

ITERATIVE METHODS FOR NEARLY SINGULAR LINEAR SYSTEMS*

WILLIAM W. HAGER†

Abstract. Iterative methods are developed and studied for near-singular linear systems $\mathbf{C}\mathbf{x} = \mathbf{b}$. Our approach, called the transformed minimal residual algorithm (TMRES), is derived from any convergent iterative scheme $\mathbf{S}\mathbf{x}_{k+1} = \mathbf{T}\mathbf{x}_k + \mathbf{b}$ associated with a splitting $\mathbf{C} = \mathbf{S} - \mathbf{T}$. In each step of TMRES, the transformed residual $\mathbf{S}^{-1}(\mathbf{b} - \mathbf{C}\mathbf{x})$ is minimized over a Krylov space generated by $\mathbf{S}^{-1}\mathbf{T}$. The original iterative scheme typically converges slowly when \mathbf{C} is nearly singular, while a Krylov space generated by $\mathbf{S}^{-1}\mathbf{T}$ often contains a much better approximation to a solution. TMRES is algebraically equivalent to the generalized minimal residual algorithm (GMRES) preconditioned by \mathbf{S}^{-1} , although there are numerical differences since a different matrix $\mathbf{S}^{-1}\mathbf{C}$ is used to generate the Krylov space in preconditioned GMRES. Special attention is given to sparsity and convergence issues related to linear systems of the form $(\mathbf{A}\mathbf{A}^T + \sigma\mathbf{I})\mathbf{x} = \mathbf{b}$, where $\sigma \geq 0$.

Key words. singular linear system, ill-conditioned system, Krylov space, matrix splitting, preconditioning, generalized minimal residual, successive overrelaxation, Gauss–Seidel, conjugate gradients, linear programming, sparse matrices

AMS subject classifications. 65F10, 65F50, 65Y20

PII. S106482759834634X

1. Introduction. Iterative methods are developed and analyzed for singular or near-singular linear systems $\mathbf{C}\mathbf{x} = \mathbf{b}$, where \mathbf{C} is an $m \times m$ matrix, possibly nonsymmetric. In the generalized minimal residual algorithm (GMRES) of Saad and Schultz [43], the residual $\mathbf{r}(\mathbf{x}) = \mathbf{b} - \mathbf{C}\mathbf{x}$ is minimized over a Krylov space generated by the matrix \mathbf{C} . Our transformed minimal residual approach (TMRES) is based on a splitting $\mathbf{C} = \mathbf{S} - \mathbf{T}$, where \mathbf{S} is nonsingular and the spectral radius of $\mathbf{S}^{-1}\mathbf{T}$ is less than one. The transformed residual $\mathbf{S}^{-1}\mathbf{r}(\mathbf{x})$ is minimized over $\mathbf{x} = \mathbf{x}_0 + \mathbf{z}$ with \mathbf{z} in the Krylov space

$$\mathcal{K}(\mathbf{S}^{-1}\mathbf{T}, \mathbf{g}, k) = \text{span}\{\mathbf{g}, (\mathbf{S}^{-1}\mathbf{T})\mathbf{g}, (\mathbf{S}^{-1}\mathbf{T})^2\mathbf{g}, \dots, (\mathbf{S}^{-1}\mathbf{T})^{k-1}\mathbf{g}\}$$

generated by $\mathbf{S}^{-1}\mathbf{T}$ starting from the vector $\mathbf{g} = \mathbf{S}^{-1}(\mathbf{b} - \mathbf{C}\mathbf{x}_0)$. Even though the associated iterative scheme $\mathbf{S}\mathbf{x}_{k+1} = \mathbf{T}\mathbf{x}_k + \mathbf{b}$ converges slowly when \mathbf{C} is nearly singular, we observe that $\mathcal{K}(\mathbf{S}^{-1}\mathbf{T}, \mathbf{g}, k)$ often yields a good approximation to a solution of $\mathbf{C}\mathbf{x} = \mathbf{b}$ for relatively small k . Theoretically, the Krylov space generated by $\mathbf{S}^{-1}\mathbf{T}$ is the same as that generated by the GMRES preconditioned matrix $\mathbf{S}^{-1}\mathbf{C}$. Numerically, these spaces differ since the matrices $\mathbf{S}^{-1}\mathbf{T}$ and $\mathbf{S}^{-1}\mathbf{C}$ are different.

The splittings that we consider include successive overrelaxation (SOR), damped Jacobi (see Hageman and Young [22]), and a new splitting applicable to situations where the columns in a matrix are selected from the columns of a larger matrix. Some early work on preconditioned conjugate gradient methods generated by splittings was developed by Concus, Golub, and O’Leary [12, 13]. Work leading up to GMRES includes that of Paige and Saunders [40], who developed a minimum residual algorithm MINRES for symmetric systems. Other related work includes [31] and [50]. In [18]

*Received by the editors October 26, 1998; accepted for publication (in revised form) January 3, 2000; published electronically August 24, 2000. This work was supported by the National Science Foundation.

<http://www.siam.org/journals/sisc/22-2/34634.html>

†Department of Mathematics, University of Florida, Gainesville, FL 32611 (hager@math.ufl.edu, <http://www.math.ufl.edu/~hager>).

Greenbaum examines various preconditioners for the conjugate gradient method in the context of partial differential equations. In [7] Brown and Walker examine the convergence (or lack of it) for GMRES applied to singular or near-singular matrices. The use of deflation with GMRES is studied in [37].

In [4] Baglama et al. develop preconditioned restarted GMRES algorithms in which the preconditioner is generated by Sorensen's implicitly restarted Arnoldi method [46]. Their approach approximates an invariant subspace of the matrix associated with eigenvalues close to the origin. This subspace is used in a preconditioner that moves these small eigenvalues to one, leading to more rapid convergence in the GMRES algorithm. In this paper, we observe that for any convergent splitting $\mathbf{C} = \mathbf{S} - \mathbf{T}$, the eigenvalues of \mathbf{C} close to the origin are often associated with eigenvalues of $\mathbf{S}^{-1}\mathbf{T}$ of nearly largest magnitude; as a result, a good approximation to the solution of a nearly singular system $\mathbf{C}\mathbf{x} = \mathbf{b}$ is often obtained from the Krylov space $\mathcal{K}(\mathbf{S}^{-1}\mathbf{T}, \mathbf{g}, k)$ for relatively small k .

One possible application of the methods developed in this paper is to quadratic programming with a sphere constraint. Problems of this form must be solved in each iteration of the trust region algorithm [8, 10, 16, 38, 42]. The first-order optimality system for these quadratic programs leads to a linear system whose matrix becomes more singular as we increase either the radius of the sphere or the norm of the linear term in the cost function. Hence, the quadratic programs arising in trust region methods can lead to near-singular linear systems for which TMRES is well suited. Other possible application areas include homotopy continuation methods [2], nonlinear eigenvalue problems [11], and seismic inversion problems.

Throughout this paper, we illustrate convergence properties in the near-singular setting using a prototype linear system of the form

$$(1.1) \quad (\mathbf{A}\mathbf{A}^T + \sigma\mathbf{I})\mathbf{x} = \mathbf{b},$$

where $\sigma \geq 0$ and \mathbf{A} is an $m \times n$ real matrix. Systems of this form, with σ a small positive number, are solved in each iteration of the LP dual active set algorithm [23]. In interior point methods for linear programming [33, 51], each iteration involves solving a linear system with matrix of the form $\mathbf{Z}\mathbf{\Sigma}\mathbf{Z}^T$. This system has the form (1.1) when we take $\mathbf{A} = \mathbf{Z}\mathbf{\Sigma}^{1/2}$ and $\sigma = 0$. The test matrices in this paper are obtained from the LP problems in David Gay's Netlib collection (www.netlib.org/lp). These matrices, which are all sparse, can be obtained in a variety of formats through the COAP Software Forum (www.math.ufl.edu/~coap).

For symmetric linear systems of the form (1.1), the TMRES scheme with SSOR preconditioning and an SSOR preconditioned conjugate gradient scheme [6, 12, 13] exhibit similar convergence when applied to ill-conditioned matrices emanating from linear programming. In theory, these two schemes generate iterates in the same Krylov space, but they differ in the merit function used to select the approximation from the Krylov space.

To illustrate convergence problems that are encountered when solving nearly singular linear systems, we consider a netlib/lp matrix \mathbf{A} for the (small) problem *beacofd* ($m = 173$, $n = 295$). The right side \mathbf{b} was randomly generated on the unit sphere¹ in \mathbf{R}^m , $\sigma = 0$, and the columns of \mathbf{A} in (1.1) were scaled to be unit vectors. This column scaling is the one we use in the LP dual active set algorithm. In interior

¹A vector can be randomly generated on the Euclidean unit sphere by first randomly generating its components using a Gaussian distribution with mean 0 then dividing the resulting vector by its length.

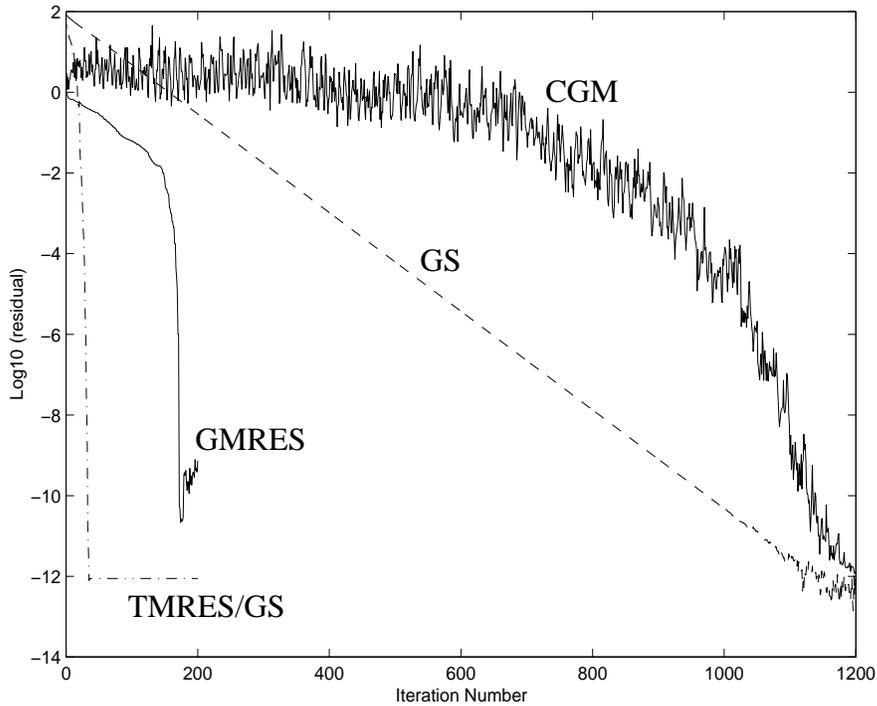


FIG. 1.1. Convergence for the conjugate-gradient method (CGM), the Gauss-Seidel method (GS), the generalized minimal residual algorithm (GMRES) without restarts, and the transformed minimal residual algorithm (TMRES) obtained from Gauss-Seidel splitting when \mathbf{A} in (1.1) is the test problem *beaconfd*.

point methods, the diagonal scaling matrix Σ becomes increasingly ill-conditioned, and the condition number of the associated $\mathbf{A}\mathbf{A}^T$ could be much larger than that for the normalized \mathbf{A} used in our experiments.

We solve the linear system (1.1) using three different iterative methods: Gauss-Seidel (GS), conjugate gradients (CGM), and GMRES. Although various Lanczos and conjugate gradient type methods [41, 44, 45] have been developed for special versions of (1.1), we do not take into account special structure for (1.1) in our computations; rather, we treat this system as having the form $\mathbf{C}\mathbf{x} = \mathbf{b}$, where \mathbf{C} is an $m \times m$ symmetric, positive definite matrix. All of the schemes, GS, CGM, and GMRES, are guaranteed to converge (see [49, p. 77], [36, Chap. 8], and [43]). Moreover, the conjugate gradient method yields the best approximation to the solution from an associated Krylov space, where the approximation quality is measured using the C -norm defined by $\|\mathbf{x}\|_C = \sqrt{\mathbf{x}^T \mathbf{C} \mathbf{x}}$. GMRES minimizes the residual norm $\|\mathbf{b} - \mathbf{C}\mathbf{x}\|$, where $\|\cdot\|$ is the Euclidean norm, over the same Krylov space. According to the convergence theory, both CGM and GMRES solve $\mathbf{C}\mathbf{x} = \mathbf{b}$ in at most m iterations, assuming exact arithmetic.

With the starting guess $\mathbf{x} = \mathbf{0}$, we plot in Figure 1.1 the iteration number versus the base 10 logarithm of the residual norm for the associated iterate. The computations were done in double precision in Matlab on a Sun workstation. For this small problem, the Gauss-Seidel and conjugate gradient methods converge slowly to the solution, taking nearly 1200 iterations. The generalized minimal residual algorithm

requires about 173 iterations, the number of rows in \mathbf{A} . Paige and Saunders' MINRES routine would probably yield results similar to GMRES because $\mathbf{C} = \mathbf{A}\mathbf{A}^T + \sigma\mathbf{I}$ is symmetric and the two routines are algebraically equivalent in this case. TMRES will be developed in sections 3 and 4. Its convergence behavior shown in Figure 1.1 for a Gauss–Seidel splitting is typical for many netlib/lp test problems.

With GMRES and TMRES, restarts may be essential to reduce the storage requirements when the matrix is nonsymmetric. In this paper, we do not analyze restarts.

2. Convergence properties. Let us study more closely the convergence behavior illustrated in Figure 1.1. The matrix associated with the test problem beaconfd, like many of the problems in netlib/lp, has some rows that are nearly linear combinations of other rows. The 10 smallest and largest eigenvalues of $\mathbf{A}\mathbf{A}^T$ are the following:

0.00000177	2.8954
0.00000415	3.5490
0.00001338	4.9565
0.00002115	5.1235
0.00002367	6.5875
0.00004273	7.1127
0.00004399	7.5612
0.00004847	10.9200
0.00005424	43.9156
0.00006267	63.2578

The matrix $\mathbf{A}\mathbf{A}^T$ is nearly singular in the sense that the ratio between the largest and the smallest eigenvalue of $\mathbf{A}\mathbf{A}^T$ is of order 10^7 , much larger than 1.

Let $\mathbf{C} = \mathbf{L} + \mathbf{U}$ be symmetric and positive definite, where

$$(2.1) \quad l_{ij} = c_{ij} \text{ and } u_{ij} = 0 \text{ for } i \geq j, \quad l_{ij} = 0 \text{ and } u_{ij} = c_{ij} \text{ for } i < j.$$

The Gauss–Seidel method for solving $\mathbf{C}\mathbf{x} = \mathbf{b}$ is given by the iteration

$$\mathbf{x}_{k+1} = -\mathbf{L}^{-1}(\mathbf{U}\mathbf{x}_k - \mathbf{b}).$$

If \mathbf{x}_* is the exact solution to $\mathbf{C}\mathbf{x} = \mathbf{b}$, then the error $\mathbf{e}_k = \mathbf{x}_k - \mathbf{x}_*$ satisfies the recurrence $\mathbf{e}_{k+1} = \mathbf{M}\mathbf{e}_k$, where $\mathbf{M} = -\mathbf{L}^{-1}\mathbf{U}$. When \mathbf{C} is positive definite, the spectral radius of \mathbf{M} , denoted $\rho(\mathbf{M})$, is strictly less than 1 [49, p. 77]. The spectral radius measures the convergence speed in the sense that the magnitude of error components associated with eigenvectors whose eigenvalue magnitudes are equal to the spectral radius is multiplied by the spectral radius in each iteration.

If \mathbf{C} is singular and \mathbf{z} is a nonzero vector such that $\mathbf{C}\mathbf{z} = \mathbf{0}$, then the relation $(\mathbf{L} + \mathbf{U})\mathbf{z} = \mathbf{0}$ implies that $\mathbf{z} = -\mathbf{L}^{-1}\mathbf{U}\mathbf{z}$, or $\mathbf{z} = \mathbf{M}\mathbf{z}$. Hence, \mathbf{z} is an eigenvector of \mathbf{M} corresponding to the eigenvalue 1, and the spectral radius of \mathbf{M} is at least 1. Consequently, the error \mathbf{e}_k in the Gauss–Seidel iteration may not tend to zero. If \mathbf{C} is nearly singular, then the distance between the eigenvalues of \mathbf{M} and 1 can be estimated using Gerschgorin's theorem.

PROPOSITION 2.1. *Let \mathbf{C}_0 be a square singular matrix, and consider the perturbed matrix $\mathbf{C} = \mathbf{C}_0 + \tau\mathbf{I}$. Let \mathbf{S} be any nonsingular matrix for which $\mathbf{S}^{-1}\mathbf{C}_0$ is diagonalizable: $\mathbf{S}^{-1}\mathbf{C}_0 = \mathbf{F}\mathbf{\Phi}\mathbf{F}^{-1}$, where $\mathbf{\Phi}$ is a diagonal matrix containing the eigenvalues. If \mathbf{T} is chosen so that $\mathbf{C} = \mathbf{S} - \mathbf{T}$ (i.e., $\mathbf{T} = \mathbf{S} - \mathbf{C}_0 - \tau\mathbf{I}$), then for τ sufficiently close*

to zero, the eigenvalues $\mu_1, \mu_2, \dots, \mu_n$ of $\mathbf{S}^{-1}\mathbf{T}$ satisfy

$$\min_i |\mu_i - 1| \leq |\tau| \|\mathbf{F}\|_1 \|\mathbf{F}^{-1}\|_1 \|\mathbf{S}^{-1}\|_1,$$

where $\|\cdot\|_1$ is the matrix 1-norm (largest absolute column sum).

Proof. Observe that

$$(2.2) \quad \mathbf{F}^{-1}(\mathbf{S}^{-1}\mathbf{T})\mathbf{F} = \mathbf{F}^{-1}(\mathbf{I} - \mathbf{S}^{-1}\mathbf{C}_0 - \tau\mathbf{S}^{-1})\mathbf{F} = \mathbf{I} - \mathbf{\Phi} - \tau\mathbf{F}^{-1}\mathbf{S}^{-1}\mathbf{F}.$$

Since \mathbf{C}_0 is singular, one of the diagonal elements of $\mathbf{\Phi}$ is zero. Hence, one of the diagonal elements of $\mathbf{I} - \mathbf{\Phi}$ is 1. Let δ be the minimum absolute difference between distinct diagonal elements of $\mathbf{\Phi}$ and choose τ close enough to zero that

$$|\tau| \|\mathbf{F}\|_1 \|\mathbf{F}^{-1}\|_1 \|\mathbf{S}^{-1}\|_1 < \delta.$$

By Gerschgorin's theorem and the fact that $\mathbf{S}^{-1}\mathbf{T}$ and $\mathbf{F}^{-1}(\mathbf{S}^{-1}\mathbf{T})\mathbf{F}$ are similar, it follows from (2.2) that $\mathbf{S}^{-1}\mathbf{T}$ has an eigenvalue in the sphere in the complex plane with center 1 and radius $|\tau| \|\mathbf{F}\|_1 \|\mathbf{F}^{-1}\|_1 \|\mathbf{S}^{-1}\|_1$. This completes the proof. \square

Loosely speaking, eigenvalues of \mathbf{C} near 0 correspond to eigenvalues of $\mathbf{M} = \mathbf{S}^{-1}\mathbf{T}$ near 1 and to slow convergence of the Gauss-Seidel method.

A number of papers [32, 35, 47] have appeared in the literature relating the spectrum of the conjugate gradient method to its convergence. More recent work is surveyed in [1]. Here we focus on the effect of near singularity on convergence. The following result is the key to understanding the behavior for CGM depicted in Figure 1.1.

PROPOSITION 2.2. *If \mathbf{C} is a symmetric matrix, then as σ tends to zero, we have*

$$\sigma(\sigma\mathbf{I} + \mathbf{C})^{-1} = \mathbf{P} + O(\sigma),$$

where \mathbf{P} denotes the orthogonal projection into the null space of \mathbf{C} and $O(\sigma)$ denotes a term that can be bounded in magnitude by a constant times $|\sigma|$.

Proof. Let $\mathbf{C} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$ denote the diagonalization of \mathbf{C} , where \mathbf{Q} is the orthogonal matrix of eigenvectors and $\mathbf{\Lambda}$ is a diagonal matrix whose diagonal elements $\lambda_1, \lambda_2, \dots, \lambda_m$ are the eigenvalues. We assume that the eigenvalues are ordered so that

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_k| > 0 \quad \text{and} \quad \lambda_i = 0 \text{ for } i > k.$$

Hence,

$$(2.3) \quad \sigma(\sigma\mathbf{I} + \mathbf{C})^{-1} = \sigma\mathbf{Q}(\mathbf{\Lambda} + \sigma\mathbf{I})^{-1}\mathbf{Q}^T$$

and

$$(2.4) \quad \sigma(\mathbf{\Lambda} + \sigma\mathbf{I})^{-1} = \mathbf{J} + O(\sigma),$$

where

$$\mathbf{J} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}.$$

Above, the lower right \mathbf{I} is $(m - k) \times (m - k)$, and $\mathbf{0}$ stands for a block of zeros. Since $\mathbf{Q}\mathbf{J}\mathbf{Q}^T = \mathbf{P}$, (2.3) and (2.4) complete the proof. \square

The CGM applied to the linear system $\mathbf{C}\mathbf{x} = \mathbf{b}$, where \mathbf{C} is symmetric and positive definite, is the following [25, p. 340]: $\mathbf{d}_0 = \mathbf{r}_0 = \mathbf{b} - \mathbf{C}\mathbf{x}_0$ and for $k \geq 0$,

$$\begin{aligned} \mathbf{p}_k &\leftarrow \mathbf{C}\mathbf{d}_k, \\ \mathbf{x}_{k+1} &\leftarrow \mathbf{x}_k + \alpha_k \mathbf{d}_k, \quad \text{where } \alpha_k = \mathbf{r}_k^\top \mathbf{r}_k / \mathbf{d}_k^\top \mathbf{p}_k, \\ \mathbf{r}_{k+1} &\leftarrow \mathbf{r}_k - \alpha_k \mathbf{p}_k, \\ \mathbf{d}_{k+1} &\leftarrow \mathbf{r}_{k+1} + \beta_k \mathbf{d}_k, \quad \text{where } \beta_k = \mathbf{r}_{k+1}^\top \mathbf{r}_{k+1} / \mathbf{r}_k^\top \mathbf{r}_k. \end{aligned}$$

In each step of the conjugate gradient method, the current direction \mathbf{d}_k is multiplied by \mathbf{C} in the process of obtaining the new direction \mathbf{d}_{k+1} . By induction, $\mathbf{x}_k = \mathbf{x}_0 + \mathbf{z}_k$, where \mathbf{z}_k lies in the Krylov space $\mathcal{K}(\mathbf{C}, \mathbf{r}_0, k)$.

Let $\mathbf{x}_* = \mathbf{C}^{-1}\mathbf{b}$ denote the solution to $\mathbf{C}\mathbf{x} = \mathbf{b}$. The conjugate gradient method is designed to minimize $\|\mathbf{x}_k - \mathbf{x}_*\|_C$ over $\mathbf{x}_k = \mathbf{x}_0 + \mathbf{z}_k$ with $\mathbf{z}_k \in \mathcal{K}(\mathbf{C}, \mathbf{r}_0, k)$. Observe that

$$\|\mathbf{x} - \mathbf{x}_*\|_C = \sqrt{(\mathbf{x} - \mathbf{x}_*)^\top \mathbf{C} (\mathbf{x} - \mathbf{x}_*)} = \sqrt{(\mathbf{C}\mathbf{x} - \mathbf{b})^\top \mathbf{C}^{-1} (\mathbf{C}\mathbf{x} - \mathbf{b})} = \sqrt{\mathbf{r}(\mathbf{x})^\top \mathbf{C}^{-1} \mathbf{r}(\mathbf{x})},$$

where $\mathbf{r}(\mathbf{x}) = \mathbf{b} - \mathbf{C}\mathbf{x}$. If $(\lambda_i, \mathbf{q}_i)$, $i = 1, \dots, m$, are orthonormal eigenpairs of \mathbf{C} , then

$$(2.5) \quad \|\mathbf{x} - \mathbf{x}_*\|_C^2 = \sum_{i=1}^m c_i(\mathbf{x})^2 / \lambda_i, \quad \text{where } c_i(\mathbf{x}) = \mathbf{r}(\mathbf{x})^\top \mathbf{q}_i.$$

Since the reciprocal of the eigenvalues appears in this expression, the components c_i of the residual \mathbf{r} associated with the smallest eigenvalues are amplified the most in (2.5).

Suppose that we apply CGM to a linear system of the form (1.1), where the rows of \mathbf{A} are linearly dependent and σ is small. By Proposition 2.2, the solution $\mathbf{x} \approx \mathbf{P}\mathbf{b}/\sigma$, the projection of \mathbf{b} into the null space of \mathbf{A}^\top , divided by σ . On the other hand, the k th conjugate gradient iterate lies in the Krylov space $\mathcal{K}(\mathbf{C}, \mathbf{b}, k)$ if $\mathbf{x}_0 = \mathbf{0}$. The vectors forming this space are all contained in the range of \mathbf{C} , and if σ is small, then each of these vectors is nearly in the range of \mathbf{A} . We saw in (2.5) that CGM attaches the greatest weight to those components of the error associated with the smallest eigenvalues. Hence, the conjugate gradient method is using vectors that nearly lie in the range of \mathbf{A} to approximate solution components that are nearly orthogonal to the range. In theory, CGM is guaranteed to reach the solution in a finite number of steps. This convergence must involve the subtraction of nearly equal numbers followed by the division of numbers that are nearly zero.

For the example depicted in Figure 1.1 with $\sigma = 0$, the effect is very similar. The rows of \mathbf{A} are nearly dependent, the solution nearly lies in the space spanned by eigenvectors associated with the smallest eigenvalues of $\mathbf{A}\mathbf{A}^\top$, while the conjugate gradient iterates nearly lie in the space spanned by eigenvectors associated with the largest eigenvalues of $\mathbf{A}\mathbf{A}^\top$. For a nice survey of results concerning the convergence of the conjugate gradient method in finite precision arithmetic, see the paper [20] of Greenbaum and Strakoš. There they observe that the finite precision conjugate gradient method behaves similarly to the exact algorithm applied to a matrix with nearby eigenvalues.

In the generalized minimal residual algorithm applied to $\mathbf{C}\mathbf{x} = \mathbf{b}$, the k th approximation \mathbf{x}_k to the solution is obtained by solving the problem

$$(2.6) \quad \min \|\mathbf{r}(\mathbf{x})\| \quad \text{subject to } \mathbf{x} = \mathbf{x}_0 + \mathbf{z}, \quad \mathbf{z} \in \mathcal{K}(\mathbf{C}, \mathbf{r}_0, k).$$

Thus when \mathbf{C} is symmetric, the GMRES iterates are contained in the same Krylov space used by the conjugate gradient method. Unlike CGM, GMRES does not weight the components of the residual in (2.6). As k increases, the components of the vector $\mathbf{C}^k \mathbf{r}_0$ associated with the large eigenvalues of \mathbf{C} become much larger in magnitude than the components associated with the small eigenvalues. Hence, we expect that GMRES approximates the solution components associated with the largest eigenvalues first. As the number of iterations approaches the dimension of the matrix, GMRES generates vectors orthogonal to the eigenvectors associated with the large eigenvalues. These orthogonal vectors correspond to the space associated with the eigenvectors of the smallest eigenvalues. Observe in Figure 1.1 that when the approximating space approaches the space corresponding to the smallest eigenvalues, the GMRES error quickly decreases. In order to improve on GMRES in this near-singular setting, we will strive to reach this space sooner. Note that for nonsymmetric matrices, the convergence of GMRES depends on both eigenvalues and departure from normality. For example, see [9, 19, 39] for related results and discussion.

3. Krylov, Arnoldi, and TMRES. With CGM or GMRES, the approximate solution lies in a Krylov space, and an orthonormal basis for this space is generated by the Arnoldi (Gram–Schmidt) process. The following algorithm generates such an orthonormal basis $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ for the Krylov space $\mathcal{K}(\mathbf{M}, \mathbf{g}, k)$:

ALGORITHM 1 (Arnoldi).

```

 $\mathbf{v}_1 \leftarrow \mathbf{g} / \|\mathbf{g}\|$ 
for  $j = 1 : k - 1$ 
   $\mathbf{s} \leftarrow \mathbf{M}\mathbf{v}_j$ 
  for  $i = 1 : j$ 
     $h_{ij} \leftarrow \mathbf{s}^T \mathbf{v}_i \quad (= \mathbf{v}_i^T \mathbf{M}\mathbf{v}_j)$ 
     $\mathbf{s} \leftarrow \mathbf{s} - h_{ij} \mathbf{v}_i$ 
  end
   $h_{j+1,j} \leftarrow \|\mathbf{s}\|$ 
   $\mathbf{v}_{j+1} \leftarrow \mathbf{s} / h_{j+1,j}$ 
end

```

end Algorithm 1

Obviously, if $h_{k+1,k} = 0$ for some k , the Arnoldi process should stop because

$$(3.1) \quad \mathcal{K}(\mathbf{M}, \mathbf{g}, j+1) = \mathcal{K}(\mathbf{M}, \mathbf{g}, j) \quad \text{for all } j \geq k.$$

CGM or GMRES corresponds to the choice $\mathbf{M} = \mathbf{C}$.

In order to more quickly generate vectors near the space spanned by the eigenvectors of \mathbf{C} associated with the smallest eigenvalues, we consider the matrix $\mathbf{M} = \mathbf{S}^{-1}\mathbf{T}$ associated with a splitting $\mathbf{C} = \mathbf{S} - \mathbf{T}$, where \mathbf{S} is nonsingular. This splitting leads to an iterative method

$$(3.2) \quad \mathbf{S}\mathbf{x}_{k+1} = \mathbf{T}\mathbf{x}_k + \mathbf{b},$$

which converges to a solution of $\mathbf{C}\mathbf{x} = \mathbf{b}$ for all choices of the initial condition \mathbf{x}_0 if and only if $\rho(\mathbf{S}^{-1}\mathbf{T}) < 1$. The terminology “convergent splitting” will mean that $\rho(\mathbf{S}^{-1}\mathbf{T}) < 1$. From the discussion of section 2, we know that the eigenpairs of \mathbf{C} whose eigenvalues are near zero correspond to eigenpairs of $\mathbf{S}^{-1}\mathbf{T}$ whose eigenvalues are near 1. Hence, the eigenvectors of $\mathbf{S}^{-1}\mathbf{T}$ associated with its eigenvalues of largest magnitude include approximations to the eigenvectors of \mathbf{C} associated with its eigenvalues of smallest magnitude. Suppose that $\mathbf{S}^{-1}\mathbf{T}$ has m linearly independent eigenvectors $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m$ and associated eigenvalues $\mu_1, \mu_2, \dots, \mu_m$. If

$\mathbf{g} = c_1\mathbf{f}_1 + c_2\mathbf{f}_2 + \cdots + c_m\mathbf{f}_m$ is the expansion of \mathbf{g} in terms of the eigenvectors, then

$$\mathbf{M}^k\mathbf{g} = (\mathbf{S}^{-1}\mathbf{T})^k\mathbf{g} = c_1\mu_1^k\mathbf{f}_1 + c_2\mu_2^k\mathbf{f}_2 + \cdots + c_m\mu_m^k\mathbf{f}_m.$$

It follows that the components of \mathbf{g} associated with the absolute largest eigenvalues of \mathbf{M} are amplified more rapidly than the components of \mathbf{g} associated with the absolute smallest eigenvalues. Loosely speaking, the Krylov space $\mathcal{K}(\mathbf{M}, \mathbf{g}, k)$ converges, as k grows, to those eigenvectors associated with the absolute largest eigenvalues of \mathbf{M} much more quickly than to those eigenvectors associated with the absolute smallest eigenvalues of \mathbf{M} . Since those eigenvectors associated with the large eigenvalues of \mathbf{M} correspond to eigenvalues of \mathbf{C} associated with its smallest eigenvalues, we converge to that space associated with the small eigenvalues of \mathbf{C} that is critical in the near-singular setting.

After generating a Krylov space and a basis for it using Algorithm 1, we now return to the linear system $\mathbf{C}\mathbf{x} = \mathbf{b}$. If we were to minimize the norm of the residual $\mathbf{r} = \mathbf{b} - \mathbf{C}\mathbf{x}$ over the space spanned by the basis vectors \mathbf{v}_i , $1 \leq i \leq k$, we would need to multiply each \mathbf{v}_i by \mathbf{C} . To circumvent this multiplication, we minimize the transformed residual \mathbf{t} given by

$$\mathbf{t}(\mathbf{x}) = \mathbf{S}^{-1}\mathbf{r}(\mathbf{x}) = \mathbf{S}^{-1}(\mathbf{b} - \mathbf{C}\mathbf{x}) = \mathbf{S}^{-1}\mathbf{b} - (\mathbf{I} - \mathbf{M})\mathbf{x}.$$

Now when \mathbf{x} is expanded in the basis vectors \mathbf{v}_i and we minimize the norm of \mathbf{t} , we need to compute the product $\mathbf{M}\mathbf{v}_i$, which is available from Algorithm 1.

At step k , the TMRES approximation \mathbf{x}_k to $\mathbf{x} = \mathbf{C}^{-1}\mathbf{b}$ is the solution to the transformed problem

$$(3.3) \quad \min \|\mathbf{t}(\mathbf{x})\| \quad \text{subject to } \mathbf{x} = \mathbf{z} + \mathbf{x}_0, \quad \mathbf{z} \in \mathcal{K}(\mathbf{M}, \mathbf{g}, k).$$

Making the special choice $\mathbf{g} = \mathbf{S}^{-1}(\mathbf{b} - \mathbf{C}\mathbf{x}_0)$, and substituting $\mathbf{x} = \mathbf{z} + \mathbf{x}_0$, we obtain the problem

$$(3.4) \quad \min \|\mathbf{g} - (\mathbf{I} - \mathbf{M})\mathbf{z}\| \quad \text{subject to } \mathbf{z} \in \mathcal{K}(\mathbf{M}, \mathbf{g}, k).$$

If \mathbf{V}_k denotes the matrix whose columns are the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ generated by the Arnoldi process, then (3.4) reduces to the following equivalent problem:

$$(3.5) \quad \min_{\mathbf{y}} \|\mathbf{g} - (\mathbf{I} - \mathbf{M})\mathbf{V}_k\mathbf{y}\|.$$

The solution to the least-squares problem (3.5) can be obtained by a process similar to that described in [43]. If \mathbf{H} is the $(k+1) \times k$ upper Hessenberg matrix whose elements are given by the Arnoldi process, then the following relation holds:

$$(3.6) \quad \mathbf{M}\mathbf{V}_k = \mathbf{V}_{k+1}\mathbf{H}.$$

Let \mathbf{e} denote the vector whose first component is $\|\mathbf{g}\|$ and whose remaining components are zero. After substituting $\mathbf{g} = \mathbf{V}_{k+1}\mathbf{e}$ in (3.5) and utilizing (3.6), we obtain the equivalent problem

$$(3.7) \quad \min_{\mathbf{y}} \|\mathbf{V}_{k+1}\mathbf{e} - \mathbf{V}_k\mathbf{y} + \mathbf{V}_{k+1}\mathbf{H}\mathbf{y}\|.$$

Since the columns of \mathbf{V}_{k+1} are orthonormal, we have

$$\begin{aligned}\|\mathbf{V}_{k+1}\mathbf{e} - \mathbf{V}_k\mathbf{y} + \mathbf{V}_{k+1}\mathbf{H}\mathbf{y}\| &= \|\mathbf{V}_{k+1}\mathbf{e} - \mathbf{V}_{k+1}\begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix} + \mathbf{V}_{k+1}\mathbf{H}\mathbf{y}\| \\ &= \|\mathbf{V}_{k+1}\mathbf{e} + \mathbf{V}_{k+1}\left(\mathbf{H} - \begin{bmatrix} \mathbf{I} \\ \mathbf{0} \end{bmatrix}\right)\mathbf{y}\| \\ &= \|\mathbf{e} + \bar{\mathbf{H}}\mathbf{y}\|,\end{aligned}$$

where $\bar{\mathbf{H}}$ is the same as the Hessenberg matrix \mathbf{H} generated by Algorithm 1 except that 1's have been subtracted from the diagonal elements. With this substitution, (3.7) simplifies to

$$(3.8) \quad \min_{\mathbf{y}} \|\mathbf{e} + \bar{\mathbf{H}}\mathbf{y}\|.$$

Finally, the vector $\mathbf{e} + \bar{\mathbf{H}}\mathbf{y}$ is multiplied on the left by a series of Givens rotations reducing $\bar{\mathbf{H}}$ to triangular form \mathbf{R} and reducing \mathbf{e} to a vector \mathbf{f} . Our approximation to the solution of $\mathbf{C}\mathbf{x} = \mathbf{b}$ is $\mathbf{x}_k = \mathbf{x}_0 + \mathbf{V}_k\mathbf{y}$, where \mathbf{y} is the solution of the triangular system $(\mathbf{R}\mathbf{y})_{1:k} = \mathbf{f}_{1:k}$. The minimum norm in (3.8) is $|f_{k+1}|$.

In the following statement of TMRES, we overwrite \mathbf{H} and \mathbf{e} with \mathbf{R} and \mathbf{f} , respectively. This algorithm is basically the same as preconditioned GMRES; in GMRES preconditioned by \mathbf{S}^{-1} , the multiplication by $\mathbf{S}^{-1}\mathbf{T}$ is replaced by $\mathbf{S}^{-1}\mathbf{C}$ and the diagonal of \mathbf{H} is not modified. The function Givens(\mathbf{a}) below generates a 2×2 rotation matrix \mathbf{Q} with the property that $(\mathbf{Q}\mathbf{a})_2 = 0$.

ALGORITHM 2 (TMRES for $\mathbf{C}\mathbf{x} = \mathbf{b}$).

```

C = S - T, S nonsingular,  $\rho(\mathbf{S}^{-1}\mathbf{T}) \leq 1$ 
g ←  $\mathbf{S}^{-1}(\mathbf{b} - \mathbf{C}\mathbf{x}_0)$ , v1 ← g/g, e ← 0,  $e_1 \leftarrow \|\mathbf{g}\|$ 
for  $j = 1, 2, \dots$  until convergence
  s ←  $\mathbf{S}^{-1}(\mathbf{T}\mathbf{v}_j)$ 
  for  $i = 1 : j$ 
     $h_{ij} \leftarrow \mathbf{s}^\top \mathbf{v}_i$ 
    s ← s -  $\mathbf{v}_i h_{ij}$ 
  end
   $h_{j+1,j} \leftarrow \|\mathbf{s}\|$ 
  v $j+1$  ← s/ $h_{j+1,j}$ 
   $h_{ii} \leftarrow h_{ii} - 1$ 
  for  $i = 1 : j - 1$ 
     $\mathbf{H}(i : i + 1, j) \leftarrow \mathbf{Q}_i \mathbf{H}(i : i + 1, j)$ 
  end
   $\mathbf{Q}_j \leftarrow \text{Givens}(\mathbf{H}(j : j + 1, j))$ 
   $\mathbf{H}(j : j + 1, j) \leftarrow \mathbf{Q}_j \mathbf{H}(j : j + 1, j)$ 
   $\mathbf{e}(j : j + 1) \leftarrow \mathbf{Q}_j \mathbf{e}(j : j + 1)$ 
end
x $j$  ←  $\mathbf{x}_0 - \mathbf{V}_j (\mathbf{H}(1 : j, 1 : j))^{-1} \mathbf{e}(1 : j)$ 
end Algorithm 2

```

Possible choices of \mathbf{S} with the property that $\rho(\mathbf{S}^{-1}\mathbf{T}) < 1$ are the following: (i) if \mathbf{C} is symmetric and positive definite, then $\mathbf{S} = \mathbf{L}$, where \mathbf{L} is the lower triangular matrix whose lower triangle matches that of \mathbf{C} (Gauss-Seidel choice); (ii) if \mathbf{C} is row diagonally dominant, then $\mathbf{S} = \mathbf{L}$ or $\mathbf{S} = \mathbf{D}$, where \mathbf{D} is the diagonal matrix whose diagonal matches that of \mathbf{C} (Jacobi choice); (iii) if $\mathbf{C} = \mathbf{A}\mathbf{A}^\top + \sigma\mathbf{I}$ where the columns of \mathbf{A} come from the columns of a larger matrix $\mathbf{B} = (\mathbf{A}|\mathbf{N})$, then $\mathbf{S} = \mathbf{B}\mathbf{B}^\top + \sigma\mathbf{I}$ (see Theorem 6.2 below).

4. Convergence analysis. The convergence of TMRES follows readily from previously established theory. Recall [29, p. 471] that the index of the matrix \mathbf{M} with respect to the zero eigenvalue is the size of the largest singular Jordan block. If $\mathcal{R}(\mathbf{C})$ denotes the range of \mathbf{C} , then we have the following result.

THEOREM 4.1. *Let \mathbf{C} be a square matrix with $\mathbf{C} = \mathbf{S} - \mathbf{T}$, where \mathbf{S} is nonsingular, and define $\mathbf{M} = \mathbf{S}^{-1}\mathbf{T}$. Then $\mathbf{C}\mathbf{x} = \mathbf{b}$ has a solution in the Krylov space $\mathcal{K}(\mathbf{M}, \mathbf{S}^{-1}\mathbf{b}, k)$ for some k if and only if $\mathbf{S}^{-1}\mathbf{b} \in \mathcal{R}((\mathbf{I} - \mathbf{M})^i)$, where i is the index of the zero eigenvalue of $\mathbf{I} - \mathbf{M}$.*

Proof. By [30, Thm. 2] the linear system $(\mathbf{I} - \mathbf{M})\mathbf{x} = \mathbf{g}$ has a solution in the Krylov space $\mathcal{K}(\mathbf{I} - \mathbf{M}, \mathbf{g}, k)$ for k sufficiently large if and only if $\mathbf{g} \in \mathcal{R}((\mathbf{I} - \mathbf{M})^i)$, where i is the index of the zero eigenvalue of $\mathbf{I} - \mathbf{M}$. If P_k denotes the set of polynomials of degree at most k , it can be shown that

$$P_k = \{p(y) : y = 1 - x, p \in P_k\}.$$

As a result, we have $\mathcal{K}(\mathbf{M}, \mathbf{g}, k) = \mathcal{K}(\mathbf{I} - \mathbf{M}, \mathbf{g}, k)$. Hence, by [30, Thm. 2] the linear system $(\mathbf{I} - \mathbf{M})\mathbf{x} = \mathbf{g}$ has a solution in the Krylov space $\mathcal{K}(\mathbf{M}, \mathbf{g}, k)$ for k sufficiently large if and only if $\mathbf{g} \in \mathcal{R}((\mathbf{I} - \mathbf{M})^i)$. We take $\mathbf{g} = \mathbf{S}^{-1}\mathbf{b}$ to complete the proof. \square

If one is not an eigenvalue of \mathbf{M} , then the index of the zero eigenvalue of $\mathbf{I} - \mathbf{M}$ is zero and the condition $\mathbf{S}^{-1}\mathbf{b} \in \mathcal{R}(\mathbf{I})$ holds trivially. If one is a nondefective eigenvalue of \mathbf{M} , then the index of the zero eigenvalue of $\mathbf{I} - \mathbf{M}$ is one and the condition $\mathbf{S}^{-1}\mathbf{b} \in \mathcal{R}(\mathbf{I} - \mathbf{M})$ is equivalent to $\mathbf{b} \in \mathcal{R}(\mathbf{C})$.

An analysis of the convergence of TMRES in terms of the spectrum of \mathbf{M} can be given along the lines of the analysis for the conjugate gradient method in [48] or [14], and for GMRES in [43]. For example, we have the following result.

THEOREM 4.2. *Suppose that $\mathbf{C} = \mathbf{S} - \mathbf{T}$, where \mathbf{S} is invertible, and that there exists a nonsingular matrix \mathbf{F} such that $\mathbf{M} = \mathbf{S}^{-1}\mathbf{T} = \mathbf{F}\mathbf{\Phi}\mathbf{F}^{-1}$, where $\mathbf{\Phi}$ is diagonal. If \mathbf{x}_k denotes a solution to (3.3), and μ_i is the i th diagonal element of $\mathbf{\Phi}$, then we have*

$$\|\mathbf{t}(\mathbf{x}_k)\| \leq \|\mathbf{F}\| \|\mathbf{F}^{-1}\| \|\mathbf{t}(\mathbf{x}_0)\| \min_{\substack{p \in P_k \\ p(0)=1}} \max_i |p(1 - \mu_i)|.$$

Proof. Simply apply [43, Prop. 4] to the transformed system $(\mathbf{I} - \mathbf{M})\mathbf{z} = \mathbf{g}$, where $\mathbf{g} = \mathbf{t}(\mathbf{x}_0)$. \square

5. More examples. When \mathbf{C} is symmetric and positive definite, the Gauss-Seidel splitting is convergent and, hence, yields an appropriate TMRES splitting. Figure 1.1 shows the iterates for the TMRES/GS algorithm applied to (1.1) where \mathbf{A} is the matrix from beaconfd and $\sigma = 0$. After about 35 iterations, the norm of the residual has been reduced by 12 orders of magnitude, providing relatively rapid convergence.

Next, we examine the convergence for progressively larger problems from the netlib/lp directory. Figure 5.1 shows the convergence for the system (1.1) where \mathbf{A} is the matrix from ffff800 ($m = 524$, $n = 1028$, $\sigma = 0$). The condition number for $\mathbf{A}\mathbf{A}^T$ is of order 10^{19} . Also, in Figure 5.1 we show the convergence of preconditioned GMRES with the preconditioner $\mathbf{S} = \mathbf{L}$, the lower triangular part of \mathbf{C} . Theoretically, the TMRES and the preconditioned GMRES curves should coincide since the Krylov spaces generated by \mathbf{M} and by $\mathbf{I} - \mathbf{M}$ are identical, algebraically. Note though that when a vector is multiplied by $\mathbf{I} - \mathbf{M}$, some subtraction of nearly equal numbers occurs: If \mathbf{f} is an eigenvector of \mathbf{M} whose associated eigenvalue μ is near one, then

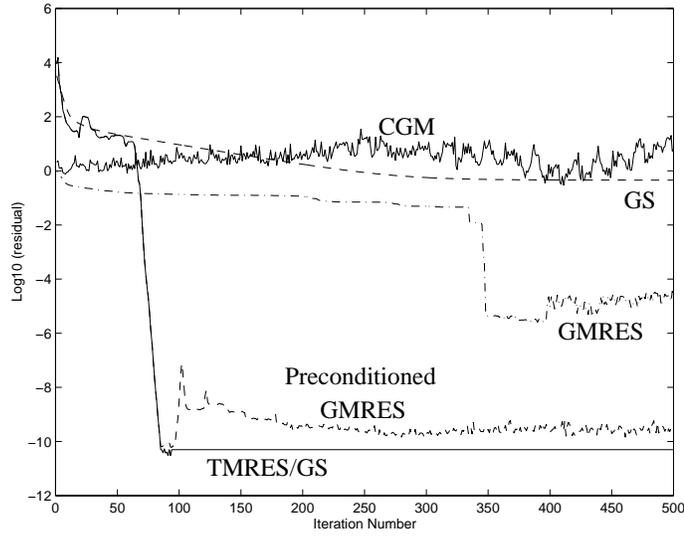


FIG. 5.1. Convergence for the test problem *ffff800*, $\sigma = 0$.

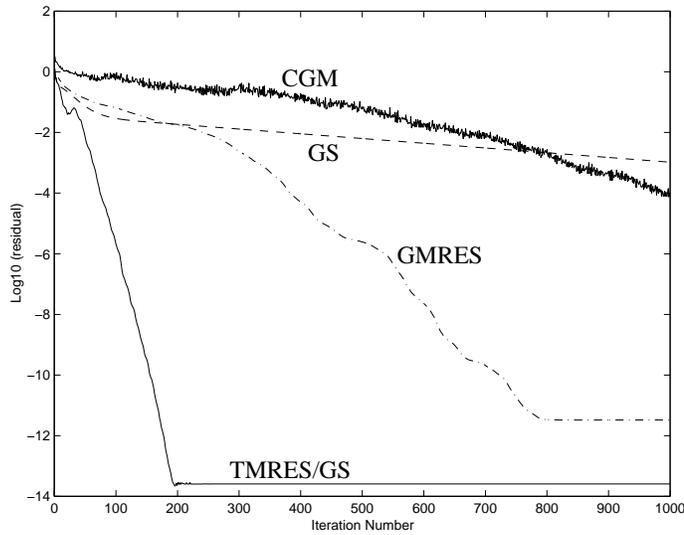


FIG. 5.2. Convergence for the test problem *greenbea*, zero rows deleted, $\sigma = 0$.

$(\mathbf{I} - \mathbf{M})\mathbf{f} = (1 - \mu)\mathbf{f} \approx \mathbf{0}$. As is well known (see [25, sects. 1–4]), the subtraction of nearly equal numbers can produce a large relative error and numerical differences between TMRES and GMRES.

Figures 5.2 and 5.3 show the convergence for *greenbea* ($m = 2392$, $n = 5598$). Since 3 rows in *greenbea* are completely zero, we ran two variations of the problem. In the variation of Figure 5.2, the zero rows are deleted, while in the variation of Figure 5.3, the zero rows are included, but $\sigma = 10^{-6}$ to keep the diagonal of \mathbf{C} nonzero. In each case, the TMRES algorithm exhibits an attractive convergence rate. Since the TMRES algorithm minimizes the norm of the transformed residual in

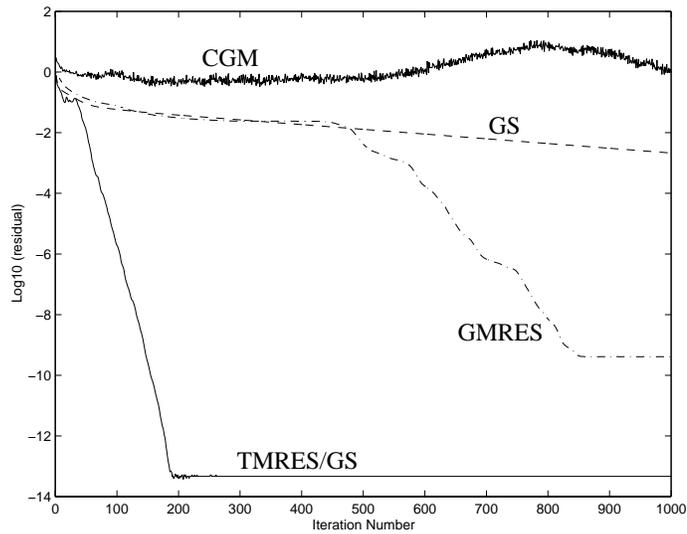


FIG. 5.3. Convergence for the test problem *greenbea*, zero rows included, $\sigma = 10^{-6}$.

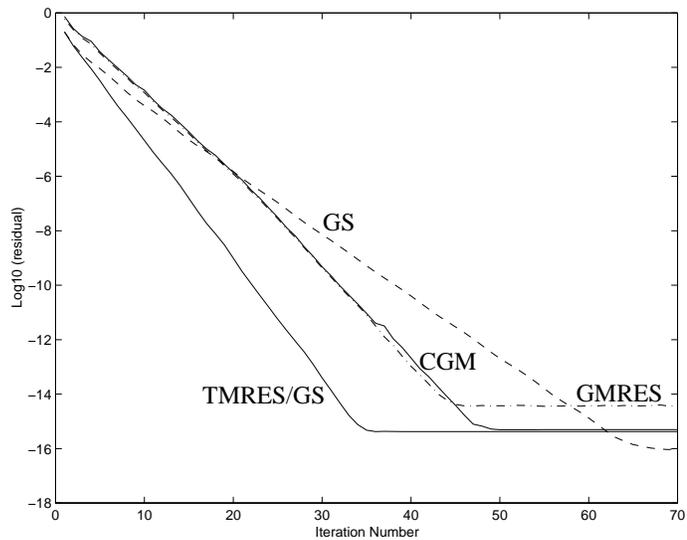


FIG. 5.4. Convergence for the test problem *greenbea*, $\sigma = 10$.

each step, the norm of the residual itself, plotted in Figures 1.1 and 5.1–5.3, is not guaranteed to decay in each iteration. In fact, GMRES is the only algorithm depicted in these figures that is guaranteed (in theory) to reduce the norm of the residual in each step.

Finally, we took $\sigma = 10$ in *greenbea* in order to examine the convergence speed when the eigenvalues of the matrix are strongly bounded away from zero and the matrix is well-conditioned. As seen in Figure 5.4, all the schemes converge quickly. The convergence of CGM and GMRES is almost identical since the ratio between the largest and smallest eigenvalue in (2.5) is of order 1, and hence, CGM and GMRES

minimize almost the same expression in each step. We emphasize that although TMRES with Gauss–Seidel splitting is effective for nearly singular linear systems such of those found in netlib/lp, its convergence when applied to a well-conditioned partial differential equation may be no better than either CGM or GMRES.

6. Symmetric TMRES. It is well known that for the Arnoldi process (Algorithm 1), the leading submatrices of the upper Hessenberg matrix \mathbf{H} are tridiagonal if $\mathbf{M} = \mathbf{S}^{-1}\mathbf{T}$ is symmetric. Hence, in this symmetric case, we can skip the evaluation of h_{ij} when $i < j - 1$. In the spirit of Hageman and Young’s terminology [22], we say that the iteration (3.2) is symmetrizable if there exists an invertible matrix \mathbf{W} such that \mathbf{WMW}^{-1} is symmetric. With the change of variables $\mathbf{x} = \mathbf{W}^{-1}\mathbf{w} + \mathbf{x}_0$, the iteration (3.2) can be written

$$(6.1) \quad \mathbf{w}_{k+1} = \mathbf{WMW}^{-1}\mathbf{w}_k + \mathbf{Wg}, \quad \mathbf{g} = \mathbf{S}^{-1}(\mathbf{b} - \mathbf{Cx}_0).$$

We now form a least-squares problem analogous to (3.4) to obtain an approximation \mathbf{x}_k to a solution of $\mathbf{Cx} = \mathbf{b}$. In particular, $\mathbf{x}_k = \mathbf{W}^{-1}\mathbf{z}_k + \mathbf{x}_0$, where \mathbf{z}_k is a solution to

$$(6.2) \quad \min \|\mathbf{Wg} - (\mathbf{I} - \mathbf{WMW}^{-1})\mathbf{z}\| \quad \text{subject to} \quad \mathbf{z} \in \mathcal{K}(\mathbf{WMW}^{-1}, \mathbf{Wg}, k).$$

This symmetric least squares problem can be solved using Algorithm 2. After taking into account the tridiagonal structure of \mathbf{H} , the resulting algorithm is closely connected with Paige and Saunders’ MINRES algorithm [40]. Although the orthonormal vectors \mathbf{v}_i for $i \leq j$ are needed in Algorithm 2, the following idea of Paige and Saunders [40, 41] can be used to avoid their storage. Instead of evaluating

$$\mathbf{V}_j (\mathbf{H}(1 : j, 1 : j))^{-1} \mathbf{e}(1 : j)$$

as we do in Algorithm 2, Paige and Saunders evaluate

$$(\mathbf{V}_j \mathbf{H}(1 : j, 1 : j))^{-1} \mathbf{e}(1 : j).$$

Let \mathbf{R} denote the upper triangular matrix $\mathbf{H}(1 : j, 1 : j)$ generated by Algorithm 2, and define $\mathbf{Z} = \mathbf{V}_j \mathbf{R}^{-1}$. Since the Hessenberg matrix used to generate \mathbf{R} is tridiagonal, \mathbf{R} has three bands, a diagonal band and two superdiagonal bands. Hence, the j th column \mathbf{z}_j of \mathbf{Z} is given by the recurrence

$$\mathbf{z}_j = (\mathbf{v}_j - \mathbf{z}_{j-1}r_{j-1,j} - \mathbf{z}_{j-2}r_{j-2,j})/r_{jj}.$$

The complete algorithm is the following.

ALGORITHM 3 (TMRES for $\mathbf{Cx} = \mathbf{b}$, $\mathbf{WS}^{-1}\mathbf{CW}^{-1}$ symmetric).

$\mathbf{g} \leftarrow \mathbf{S}^{-1}(\mathbf{b} - \mathbf{Cx}_0)$, $\mathbf{e} \leftarrow \mathbf{0}$, $e_1 \leftarrow \|\mathbf{Wg}\|$, $\mathbf{v}_1 \leftarrow \mathbf{Wg}/e_1$

$u \leftarrow 0$, $\mathbf{Q}_0 \leftarrow \mathbf{I}$, $\mathbf{Q}_{-1} \leftarrow \mathbf{0}$

for $j = 1, 2, \dots$ until convergence

$\mathbf{r} \leftarrow \mathbf{W}^{-1}\mathbf{v}_j$

$\mathbf{s} \leftarrow \mathbf{W}(\mathbf{Mr})$ ($\mathbf{C} = \mathbf{S} - \mathbf{T}$, $\mathbf{M} = \mathbf{S}^{-1}\mathbf{T}$)

$d \leftarrow \mathbf{s}^T \mathbf{v}_j$, $\bar{u} \leftarrow u$

$\mathbf{s} \leftarrow \mathbf{s} - d\mathbf{v}_j - u\mathbf{v}_{j-1}$

$u \leftarrow \|\mathbf{s}\|$

$\mathbf{v}_{j+1} \leftarrow \mathbf{s}/u$

$d \leftarrow d - 1$

```

[t  $\bar{u}$ ]  $\leftarrow$  [0  $\bar{u}$ ]  $\mathbf{Q}_{j-2}^T$ 
[ $\bar{u}$   $d$ ]  $\leftarrow$  [ $\bar{u}$   $d$ ]  $\mathbf{Q}_{j-1}^T$ 
 $\mathbf{Q}_j \leftarrow$  Givens( $d, u$ )
 $\mathbf{e}(j : j + 1) \leftarrow \mathbf{Q}_j \mathbf{e}(j : j + 1)$ 
 $d \leftarrow ([d \ u] \mathbf{Q}_j^T)_1$ 
 $\mathbf{z}_j \leftarrow (\mathbf{r} - \bar{u} \mathbf{z}_{j-1} - t \mathbf{z}_{j-2})/d$ 
 $\mathbf{x}_j \leftarrow \mathbf{x}_{j-1} - \mathbf{z}_j e_j$ 

```

end

end Algorithm 3

Since an iteration of Algorithm 3 requires only the vectors \mathbf{v}_j , \mathbf{v}_{j-1} , \mathbf{z}_{j-1} , \mathbf{z}_{j-2} , and \mathbf{x}_{j-1} , all the preceding vectors can be discarded. As in Algorithm 2, for a nearly singular linear system $\mathbf{C}\mathbf{x} = \mathbf{b}$, the iteration (6.1) should be devised in such a way that $\rho(\mathbf{M}) \leq 1$. If \mathbf{z} lies in the null space of \mathbf{C} , then $\mathbf{W}\mathbf{z}$ is an eigenvector of $\mathbf{W}\mathbf{M}\mathbf{W}^{-1} = \mathbf{W}(\mathbf{S}^{-1}\mathbf{T})\mathbf{W}^{-1}$ whose eigenvalue is 1. Thus the eigenvalues of smallest modulus in the original problem are associated with eigenvalues of largest modulus in the transformed problem.

There are various ways to choose a symmetrizing matrix \mathbf{W} . If $\mathbf{C} = \mathbf{S} - \mathbf{T}$ is a splitting for which $\rho(\mathbf{M}) = \rho(\mathbf{S}^{-1}\mathbf{T}) < 1$ where \mathbf{S} is symmetric and positive definite and \mathbf{T} is symmetric, then we can take $\mathbf{W} = \mathbf{S}^{1/2}$. With this choice for \mathbf{W} , we can avoid squares roots in Algorithm 3 if vectors like \mathbf{s} and \mathbf{v}_j are replaced by new variables that are equal to \mathbf{W}^{-1} times old variables. The resulting algorithm is the following.

ALGORITHM 3a (TMRES for $\mathbf{C}\mathbf{x} = \mathbf{b}$, $\mathbf{W}\mathbf{S}^{-1}\mathbf{C}\mathbf{W}^{-1}$ symmetric).

```

 $\mathbf{g} \leftarrow \mathbf{S}^{-1}(\mathbf{b} - \mathbf{C}\mathbf{x}_0)$ ,  $\mathbf{e} \leftarrow \mathbf{0}$ ,  $e_1 \leftarrow \|\mathbf{W}\mathbf{g}\|$ ,  $\mathbf{v}_1 \leftarrow \mathbf{g}/e_1$ 
 $u \leftarrow 0$ ,  $\mathbf{Q}_0 \leftarrow \mathbf{I}$ ,  $\mathbf{Q}_{-1} \leftarrow \mathbf{0}$ 
for  $j = 1, 2, \dots$  until convergence
   $\mathbf{r} \leftarrow \mathbf{v}_j$ 
   $\mathbf{s} \leftarrow \mathbf{M}\mathbf{r}$  ( $\mathbf{C} = \mathbf{S} - \mathbf{T}$ ,  $\mathbf{M} = \mathbf{S}^{-1}\mathbf{T}$ )
   $d \leftarrow (\mathbf{W}\mathbf{s})^T(\mathbf{W}\mathbf{v}_j)$ ,  $\bar{u} \leftarrow u$ 
   $\mathbf{s} \leftarrow \mathbf{s} - d\mathbf{v}_j - u\mathbf{v}_{j-1}$ 
   $u \leftarrow \|\mathbf{W}\mathbf{s}\|$ 
  Continue as in Algorithm 3

```

end

end Algorithm 3a

In this variation of Algorithm 3, \mathbf{W} enters as a product $\mathbf{W}^T\mathbf{W}$. Hence, when $\mathbf{W} = \mathbf{S}^{1/2}$, we have $\mathbf{W}^T\mathbf{W} = \mathbf{S}$, and the square root is gone.

As an illustration of this symmetrization, suppose that \mathbf{C} is symmetric and positive definite and \mathbf{D} is a block diagonal matrix whose diagonal blocks match those of \mathbf{C} . The damped Jacobi iteration (see [22]) corresponding to the splitting $\mathbf{S} = \omega\mathbf{D}$ is convergent when the damping parameter ω is sufficiently large, and for this splitting, \mathbf{S} is positive definite. To determine a suitable value for ω , observe that

$$\mathbf{S}^{-1}\mathbf{T} = \mathbf{I} - \frac{1}{\omega}\mathbf{D}^{-1}\mathbf{C},$$

where $\mathbf{D}^{-1}\mathbf{C}$ is similar to the symmetric, positive definite matrix $\mathbf{D}^{-1/2}\mathbf{C}\mathbf{D}^{-1/2}$. The eigenvalues of $\mathbf{S}^{-1}\mathbf{T}$ are real and less than 1 for any choice of $\omega > 0$, and for $\omega > \frac{1}{2}\rho(\mathbf{D}^{-1}\mathbf{C})$, the eigenvalues of $\mathbf{S}^{-1}\mathbf{T}$ are greater than -1 . Since the spectral radius of a matrix is bounded by any matrix norm, we have $\rho(\mathbf{M}) \leq \|\mathbf{M}\|_1$. An estimate for the 1-norm of a matrix can be obtained using the algorithm developed

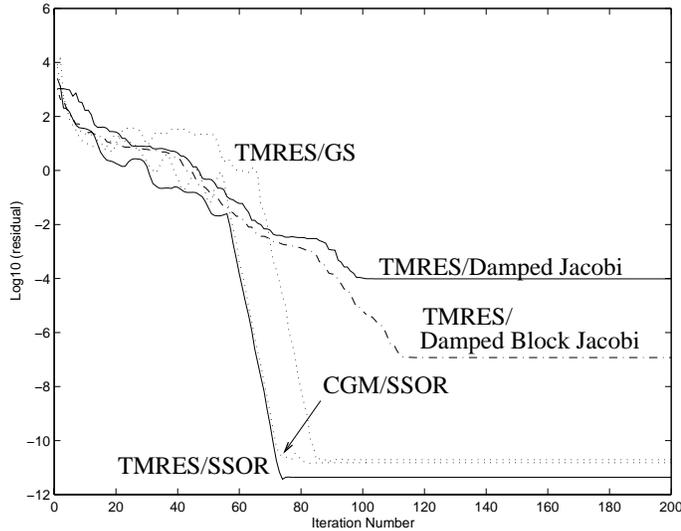


FIG. 6.1. Convergence for TMRES using the test problem ffff800 and various symmetrization schemes.

in [24]. Also see [25, p. 139] and [15, Alg. 7] for symbolic formulations of this algorithm. A Matlab version of this 1-norm estimation algorithm, due to Tim Davis, is available in Mathwork's contributed m-files ftp site, ftp.mathworks.com, as the file pub/contrib/v4/linalg/normest1.m, while Nick Higham gives a Fortran implementation in ACM TOMS Algorithm 674 [28].

In Figure 6.1, we show the convergence of TMRES using the damped Jacobi scheme, where \mathbf{D} is the diagonal of \mathbf{C} and ω is $\frac{2}{3}\|\mathbf{D}^{-1}\mathbf{C}\|_1$, for the test problem ffff800. We also show in Figure 6.1 convergence for TMRES/GS and for a damped Jacobi scheme associated with a block diagonal \mathbf{D} gotten from the Chaco partitioning code [27] of Hendrickson and Rothberg by permuting the rows and columns in order to maximize the number of nonzero elements of \mathbf{C} in diagonal blocks of size 8×8 and 9×9 (60 blocks altogether). Similar partitions are generated by the code Metis [34]. Observe that the TMRES/Block Jacobi scheme converges slightly faster than TMRES/Jacobi, but not as fast as TMRES/GS.

For schemes like Gauss-Seidel or successive overrelaxation (SOR), the \mathbf{S} matrix is not symmetric, and the square root $\mathbf{S}^{1/2}$ may be complex or nonsymmetric. However, when $\mathbf{C} = \mathbf{S} - \mathbf{T}$ is symmetric, a two-step symmetrization is possible using the original iteration followed by its transpose:

$$\mathbf{S}\mathbf{x}_{k+1/2} = \mathbf{T}\mathbf{x}_k + \mathbf{b}, \quad \mathbf{S}^T\mathbf{x}_{k+1} = \mathbf{T}^T\mathbf{x}_{k+1/2} + \mathbf{b}.$$

If the two steps are combined into one step, then

$$\mathbf{x}_{k+1} = \mathbf{M}_2\mathbf{x}_k + \mathbf{g}, \quad \text{where } \mathbf{M}_2 = \mathbf{S}^{-T}\mathbf{T}^T\mathbf{S}^{-1}\mathbf{T} \quad \text{and} \quad \mathbf{g} = \mathbf{S}^{-T}(\mathbf{T}^T\mathbf{S}^{-1} + \mathbf{I})\mathbf{b}.$$

In particular, for the SOR splitting of \mathbf{C} , this two-step scheme is the same as symmetric SOR (SSOR). As we now show, these two-step schemes are always symmetrizable when \mathbf{C} is symmetric and positive definite.

PROPOSITION 6.1. *If $\mathbf{C} = \mathbf{S} - \mathbf{T}$, where \mathbf{S} is nonsingular and \mathbf{C} is symmetric and positive definite, then $\mathbf{M}_2 = \mathbf{S}^{-T}\mathbf{T}^T\mathbf{S}^{-1}\mathbf{T}$ is symmetrizable with $\mathbf{W} = \mathbf{C}^{1/2}$.*

Proof. Substituting $\mathbf{T} = \mathbf{S} - \mathbf{C}$, we obtain

$$\mathbf{M}_2 = \mathbf{I} - \mathbf{S}^{-\top} \mathbf{C} - \mathbf{S}^{-1} \mathbf{C} + \mathbf{S}^{-\top} \mathbf{C} \mathbf{S}^{-1} \mathbf{C}.$$

Hence, $\mathbf{C}^{1/2} \mathbf{M}_2 \mathbf{C}^{-1/2}$ is symmetric. \square

Let \mathbf{D} be the diagonal of a symmetric, positive definite matrix \mathbf{C} . For the SOR splitting $\mathbf{S} = \mathbf{L} + \rho \mathbf{D}$, where $\rho = (1 - \omega)/\omega$ and $0 < \omega < 2$, another possible symmetrization is $\mathbf{W} = \mathbf{D}^{-1/2} (\mathbf{L} + \rho \mathbf{D})^\top$ (see [22, p. 31]). In Figure 6.1, the curve labeled TMRES/SSOR, we show the convergence of symmetrized TMRES for the Gauss-Seidel/SOR splitting $\mathbf{S} = \mathbf{L}$ ($\rho = 0$). Also in Figure 6.1 we show the convergence of a preconditioned conjugate gradient method [6, 12, 13] with SSOR preconditioner ($\omega = 1$). Observe that the convergence of TMRES/SSOR and CGM/SSOR are very similar. Both schemes generate iterates in the same Krylov space, but they differ in the merit function used to select the approximate solution.

The following identities can be used to streamline the implementation of the SSOR preconditioned iteration:

$$(6.3) \quad \begin{aligned} \mathbf{M}_2 &= \mathbf{I} - (2\rho + 1)(\mathbf{S}^{-\top} \mathbf{D} \mathbf{S}^{-1}) \mathbf{C}, \\ \mathbf{W} \mathbf{M}_2 \mathbf{W}^{-1} &= \mathbf{I} - (2\rho + 1) \mathbf{D}^{1/2} \mathbf{S}^{-1} \mathbf{C} \mathbf{S}^{-\top} \mathbf{D}^{1/2}. \end{aligned}$$

From this relation, we see that if \mathbf{z} lies in the null space of \mathbf{C} , then $\mathbf{W} \mathbf{z}$ is an eigenvector of $\mathbf{W} \mathbf{M}_2 \mathbf{W}^{-1}$ with eigenvalue 1. Hence, eigenvalues of smallest modulus for \mathbf{C} again correspond to eigenvalues of largest modulus for $\mathbf{W} \mathbf{M}_2 \mathbf{W}^{-1}$.

In some optimization applications, such as the LP dual active set algorithm, we need to solve many different systems of the form (1.1) where the columns of \mathbf{A} come from the columns of a larger matrix \mathbf{B} : $\mathbf{B} = (\mathbf{A} | \mathbf{N})$. As a consequence of the following result, which is basically a reformulation of the classical convergence theorem used for the SOR scheme, the splitting obtained by taking $\mathbf{S} = \mathbf{B} \mathbf{B}^\top + \sigma \mathbf{I}$ and $\mathbf{T} = \mathbf{N} \mathbf{N}^\top$ is convergent.

THEOREM 6.2. *Suppose that \mathbf{C} is symmetric and positive definite and $\mathbf{C} = \mathbf{S} - \mathbf{T}$, where \mathbf{S} is invertible. If $\mathbf{S} + \mathbf{T}^\top$ is positive definite, then $\rho(\mathbf{M}) < 1 > \rho(\mathbf{M}_2)$.*

Proof. Our proof that $\rho(\mathbf{M}) < 1$ is essentially that which appears in [49, p. 77] recast in terms of a matrix splitting. A proof that $\rho(\mathbf{M}_2) < 1$ for the SOR iteration appears in [21] (also see the classic reference [52]), while here we obtain a more general result.

Let \mathbf{f} be any vector and define $\mathbf{g} = \mathbf{M} \mathbf{f}$ and $\boldsymbol{\delta} = \mathbf{f} - \mathbf{g}$. From the identity $\mathbf{S} \mathbf{g} = \mathbf{T} \mathbf{f}$, we obtain the following relations:

$$(\mathbf{S} - \mathbf{T}) \mathbf{g} + \mathbf{T} \mathbf{g} = \mathbf{T} \mathbf{f} \quad \text{or} \quad \mathbf{C} \mathbf{g} = \mathbf{T} \boldsymbol{\delta},$$

and

$$\mathbf{S} \mathbf{g} = (\mathbf{T} - \mathbf{S}) \mathbf{f} + \mathbf{S} \mathbf{f} \quad \text{or} \quad \mathbf{C} \mathbf{f} = \mathbf{S} \boldsymbol{\delta}.$$

Multiplying the first relation by \mathbf{g}^* (the conjugate transpose of \mathbf{g}) and the second relation by \mathbf{f}^* , subtracting, and exploiting the identity $\mathbf{S} \mathbf{g} = \mathbf{T} \mathbf{f}$, we have

$$(6.4) \quad \begin{aligned} \mathbf{f}^* \mathbf{C} \mathbf{f} - \mathbf{g}^* \mathbf{C} \mathbf{g} &= (\mathbf{f}^* \mathbf{S} - \mathbf{g}^* \mathbf{T}) \boldsymbol{\delta} \\ &= (\mathbf{f}^* \mathbf{S} - \mathbf{g}^* \mathbf{T}) \boldsymbol{\delta} + (\mathbf{T} \mathbf{f} - \mathbf{S} \mathbf{g})^* \boldsymbol{\delta} \\ &= \mathbf{f}^* (\mathbf{S} + \mathbf{T}^\top) \boldsymbol{\delta} - \mathbf{g}^* (\mathbf{T} + \mathbf{S}^\top) \boldsymbol{\delta}. \end{aligned}$$

Since \mathbf{C} is symmetric,

$$\mathbf{C}^\top = \mathbf{S}^\top - \mathbf{T}^\top = \mathbf{C} = \mathbf{S} - \mathbf{T}.$$

Hence, $\mathbf{S} + \mathbf{T}^\top = \mathbf{S}^\top + \mathbf{T}$, and substituting this in (6.4) gives

$$(6.5) \quad \mathbf{f}^* \mathbf{C} \mathbf{f} - \mathbf{g}^* \mathbf{C} \mathbf{g} = \delta^* (\mathbf{S} + \mathbf{T}^\top) \delta.$$

If \mathbf{f} is an eigenvector associated with an absolute largest eigenvalue λ of \mathbf{M} , then $\mathbf{g} = \mathbf{M} \mathbf{f} = \lambda \mathbf{f}$, and it follows from (6.5) that

$$(1 - |\lambda|^2) \mathbf{f}^* \mathbf{C} \mathbf{f} = \delta^* (\mathbf{S} + \mathbf{T}^\top) \delta \geq 0$$

since $\mathbf{S} + \mathbf{T}^\top$ is positive definite. Since \mathbf{C} is also positive definite, we conclude that $|\lambda| \leq 1$. If $|\lambda| = 1$, then $\delta = \mathbf{0}$, or $\mathbf{g} = \mathbf{f} = \mathbf{M} \mathbf{f}$. Rearranging this last identity gives $\mathbf{C} \mathbf{f} = \mathbf{0}$, which is impossible since \mathbf{C} is invertible. Hence $|\lambda| < 1$ and $\rho(\mathbf{M}) < 1$.

With \mathbf{h} defined by $\mathbf{h} = \mathbf{S}^{-\top} \mathbf{T}^\top \mathbf{g} = \mathbf{S}^{-\top} \mathbf{T}^\top \mathbf{S}^{-1} \mathbf{T} \mathbf{f} = \mathbf{M}_2 \mathbf{f}$, the same manipulations used to obtain (6.5), but with \mathbf{S} and \mathbf{T} replaced by \mathbf{S}^\top and \mathbf{T}^\top , yield the relation

$$\mathbf{g}^* \mathbf{C} \mathbf{g} - \mathbf{h}^* \mathbf{C} \mathbf{h} = \epsilon^* (\mathbf{S} + \mathbf{T}^\top) \epsilon, \quad \epsilon = \mathbf{g} - \mathbf{h}.$$

Adding this relation to (6.5), we obtain

$$\mathbf{f}^* \mathbf{C} \mathbf{f} - \mathbf{h}^* \mathbf{C} \mathbf{h} = \delta^* (\mathbf{S} + \mathbf{T}^\top) \delta + \epsilon^* (\mathbf{S} + \mathbf{T}^\top) \epsilon \geq 0.$$

Now, if \mathbf{f} is an eigenvector associated with the absolute largest eigenvalue μ of \mathbf{M}_2 , then $|\mu| \leq 1$ and $|\mu| = 1$ if and only if $\delta = \mathbf{0} = \mathbf{f} - \mathbf{g}$ and $\epsilon = \mathbf{0} = \mathbf{g} - \mathbf{h}$. Again, the identity $\mathbf{g} = \mathbf{f} = \mathbf{M} \mathbf{f}$ implies that \mathbf{C} is singular. Hence $|\mu| < 1$ and $\rho(\mathbf{M}_2) < 1$. \square

7. Sparsity considerations. If TMRES is implemented using either a Gauss-Seidel or an SOR splitting, then for a linear system of the form (1.1), it would appear that the product $\mathbf{A} \mathbf{A}^\top$ must be evaluated. After forming this product, the number of multiplications and additions needed to perform the iteration is $\text{nnz}(\mathbf{A} \mathbf{A}^\top)$, where nnz is (Matlab) notation for the number of nonzero elements. When \mathbf{A} is sparse, $\mathbf{A} \mathbf{A}^\top$ often has more nonzero entries than \mathbf{A} itself. Hence, for some sparse matrices, it may be more efficient to express an iteration for (1.1) in terms of \mathbf{A} itself rather than $\mathbf{A} \mathbf{A}^\top$. In [6] (also see [5, p. 284]) Björck and Elfving show that for SSOR preconditioned conjugate gradient iterations, the matrix-vector product $(\mathbf{W} \mathbf{M}_2 \mathbf{W}^{-1}) \mathbf{x}$ associated with the SSOR matrix (6.3) and $\mathbf{A} \mathbf{A}^\top$ can be evaluated in about $4 \text{nnz}(\mathbf{A})$ multiplications since products of the form $\mathbf{S}^{-\top} \mathbf{x}$ and $\mathbf{A}(\mathbf{S}^{-\top} \mathbf{x})$ can be simultaneously evaluated in $2 \text{nnz}(\mathbf{A})$ multiplications altogether. A similar approach applies to SOR splittings as we now observe.

If \mathbf{y} and \mathbf{z} denote the iterates \mathbf{x}_k and \mathbf{x}_{k+1} , respectively, in (3.2) (so that $\mathbf{S} \mathbf{z} = \mathbf{T} \mathbf{y} + \mathbf{b}$), then for the SOR iteration with relaxation parameter $0 < \omega < 2$, we have

$$z_i - y_i = \frac{\omega \left(b_i - \sum_{j=1}^{i-1} c_{ij} z_j - \sum_{j=i}^m c_{ij} y_j \right)}{c_{ii}}, \quad i = 1 \text{ to } m.$$

Since $\mathbf{C} = \mathbf{A} \mathbf{A}^\top + \sigma \mathbf{I}$ for the system (1.1), it follows that

$$c_{ij} = \mathbf{a}_i^\top \mathbf{a}_j \text{ for } i \neq j, \quad c_{ii} = \sigma + \|\mathbf{a}_i\|^2,$$

TABLE 1
Method and approximate number of multiplications and additions.

Method	Operations
TMRES/SSOR	4nnz(\mathbf{A})
CGM/SSOR	4nnz(\mathbf{A})
TMRES/SOR	2nnz(\mathbf{A})
GS	2nnz(\mathbf{A})
GMRES	2nnz(\mathbf{A})
TMRES/Jacobi	2nnz(\mathbf{A})
CGM	2nnz(\mathbf{A})

where \mathbf{a}_i denotes the i th column of \mathbf{A}^T . Let us define $\mathbf{q}_i \in \mathbf{R}^n$ by

$$\mathbf{q}_i = \sum_{j=1}^{i-1} \mathbf{a}_j z_j + \sum_{j=i}^m \mathbf{a}_j y_j.$$

From this definition, it follows that

$$z_i - y_i = \omega(b_i - \sigma y_i - \mathbf{q}_i^T \mathbf{a}_i) / c_{ii}, \quad \text{where } \mathbf{q}_{i+1} = \mathbf{q}_i + \mathbf{a}_i(z_i - y_i).$$

Hence, an iteration of SOR, overwriting the current iterate \mathbf{x}_k by the new iterate \mathbf{x}_{k+1} , can be implemented in the following way.

ALGORITHM 4 (SOR for $(\mathbf{A}\mathbf{A}^T + \sigma\mathbf{I})\mathbf{x} = \mathbf{b}$, $\mathbf{q} = \mathbf{A}^T\mathbf{x}_k$).

for $i = 1 : m$

$$d \leftarrow \omega(b_i - \sigma x_i - \mathbf{q}^T \mathbf{a}_i) / (\sigma + \|\mathbf{a}_i\|^2)$$

$$\mathbf{q} \leftarrow \mathbf{q} + \mathbf{a}_i d$$

$$x_i \leftarrow x_i + d$$

end

end Algorithm 4

After completing the loop in Algorithm 4, \mathbf{q} contains $\mathbf{A}^T\mathbf{x}_{k+1}$ and \mathbf{x} contains \mathbf{x}_{k+1} , assuming \mathbf{x} initially stores \mathbf{x}_k . Note that this SOR iteration requires $2\text{nnz}(\mathbf{A})$ multiplications and additions.

8. Conclusions. In Table 1 we give work estimates for various iterative schemes studied in this paper and for the prototype system (1.1). We assume that the product $\mathbf{A}\mathbf{A}^T$ is not formed, and that the computations utilize \mathbf{A} by itself. Table 1 only gives the matrix-vector product work; for Algorithm 3 and symmetric schemes such as CGM, TMRES/SSOR, CGM/SSOR, and TMRES/Jacobi, each iteration involves a small number (between 5 and 10) of additional vector-vector or scalar-vector products. For schemes based on the nonsymmetric Algorithm 2, there are about $(k-1)(k+2)$ additional vector-vector or scalar-vector products.

The numerical experiments in this paper show that iterative schemes may converge slowly when the matrix is nearly singular. In the TMRES approach, we start with a convergent splitting $\mathbf{C} = \mathbf{S} - \mathbf{T}$ and compute an orthonormal basis for a Krylov space generated by $\mathbf{M} = \mathbf{S}^{-1}\mathbf{T}$. Since TMRES is algebraically equivalent to GMRES preconditioned by \mathbf{S}^{-1} , TMRES converges in theory if and only if GMRES preconditioned by \mathbf{S}^{-1} converges. We observed that when $\rho(\mathbf{S}^{-1}\mathbf{T}) < 1$, a small dimensional Krylov space often contains a good approximation to the solution of a nearly singular linear system. For the prototype system (1.1), TMRES/SSOR and CGM/SSOR exhibited similar (rapid) convergence and were more efficient than unpreconditioned

schemes. The efficiency of TMRES/Jacobi was comparable to that of the SSOR-based schemes when its lower operation count $2\text{nnz}(\mathbf{A})$ was taken into account.

Acknowledgment. The author gratefully acknowledges the many thoughtful and perceptive comments of the referees and their pointers to related work.

REFERENCES

- [1] L. ADAMS AND J. L. NAZARETH, EDs., *Linear and Nonlinear Conjugate Gradient-Related Methods*, SIAM, Philadelphia, PA, 1996.
- [2] E. ALLGOWER AND K. GEORG, *Simplicial and continuation methods for approximating fixed points and solutions to systems of equations*, SIAM Rev., 22 (1980), pp. 28–85.
- [3] W. E. ARNOLDI, *The principle of minimized iteration in the solution of the matrix eigenvalue problem*, Quart. Appl. Math., 9 (1951), pp. 17–29.
- [4] J. BAGLAMA, D. CALVETTI, G. H. GOLUB, AND L. REICHEL, *Adaptively preconditioned GMRES algorithms*, SIAM J. Sci. Comput., 20 (1998), pp. 243–269.
- [5] A. BJÖRCK, *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia, PA, 1996.
- [6] A. BJÖRCK AND T. ELFVING, *Accelerated projection methods for computing pseudoinverse solutions of systems of linear equations*, BIT, 19 (1979), pp. 145–163.
- [7] P. N. BROWN AND H. F. WALKER, *GMRES on (nearly) singular systems*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 37–51.
- [8] R. H. BYRD, R. B. SCHNABEL, AND G. A. SCHULTZ, *A trust region algorithm for nonlinearly constrained optimization*, SIAM J. Numer. Anal., 24 (1987), pp. 1152–1170.
- [9] S. L. CAMPBELL, I. C. F. IPSEN, C. T. KELLEY, AND C. D. MEYER, *GMRES and the minimal polynomial*, BIT, 36 (1996), pp. 664–675.
- [10] M. R. CELIS, J. E. DENNIS, AND R. A. TAPIA, *A trust region strategy for nonlinear equality constrained optimization*, in Numerical Optimization, SIAM, Philadelphia, PA, 1985, pp. 71–82.
- [11] T. F. CHAN, *Newton-like pseudo-arclength methods for computing simple turning points*, SIAM J. Sci. Statist. Comput., 5 (1984), pp. 135–148.
- [12] P. CONCUS, G. H. GOLUB, AND D. P. O’LEARY, *A generalized conjugate gradient method for the numerical solution of elliptic partial differential equations*, in Sparse Matrix Computations, J. R. Bunch and D. J. Rose, eds., Academic Press, New York, 1976, pp. 309–332.
- [13] P. CONCUS, G. H. GOLUB, AND D. P. O’LEARY, *Numerical solution of nonlinear elliptic partial differential equations by a generalized conjugate gradient method*, Computing, 19 (1978), pp. 321–339.
- [14] J. W. DANIEL, *The conjugate gradient method for linear and nonlinear operator equations*, SIAM J. Numer. Anal., 4 (1967), pp. 10–26.
- [15] T. A. DAVIS AND W. W. HAGER, *Modifying a sparse Cholesky factorization*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 606–627.
- [16] M. EL-ALEM, *A global convergence theory for the Celis-Dennis-Tapia trust-region algorithm for constrained optimization*, SIAM J. Numer. Anal., 28 (1991), pp. 266–290.
- [17] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, 2nd ed., Johns Hopkins University Press, Baltimore, MD, 1989.
- [18] A. GREENBAUM, *Comparison of splittings used with the conjugate gradient algorithm*, Numer. Math., 33 (1979), pp. 181–194.
- [19] A. GREENBAUM, V. PTÁK, AND Z. STRAKOŠ, *Any nonincreasing convergence curve is possible for GMRES*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 465–469.
- [20] A. GREENBAUM AND Z. STRAKOŠ, *Predicting the behavior of finite precision Lanczos and conjugate gradient computations*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 121–137.
- [21] G. J. HABETLER AND E. L. WACHSPRESS, *Symmetric successive overrelaxation in solving diffusion difference equations*, Math. Comp., 15 (1961), pp. 356–362.
- [22] L. A. HAGEMAN AND D. M. YOUNG, *Applied Iterative Methods*, Academic Press, New York, 1981.
- [23] W. W. HAGER, *The LP dual active set algorithm*, in High Performance Algorithms and Software in Nonlinear Optimization, R. D. Leone, A. Murli, P. M. Pardalos, and G. Toraldo, eds., Kluwer, Dordrecht, the Netherlands, 1998, pp. 243–254.
- [24] W. W. HAGER, *Condition estimates*, SIAM J. Sci. Statist. Comput., 5 (1984), pp. 311–316.
- [25] W. W. HAGER, *Applied Numerical Linear Algebra*, Prentice-Hall, Englewood Cliffs, NJ, 1988 (available from W. W. Hager, Department of Mathematics, University of Florida, Gainesville).

- [26] W. W. HAGER, *Minimizing a Quadratic over a Sphere*, Department of Mathematics, University of Florida, Gainesville, <http://www.math.ufl.edu/~hager/papers/sphere.ps> (May 1999).
- [27] B. HENDRICKSON AND E. ROTHBERG, *Improving the Runtime and Quality of Nested Dissection Ordering*, Technical report, Sandia National Laboratories, Albuquerque, NM, 1997.
- [28] N. J. HIGHAM, *Fortran codes for estimating the one-norm of a real or complex matrix with applications to condition estimation*, ACM Trans. Math. Software, 14 (1988), pp. 381–396.
- [29] R. A. HORN AND C. R. JOHNSON, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1991.
- [30] I. C. F. IPSEN AND C. D. MEYER, *The idea behind Krylov methods*, Amer. Math. Monthly, 105 (1998), pp. 889–899.
- [31] K. C. JEA AND D. M. YOUNG, *Generalized conjugate gradient acceleration of nonsymmetrizable iterative methods*, Linear Algebra Appl., 34 (1980), pp. 159–194.
- [32] A. JENNINGS, *Influence of the eigenvalue spectrum on the convergence rate of the conjugate gradient method*, J. Inst. Math. Appl., 20 (1977), pp. 61–72.
- [33] N. KARMARKAR, *A new polynomial time algorithm for linear programming*, Combinatorica, 4 (1984), pp. 373–395.
- [34] G. KARYPIS AND V. KUMAR, *METIS: Unstructured Graph Partitioning and Sparse Matrix Ordering System*, Technical report, Department of Computer Science, University of Minnesota, Minneapolis, 1995.
- [35] D. G. LUENBERGER, *Convergence rate of a penalty-function scheme*, J. Optim. Theory Appl., 7 (1971), pp. 39–51.
- [36] D. G. LUENBERGER, *Linear and Nonlinear Programming*, Addison-Wesley, Reading, MA, 1984.
- [37] J. C. MEZA AND W. W. SYMES, *Deflated Krylov subspace methods for nearly singular linear systems*, J. Optim. Theory Appl., 72 (1992), pp. 441–457.
- [38] J. J. MORÉ, *Recent developments in algorithms and software for trust region methods*, in A. Bachem, M. Grotschel, and B. Korte, eds., *Mathematical Programming: State of the Art*, Springer-Verlag, Berlin, 1983, pp. 258–287.
- [39] N. M. NACHTIGAL, S. C. REDDY, AND L. N. TREFETHEN, *How fast are nonsymmetric matrix iterations?*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 778–795.
- [40] C. C. PAIGE AND M. A. SAUNDERS, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., 12 (1975), pp. 617–629.
- [41] C. C. PAIGE AND M. A. SAUNDERS, *LSQR: An algorithm for sparse linear equations and sparse least squares*, ACM Trans. Math. Software, 8 (1982), pp. 43–71.
- [42] M. J. D. POWELL AND Y. YUAN, *A trust region algorithm for equality constrained optimization*, Math. Programming, 49 (1991), pp. 189–211.
- [43] Y. SAAD AND M. H. SCHULTZ, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.
- [44] M. A. SAUNDERS, *Solution of sparse rectangular systems using LSQR and CRAIG*, BIT, 35 (1995), pp. 588–604.
- [45] M. A. SAUNDERS, *Computing projections with LSQR*, BIT, 37 (1997), pp. 96–104.
- [46] D. C. SORENSEN, *Implicit application of polynomial filters in a k-step Arnoldi method*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 357–385.
- [47] G. W. STEWART, *The convergence of the method of conjugate gradients at isolated extreme points in the spectrum*, Numer. Math., 24 (1975), pp. 85–93.
- [48] E. L. STIEFEL, *Kernal polynomials in linear algebra and their numerical applications*, Nat. Bur. Standards Appl. Math. Ser., 49 (1958), pp. 1–22.
- [49] R. S. VARGA, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1962.
- [50] 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.
- [51] S. J. WRIGHT, *Primal-Dual Interior-Point Methods*, SIAM, Philadelphia, PA, 1997.
- [52] D. M. YOUNG, *Iterative Solution of Large Linear Systems*, Academic Press, New York, 1971.