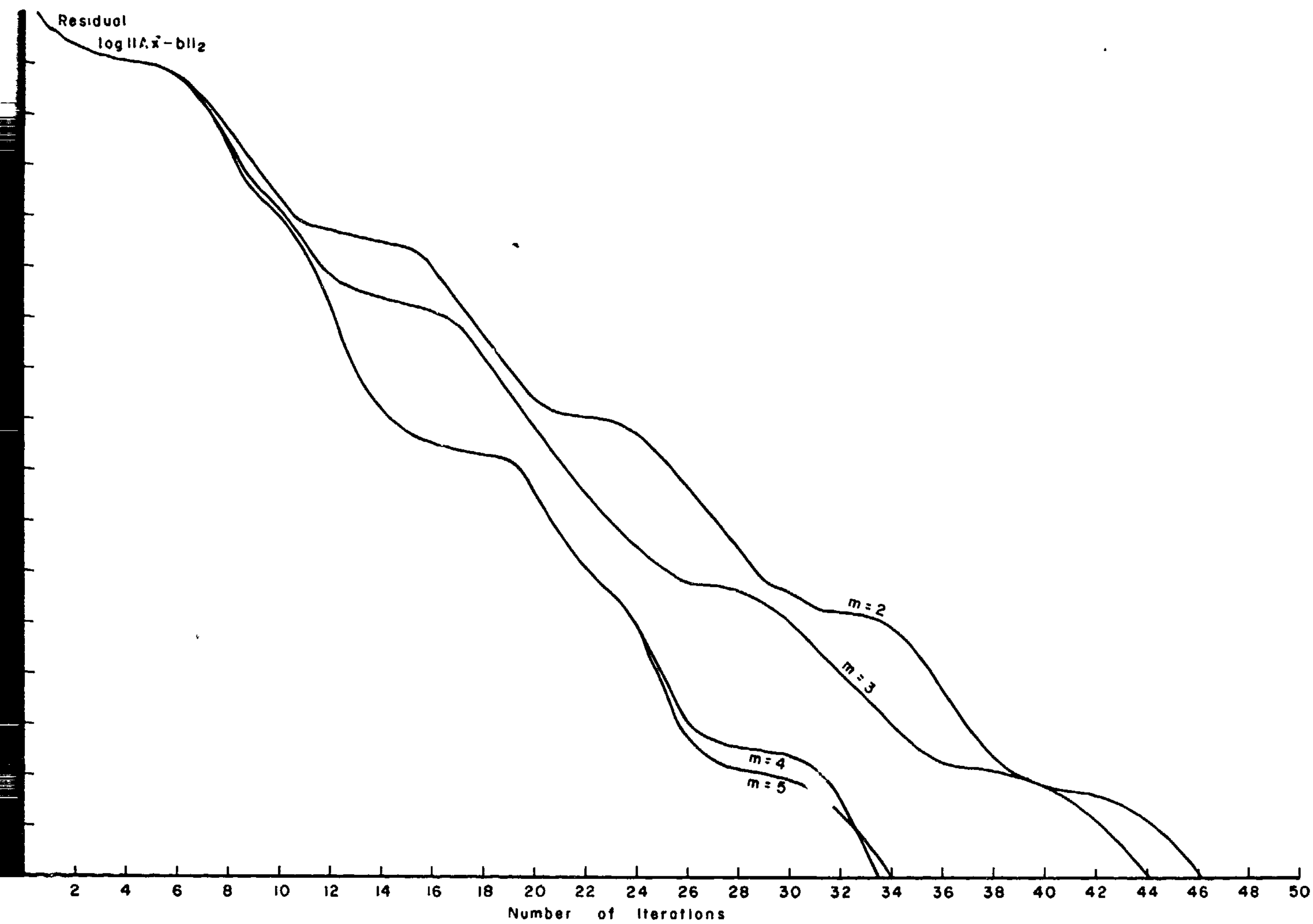


Fig. 5 - Effect of number orthogonalizations on convergence rate.

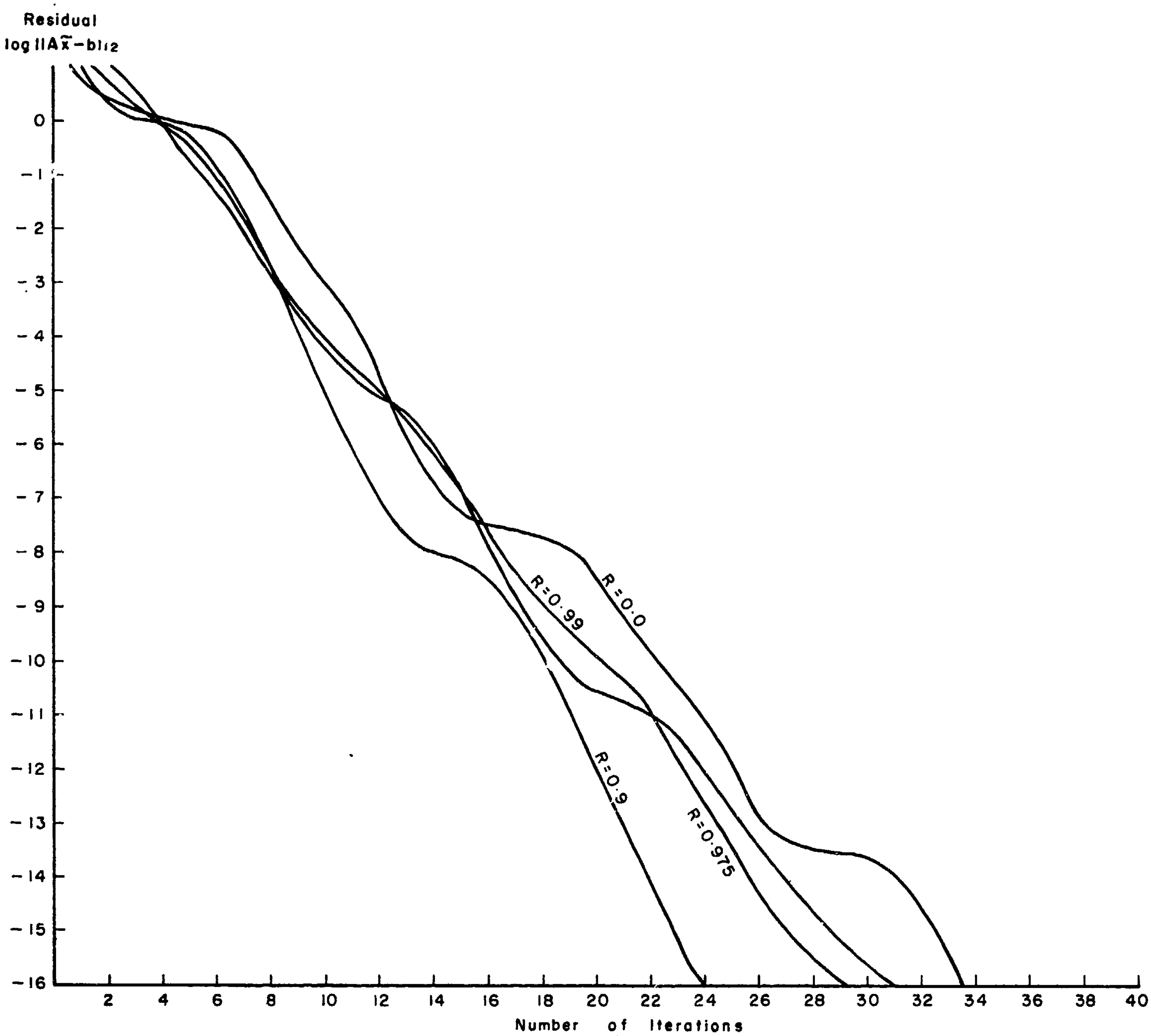


Fig. 6 - Effect of iteration parameter on convergence rate.

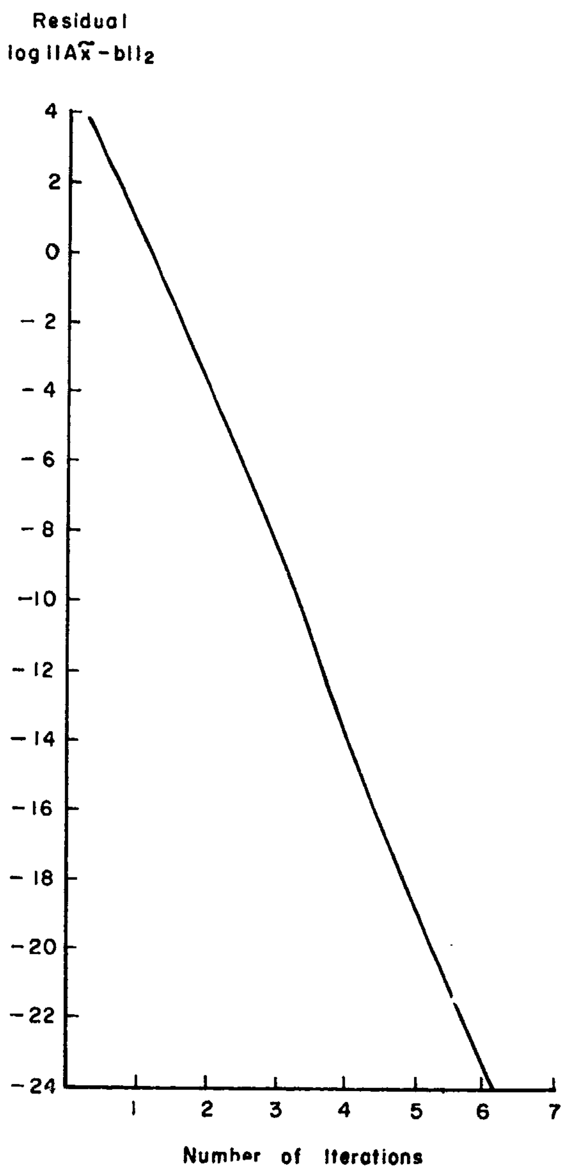


Fig. 7 - Solution of strongly non-symmetric matrix.