Presented at the Elliptic Problem Solvers Conference, Naval Postgraduate School, Monterey, CA, January 1983.

Abstract

We give an overview of recent developments in iterative methods for solving the sparse nonsymmetric linear systems that arise from the discretization of non-self-adjoint elliptic problems. We consider Krylov subspace methods, including conjugate gradient-like methods, the adaptive Chebyshev method, and a new hybrid Chebyshev/gradient technique. We discuss several preconditionings that can be used with all these methods, and compare the performance of the methods and preconditionings in solving a model elliptic problem.

Iterative Methods for Non-Self-Adjoint Elliptic Problems

Howard C. Elman

Technical Report 266

April 1983

This work was supported by the U. S. Office of Naval Research under grant N00014-82-K-0184.

1. Introduction

The numerical solution of non-self-adjoint elliptic boundary value problems requires the solution of systems of linear equations of the form

$$Ax = b, \tag{1}$$

where A is a large sparse nonsingular nonsymmetric matrix of order N. In most cases, direct methods for solving (1) entail an LU-decomposition of A, which can be prohibitively expensive for large problems. Iterative methods have low storage requirements, but must converge rapidly to be effective. Very effective iterative methods exist for solving the symmetric positive-definite linear systems arising from the discretization of self-adjoint elliptic problems (see e.g. [12]), but little of the convergence theory carries over to the nonsymmetric case. In this paper, we discuss recent progress in the development of iterative methods for solving nonsymmetric systems.

Progress has been made along two directions:

1. *basic iterative methods,* i.e. techniques for computing the i'th approximate solution x_i from previously computed quantities; and
2. *preconditioning techniques,* in which an auxiliary operator Q is used to generate an equivalent linear system such as

   $$AQ^{-1}\tilde{x} = b, \quad x = Q^{-1}\tilde{x}$$

   that is in some sense easier to solve.

In general, these two approaches are independent: any basic iterative method can be combined with any preconditioning technique.

Most of the effort in basic iterative methods has been directed toward Krylov subspace methods, in which x_i is chosen from the translated Krylov space $x_0 + S_i$, where

$$S_i = \text{span}\{r_0, Ar_0, \ldots, A^{i-1}r_0\},$$

and $r_0 = b - Ax_0$. One approach has been to generalize the conjugate gradient method (CG) [13], producing a method that needs no a priori parameter estimates. One such method, which depends explicitly on preconditioning by the symmetric part of A, was proposed by Concus and Golub [2] and Widlund [24]. More flexible methods include a collection of CG-like methods based on a truncated or restarted Gram-Schmidt-like computation, developed by several authors [1, 5, 6, 18, 23, 25, 26]; and the biconjugate gradient method, originally proposed by Fletcher [10] for symmetric indefinite systems (see [6, 12] for other references). A second important approach is Manteuffel's adaptive Chebyshev method [14, 15], which inherits certain asymptotic properties from the Chebyshev polynomials but requires estimates for the eigenvalues of A.

Most of the preconditionings used for nonsymmetric systems are straightforward generalizations of techniques for symmetric systems. For general matrices, they include incomplete factorizations, such as those developed by Dupont, Kendall and Rachford [4], Gustafsson [11], and Meijerink and van der Vorst [16]. Special techniques applicable to

discretized elliptic problems include partial elimination (construction of a "reduced system;" see [12]) and fast direct methods [3, 6].

In this paper, we describe what we regard as the most salient features of the basic iterative methods, present some recent enhancements, and briefly discuss some preconditioning techniques. In Section 2, we describe representative examples and properties of the Krylov subspace methods. In Section 3, we discuss a new hybrid Chebyshev/CG-like method recently developed by Elman, Saad and Saylor [7]. In Section 4, we discuss the incomplete factorization and reduced system preconditioning techniques. In Section 5, we present some numerical experiments in which the methods discussed are used to solve a two-dimensional non-self-adjoint elliptic boundary value problem, and in Section 6, we present some open questions.

2. Krylov Subspace Methods

In this section, we consider three Krylov subspace methods: two different generalizations of the conjugate gradient method, and the adaptive Chebyshev method.

For symmetric positive-definite systems, the conjugate gradient method computes the unique point in $x_0 + S_i$ whose A-norm $(x\text{-}x_i, A(x\text{-}x_i))^{1/2}$ is smallest. We refer to this as a "global" property of CG: at step i, it solves an i-dimensional minimization problem. CG accomplishes this with a small amount of work per step (5N multiplications plus a matrix-vector product) but without a priori parameter (e.g. eigenvalue) estimates.

These properties of rapid convergence, low cost per step, and independence of parameter estimates have led to efforts to generalize CG. One type of generalization sacrifices global properties to retain low cost per step. By way of introduction, consider the following method [5, 6]:

Algorithm 1: The generalized conjugate residual method (GCR).

Choose x_0. Compute $r_0 = b\text{-}Ax_0$. Set $p_0 = r_0$.

FOR i=0 STEP 1 UNTIL Convergence DO

$$\alpha_i = (r_i, Ap_i)/(Ap_i, Ap_i),$$

$$x_{i+1} = x_i + \alpha_i p_i,$$

$$r_{i+1} = r_i - \alpha_i Ap_i,$$

$$b_j^{(i)} = -(Ar_{i+1}, Ap_j)/(Ap_j, Ap_j),$$

$$p_{i+1} = r_{i+1} + \sum_{j=0}^{i} b_j^{(i)} p_j.$$

This method is convergent if the symmetric part $(A+A^T)/2$ is positive-definite. The directions $\{p_i\}$ are constructed to be A^TA-orthogonal, and as a result, the choice of the steplength α_i is such that x_{i+1} is the point in $x_0 + S_{i+1}$ with smallest residual norm $\|r_{i+1}\|_2$. Thus, GCR is optimal with respect to the residual norm. Unlike CG, however, GCR achieves global optimality at a

high cost: the recurrence for p_{i+1} requires O((i+1)N) operations and storage. Though optimal, GCR is not practical if more than a few (say, twenty) iterations are required.

Indeed, as Faber and Manteuffel [9] have recently shown, except in special cases there are no optimal CG-like methods with short recurrences for nonsymmetric matrices. However, nonoptimal but effective techniques that do not require parameter estimates have been developed by forcibly limiting the cost per step. For example, the direction update can be truncated so that most $k \ll N$ previous directions are used after iteration k:

$$p_{i+1} = r_{i+1} + \sum_{j=i-k+1}^{i} b_j^{(i)} p_j.$$

In this case, the iterate x_{i+1} is a *local* minimum, the point in

$$x_{i-k} + \text{span}\{p_{i-k},\ldots,p_i\}$$

whose residual norm $\|r_{i+1}\|_2$ is minimized. Alternatively, with a maximum of k directions saved, GCR can be restarted every k+1 steps. We refer to the truncated and restarted versions of GCR as Orthomin(k) and GCR(k) respectively. One matrix-vector product is required at each step (see Table 1). Let M denote the symmetric part of A, and R the skew-symmetric part $(A-A^T)/2$. When M is positive-definite, the residuals generated by both methods satisfy [5, 6]

$$\|r_i\|_2 \leq \left[1 - \frac{\lambda_{\min}(M)^2}{\lambda_{\min}(M)\lambda_{\max}(M)+\rho(R)^2}\right]^{i/2}\|r_0\|_2, \tag{2}$$

where $\lambda_{\min}(M)$ and $\lambda_{\max}(M)$ denote the smallest and largest eigenvalues of M, respectively, and $\rho(R)$ denotes the spectral radius of R. Hence, both methods are convergent. In practice, we have found this bound to be pessimistic and these methods to be effective solution techniques (see [6] and Section 5).

A second generalization of CG is the biconjugate gradient method (BCG) [10], which attempts to impose a global condition with short recurrences.

Algorithm 2: The biconjugate gradient method.

Choose x_0. Compute $r_0 = b - Ax_0$. Set $p_0 = r_0$, $\bar{r}_0 = r_0$, $\bar{p}_0 = p_0$.

FOR i=0 STEP 1 UNTIL Convergence DO

$$\alpha_i = (\bar{r}_i, r_i)/(\bar{p}_i, Ap_i),$$
$$x_{i+1} = x_i + \alpha_i p_i,$$
$$r_{i+1} = r_i - \alpha_i A p_i, \qquad \bar{r}_{i+1} = \bar{r}_i - \alpha_i A^T \bar{p}_i,$$
$$\beta_{i+1} = (\bar{r}_{i+1}, r_{i+1})/(\bar{r}_i, r_i),$$
$$p_{i+1} = r_{i+1} + \beta_i p_i, \qquad \bar{p}_{i+1} = \bar{r}_{i+1} + \beta_i \bar{p}_i.$$

The introduction of the auxiliary vectors $\{\bar{r}_i\}$ and $\{\bar{p}_i\}$ results in the construction of the auxiliary Krylov space

$$\bar{S}_i = \text{span}\{r_0, A^T r_0, \ldots, (A^T)^{i-1} r_0\}.$$

The choice of the scalars $\{\alpha_i\}$ and $\{\beta_i\}$ establish the global Galerkin conditions

$$(r_i,v) = 0, \quad v \in \overline{S}_i, \qquad (\bar{r}_i,v) = 0, \quad v \in S_i.$$

However, these relations are imposed at the cost of an extra matrix-vector product, by A^T, at each step (see Table 1). Morever, although BCG can in principle be applied to any nonsingular linear system, little is known about convergence. We know of no error norm that is monotonically decreasing, and although the method will terminate after at most N iterations, it may break down with α_i=0 before the solution is found. (See [12, 19].)

A contrasting approach is Manteuffel's adaptive Chebyshev method [14, 15], which displays rapid convergence at low cost per step but is highly dependent on parameter estimates. This method generates a sequence of iterates $\{x_i\}$ whose residuals satisfy

$$r_i = C_i(A)r_0, \tag{3}$$

where C_i is a scaled, translated i'th-degree Chebyshev polynomial:

$$C_i(z) = T_i\left(\frac{d-z}{c}\right) \Big/ T_i\left(\frac{d}{c}\right). \tag{4}$$

The scaling and translation depend on locating an ellipse E that is small in a certain sense which encloses the spectrum of A but whose interior does contain the origin; d is the center of E, and $d\pm c$ are its foci. Rapid convergence follows from the asymptotic optimality in L_∞ of the Chebyshev polynomials on ellipses in the complex plane [15]. Moreover, the three-term recurrence for Chebyshev polynomials induces a three-term recurrence

$$x_{i+1} = x_i + \alpha_i r_i + \beta_i(x_i - x_{i-1}), \tag{5}$$

in which the scalars α_i, β_i are computed with scalar arithmetic, so that the work per step is lower than for the CG-like methods. (See Table 1.)

The drawback is the explicit dependence on the ellipse parameters c and d, which are computed from the extreme eigenvalues of A. Starting with (possibly arbitrary) initial values for c, d, the adaptive Chebyshev method estimates these eigenvalues dynamically for systems where the symmetric part is positive-definite [14]. With given values for c, d, the generated residuals satisfy

$$r_i \approx S(A)^i r_0, \qquad i \to \infty, \tag{6}$$

where S(A) is the linear operator induced by

$$S(z) = S_{c,d}(z) = \frac{d-z + [(d-z)^2 - c^2]^{1/2}}{d + [d^2 - c^2]^{1/2}}.$$

That is, the residuals resemble the vectors generated by the power method for S(A). If some eigenvalue of S(A) has modulus greater than one and r_0 has a component in the corresponding eigenvector, then the residuals will diverge but will eventually become rich in that eigenvector. In this case, the residuals can be used to compute estimates for the eigenvalues of S(A). These are then used to compute eigenvalue estimates for A, which in turn are used to compute new

ellipse parameters. The Chebyshev iteration is then restarted with the new ellipse parameters. Note that since the eigenvalue computation is based on a variant of the power method, this computation will actually be facilitated if the residuals are allowed to diverge. Numerical experiments [6] indicate that the norms of the residuals generated by the adaptive Chebyshev method may increase by several orders of magnitude before good parameters are obtained, but that convergence is rapid afterwards (see Section 5).

	Orthomin(k)	GCR(k)	BCG	Chebyshev[1]
Work/Loop	(3k+4)N, Av	((3/2)k+4)N, Av	7N, Av, A^Tv	2N, Av
Storage	(2k+4)N	(2k+4)N	6N	8N

Table 1: Work and storage costs.

Table 1 contains the work (number of floating point multiplications and divisions) and storage costs (not including storage for A, b) of the methods of this section. Matrix-vector products are counted separately: Av denotes a matrix multiplication by A and A^Tv denotes a matrix multiplication by A^T.

3. A Hybrid Method

A method that attempts to combine two of the approaches of the previous section is the hybrid Chebyshev/gradient method recently devised by Elman, Saad and Saylor [7]. It uses the basic Chebyshev iteration (5) but replaces the Manteuffel adaptive procedure with a Lanczos-type method that can be used to simultaneously estimate eigenvalues and improve the solution iterate. Again, it is applicable to systems with positive-definite symmetric part.

In the hybrid method, the eigenvalue computation is performed by Arnoldi's method [20], which estimates the eigenvalues of A by reducing it to upper-Hessenberg form. Starting with an initial vector v_1 with $\|v_1\|_2=1$, this method generates an orthonormal basis for the Krylov space $\text{span}\{v_1, Av_1, \ldots, A^{k-1}v_1\}$ by an iteration of the form

$$h_{i+1,i}v_{i+1} = Av_i - \sum_{j=1}^{i} h_{ji}v_j.$$

The orthonormal matrix $V_k = [v_1,\ldots,v_k]$ and the k'th order upper-Hessenberg matrix $H_k = [h_{ij}]$ satisfy $V_k^TAV_k = H_k$. Taking this relation as an approximate similarity transformation, Arnoldi's method uses the eigenvalues of H_k as estimates for the eigenvalues of A.

Information obtained from Arnoldi's method can be used as the basis of an iterative method

[1]The overhead of 14N operations for the adaptive step is not included.

for solving (1). Given some initial guess, x_0, let v_1 denote the normalized residual $r_0/\|r_0\|_2$. Then the k Arnoldi vectors $\{v_i\}_{i=1}^k$ can be used to compute the point

$$\tilde{x} \in x_0 + \text{span}\{r_0, Ar_0, \ldots, A^{k-1}r_0\}$$

whose residual norm $\|b\text{-}A\tilde{x}\|_2$ is minimized. Let $\bar{H}_k$ denote the (k+1)xk matrix obtained by appending to H_k a row with single nonzero entry $h_{k+1,k}$ in column k. Then $\tilde{x}$ is given by

$$\tilde{x} = \sum_{i=1}^{k} \alpha_i v_i,$$

where the scalars $\{\alpha_i\}_{i=1}^k$ solve the least squares problem

$$\min \| \bar{H}_k \alpha - \|r_0\|_2 e_1 \|_2.$$

This is the generalized minimum residual (GMRES) method of Saad and Schultz [21]. It is mathematically equivalent to GCR when the symmetric part of A is positive-definite, and it is also applicable to arbitrary nonsingular matrices.

The hybrid method combines the Chebyshev method with the CG-like method GMRES, using Arnoldi's method for eigenvalue estimates. In the following simple form, the adaptive step is invoked if the residual norm generated by the Chebyshev iteration exceeds a specified tolerance τ relative to the smallest residual encountered, or after at most n Chebyshev steps. It is also used to generate initial eigenvalue estimates.[2]

Algorithm 3: The hybrid method.

Choose x_0. Compute $r_0 = b - Ax_0$.

UNTIL Convergence DO

Adaptive Step: Set v_1 = the current normalized residual, perform k Arnoldi/GMRES steps, and use the new eigenvalue estimates to update (or initialize) the ellipse parameters.

Chebyshev Step: Set i_{max}=i+n.

WHILE ($\|r_i\|_2/\|r_{min}\|_2 \leq \tau$ and $i \leq i_{max}$)
Compute x_i by the Chebyshev iteration (5).

The adaptive step uses $(k^2+4k+1)N$ multiplications and k matrix-vector products Av, plus kN storage.

There are two advantages over the original adaptive Chebyshev method. First, by not using the approximate powers of S(A) (6), the hybrid method is not as dependent on divergence of the Chebyshev iteration for eigenvalue estimates. The use of Arnoldi's method allows much more stringent monitoring of the residuals [7]. Second, in addition to providing new eigenvalue

[2]Manteuffel's method for computing ellipse parameters from eigenvalue estimates is still used [15].

estimates, the adaptive step improves the quality of the approximate solution provided to the next Chebyshev run. Indeed, the GMRES steps approximately annihilate the components in the residual of the eigenvectors corresponding to the computed eigenvalue estimates [7, 21].

4. Some Preconditioning Techniques

In this section, we briefly describe some choices for the preconditioning operator Q. See [6] for more details and other preconditionings.

One effective class of preconditionings is based on approximate factorization of A. Q is the product of a lower-triangular matrix L and an upper-triangular matrix U that resemble the factors obtained from Gaussian elimination but are forced to be sparse, so that the preconditioned matrix-vector product $AQ^{-1}v$ is not too expensive. Suppose $\mathbf{Z} \subseteq \{(i,j)|1 \leq i,j \leq N\}$ is given. In the incomplete LU-factorization (ILU) [16], L_{ij} and U_{ij} are defined to be zero for $(i,j) \in \mathbf{Z}$, and $[LU]_{ij} = A_{ij}$ for $(i,j) \notin \mathbf{Z}$. In the modified incomplete factorization, (MILU) [4, 11], an attempt is made to compensate for the error caused by restricting the fill-in of the factors. $[LU]_{ij} = A_{ij}$ for $(i,j) \notin \mathbf{Z}$ and $i \neq j$, but the diagonal L_{ii} is modified so that $\sum_{j=1}^{N} (A_{ij} - [LU]_{ij}) = 0$. Note that these preconditionings can be defined for any matrix. For example, the zero pattern $\mathbf{Z}$ can consist of the set of indices for which $A_{ij}=0$. Locations of additional fill-in for matrices arising from the finite-difference discretization of elliptic problems have been considered; see e.g. [11, 16].

When five-point (in two dimensions) or seven-point (in three dimensions) finite difference schemes are used to discretize elliptic problems, the resulting matrices are two-cyclic, i.e., the rows and columns can be permuted symmetrically so that (1) has the form

$$Ax = \begin{bmatrix} D_1 & C_1 \\ C_2 & D_2 \end{bmatrix} \begin{bmatrix} x^{(1)} \\ x^{(2)} \end{bmatrix} = \begin{bmatrix} b^{(1)} \\ b^{(2)} \end{bmatrix},$$

where D_1 and D_2 are diagonal matrices. If

$$Q = \begin{bmatrix} D_1^{-1} & 0 \\ -C_2 D_1^{-1} & I \end{bmatrix},$$

then the left-preconditioned system has the form

$$Q^{-1}Ax = \begin{bmatrix} I & D_1^{-1}C_1 \\ 0 & D_2 - C_2 D_1^{-1} C_1 \end{bmatrix} \begin{bmatrix} x^{(1)} \\ x^{(2)} \end{bmatrix} = \begin{bmatrix} D_1^{-1} b^{(1)} \\ b^{(2)} - C_2 D_1^{-1} b^{(1)} \end{bmatrix},$$

in which the block unknowns $x^{(1)}$ and $x^{(2)}$ are decoupled. Block $x^{(2)}$ can be computed by solving the reduced system

$$[D_2 - C_2 D_1^{-1} C_1]\, x^{(2)} = b^{(2)} - C_2 D_1^{-1} b^{(1)} \tag{7}$$

whose order is typically approximately N/2. Block x_1 is then given by $D_1^{-1}[b^{(1)} - C_1 x^{(2)}]$.

Note that with this reduced system preconditioning, the coefficient matrix and right hand side of (7) are formed explicitly. This enables (7) to also be preconditioned by, for example, an incomplete factorization. If no second preconditioning is used, then the reduced system need not be formed.

5. Numerical Experiments

In this section, we describe some numerical experiments for solving the two-dimensional elliptic partial differential equation

$$-(e^{-xy}u_x)_x - (e^{xy}u_y)_y + \gamma[(x+y)u_y + ((x+y)u)_y] + [1/(1+x+y)] = G, \tag{8}$$

with the right hand side G chosen so that the solution is

$$u(x,y) = x\, e^{xy} \sin(\pi x) \sin(\pi y).$$

We pose (8) on the unit square $\{0 \leq x,y \leq 1\}$ with homogeneous Dirichlet boundary conditions and discretize using centered finite differences on a uniform 47x47 grid, producing a linear system of order $N = 2209$.

We present two sets of experiments, corresponding to the values $\gamma=5$ and $\gamma=50$ for the scalar γ in (8). In each case, we solve (8) using Orthomin(1), BCG, adaptive Chebyshev, and the hybrid method with four Arnoldi vectors, together with the ILU preconditioning. In addition, we solve the reduced system using Orthomin(1) and the ILU preconditioning. As a benchmark, we show the performance of the conjugate gradient method applied to the normal equations of the ILU preconditioned system.[3] Figures 1 and 2 graph work vs. residual norm (on a logarithmic scale) for $\gamma=5$ and $\gamma=50$, respectively. All overhead for computing preconditionings and adapting is included in the operation counts. The tests were run on a VAX11-780 in double precision (55 bit mantissa). More extensive experiments and details on implementation are given in [6, 7].

6. Open Questions

Two open questions are:

1. how well do these methods generalize for indefinite problems;
2. can strong error bounds be found.

Most of the iterative methods that we have discussed are rigorously applicable only when the symmetric part of the coefficient matrix is positive-definite. There are two ways to handle indefinite systems. One is to choose an accurate enough preconditioning so that the preconditioned matrix AQ^{-1} has positive-definite symmetric part. We have, however,

[3]This method is equivalent to Paige and Saunders' LSQR method [17], which is more stable but requires more work per step. We also compared these two methods in these tests and found the residual norms to be nearly identical at each step.

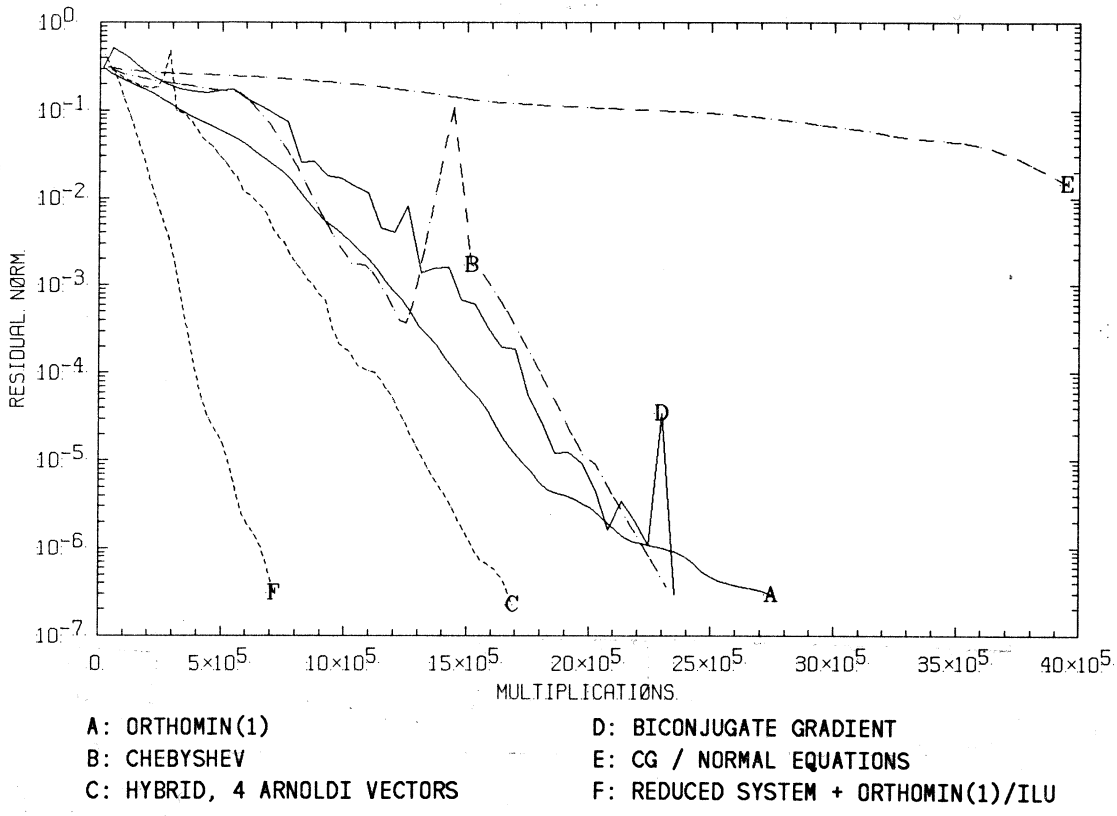

Figure 1: Residual norm vs. multiplications, γ=5, ILU preconditioning.

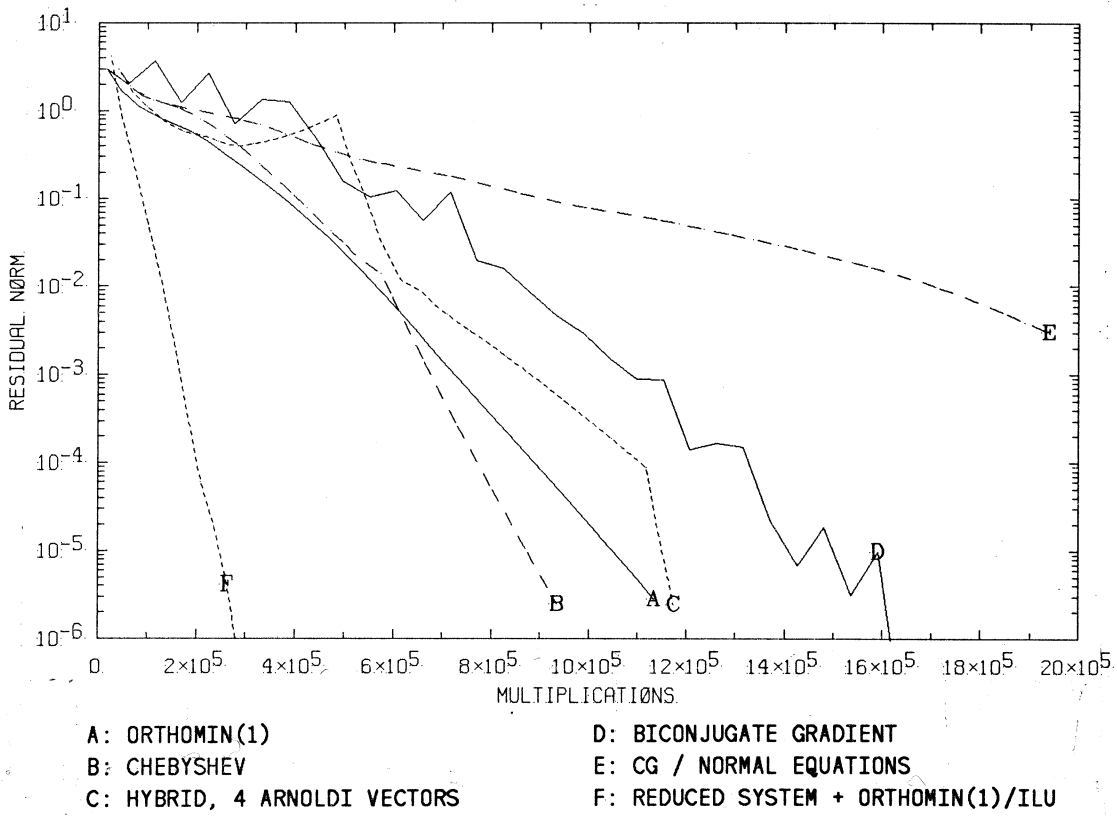

Figure 2: Residual norm vs. multiplications, γ=50, ILU preconditioning.

encountered situations where apparently effective preconditionings lead to indefinite preconditioned systems [6]. An alternative is for the basic iterative method to handle indefiniteness. Several techniques may be applicable in this domain: the GMRES method mentioned in Section 3; the incomplete orthogonalization method of Saad [18]; Orthodir, proposed by Young and Jea [12, 26]; the biconjugate gradient method; a version of Broyden's method developed for sparse systems by Engleman, Strang and Bathe [8]; and a new adaptive method devised by Smolarski and Saylor [22]. None of these ideas have been stringently tested, though, and little is known about convergence rates.

Finally, the analysis of the CG-like methods [5, 6] and of the Chebyshev method [15] prove that these techniques converge, but it is weaker than the analysis of methods for symmetric positive-definite problems. We do not know whether existing error bounds, such as (2), are tight or if there are stronger ones.

References

[1] O. Axelsson. Conjugate gradient type methods for unsymmetric and inconsistent systems of linear equations. *Linear Algebra and Appl.* 29:1-16, 1980.

[2] P. Concus and G. H. Golub. A generalized conjugate gradient method for nonsymmetric systems of linear equations. In R. Glowinski and J. L. Lions, Editors, *Lecture Notes in Economics and Mathematical Systems, Vol. 134*, Springer-Verlag, Berlin, 1976, pp. 56-65.

[3] P. Concus and G. H. Golub. Use of fast direct methods for the efficient numerical solution of nonseparable elliptic equations. *SIAM J. Numer. Anal.* 10:1103-1120, 1973.

[4] T. Dupont, R. P. Kendall and H. H. Rachford Jr. An approximate factorization procedure for solving self-adjoint elliptic difference equations. *SIAM J. Numer. Anal.* 5:559-573, 1968.

[5] S. C. Eisenstat, H. C. Elman and M. H. Schultz. Variational iterative methods for nonsymmetric systems of linear equations. *SIAM J. Numer. Anal.* 20:345-357, 1983.

[6] H. C. Elman. *Iterative Methods for Large, Sparse, Nonsymmetric Systems of Linear Equations.* Ph.D. Thesis, Department of Computer Science, Yale University, 1982. Also available as Technical Report 229.

[7] H. C. Elman, Y. Saad and P. E. Saylor. *A Hybrid Chebyshev Krylov-Subspace Method for Nonsymmetric Systems of Linear Equations.* 1983. In preparation.

[8] M. S. Engleman, G. Strang and K.-J. Bathe. The application of quasi-Newton methods in fluid mechanics. *Int. J. Numer. Meth. Eng.* 17:707-718, 1981.

[9] V. Faber and T. Manteuffel. *Conjugate Gradient Methods.* Technical Report LA-UR-82-1967, Los Alamos National Laboratory, 1982.

[10] R. Fletcher. Conjugate gradient methods for indefinite systems. In G. A. Watson, Editor, *Numerical Analysis Dundee 1975*, Springer-Verlag, New York, 1976, pp. 73-89.

[11] I. Gustafsson. A class of first order factorizations. *BIT* 18:142-156, 1978.

[12] L. A. Hageman and D. M. Young. *Applied Iterative Methods.* Academic Press, New York, 1981.

[13] M. R. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear systems. *J. Research Nat. Bur. Standards* 49:409-435, 1952.

[14] T. A. Manteuffel. Adaptive procedure for estimation of parameters for the nonsymmetric Tchebychev iteration. *Numer. Math.* 31:187-208, 1978.

[15] T. A. Manteuffel. The Tchebychev iteration for nonsymmetric linear systems. *Numer. Math.* 28:307-327, 1977.

[16] J. A. Meijerink and H. A. van der Vorst. An iterative solution method for linear systems of which the coefficient matrix is a symmetric M-matrix. *Math. Comp.* 31:148-162, 1977.

[17] C. C. Paige and M. A. Sanders. LSQR: An algorithm for sparse linear equations and sparse least squares. *ACM Trans. Mathematical Software* 8:43-71, 1982.

[18] Y. Saad. Krylov subspace methods for solving large unsymmetric linear systems. *Math. Comp.* 37:105-126, 1981.

[19] Y. Saad. The Lanczos biorthogonalization algorithm and other oblique projection methods for solving large unsymmetric systems. *SIAM J. Numer. Anal.* 19:485-506, 1982.

[20] Y. Saad. Variations on Arnoldi's method for computing eigenelements of large unsymmetric matrices. *Linear Algebra and Appl.* 34:269-295, 1980.

[21] Y. Saad and M. H. Schultz. *GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems.* Technical Report 254, Yale University Department of Computer Science, 1983.

[22] D. C. Smolarski and P. E. Saylor. *Optimum Parameters for the Solution of Linear Equations by Richardson's Iteration.* May 1982. Unpublished manuscript.

[23] P. K. W. Vinsome. Orthomin, an iterative method for solving sparse sets of simultaneous linear equations. In *Proceedings of the Fourth Symposium on Reservoir Simulation,* Society of Petroleum Engineers of AIME, 1976, pp. 149-159.

[24] O. Widlund. A Lanczos method for a class of non-symmetric systems of linear equations. *SIAM J. Numer. Anal.* 15:801-812, 1978.

[25] D. M. Young. *Accelerating non-symmetrizable iterative methods.* 1983. To appear in Proceedings of the Elliptic Problem Solvers Conference held at The Naval Postgraduate School, Monterey CA, Jan. 1983.

[26] D. M. Young and K. C. Jea. Generalized conjugate gradient acceleration of nonsymmetrizable iterative methods. *Linear Algebra and Appl.* 34:159-194, 1980.