GMRES on (Nearly) Singular Systems

Peter N. Brown
Homer F. Walker

CRPC-TR94387
January 1994

Center for Research on Parallel Computation
Rice University
P.O. Box 1892
Houston, TX 77251-1892

GMRES ON (NEARLY) SINGULAR SYSTEMS*

PETER N. BROWN† AND HOMER F. WALKER‡

Abstract. We consider the behavior of the GMRES method for solving a linear system $Ax = b$ when A is singular or nearly so, i.e., ill-conditioned. The (near) singularity of A may or may not affect the performance of GMRES, depending on the nature of the system and the initial approximate solution. For singular A, we give conditions under which the GMRES iterates converge safely to a least-squares solution or to the pseudo-inverse solution. These results also apply to any residual minimizing Krylov subspace method that is mathematically equivalent to GMRES. A practical procedure is outlined for efficiently and reliably detecting singularity or ill-conditioning when it becomes a threat to the performance of GMRES.

Key Words. GMRES method, residual minimizing methods, Krylov subspace methods, iterative linear algebra methods, singular or ill-conditioned linear systems

AMS(MOS) subject classification. 65F10

1. Introduction.

The generalized minimal residual (GMRES) method of Saad and Schultz [8] is widely used for solving a general linear system

$$Ax = b, \qquad A \in \mathbb{R}^{n \times n}, \tag{1.1}$$

and its behavior is well-understood when A is nonsingular. Our purpose here is to examine the behavior of GMRES when A is singular or nearly so, i.e., ill-conditioned, and to formulate effective ways of detecting and handling (near) singularity in practice.

Abstractly, GMRES begins with an initial approximate solution x_0 and initial residual $r_0 = b - Ax_0$ and characterizes the kth approximate solution as $x_k = x_0 + z_k$, where z_k solves

$$\min_{z \in \mathcal{K}_k} \|b - A(x_0 + z)\|_2 = \min_{z \in \mathcal{K}_k} \|r_0 - Az\|_2. \tag{1.2}$$

Here, $\mathcal{K}_k$ is the kth Krylov subspace determined by A and r_0, defined by

$$\mathcal{K}_k \equiv \text{span}\{r_0, Ar_0, \ldots, A^{k-1}r_0\}.$$

There are a number of ways of implementing GMRES, but in each, one generates a basis of $\mathcal{K}_k$ and then replaces (1.2) by an unconstrained k-dimensional least-squares problem. We shall not be more specific about the basis generating process at this point, except to assume that it successfully generates a basis if and only if $\dim \mathcal{K}_k = k$.

We shall say that GMRES *does not break down* at the kth step if $\dim A(\mathcal{K}_k) = k$. In this case, (1.2) has a unique solution. Furthermore, since $\dim \mathcal{K}_k = k$, a basis of $\mathcal{K}_k$ is successfully generated and the k-dimensional least-squares problem also has a unique solution. This definition addresses two distinct kinds of breakdown: *rank deficiency of the least-squares problem* (1.2), which occurs

* Printed February 10, 1994. Submitted to *SIAM Journal on Matrix Analysis and Applications*, January, 1994.

† Center for Computational Sciences and Engineering, Lawrence Livermore National Laboratory, Livermore, CA 94550. This research was supported in part by the Applied Mathematical Sciences subprogram of the Office of Scientific Computing, U.S. Dept. of Energy, by Lawrence Livermore National Laboratory under contract W-7405-ENG-48.

‡ Department of Mathematics and Statistics, Utah State University, Logan, UT 84322-3900. The work of this author was supported in part by United States Air Force Office of Scientific Research Grant AFOSR-91-0294 and United States Department of Energy Grant DE-FG02-92ER25136, both with Utah State University. It was done in part during visits to the Computing and Mathematics Research Division, Lawrence Livermore National Laboratory, and the Center for Research on Parallel Computation, Rice University.

when $\dim A(\mathcal{K}_k) < \dim \mathcal{K}_k$, and *degeneracy of the dimension of* $\mathcal{K}_k$, which occurs when $\dim \mathcal{K}_k < k$. It is intended to focus on essential breakdown of the method, as opposed to breakdown associated with any particular implementation or ancillary algorithm used in it. Note that if $\dim A(\mathcal{K}_k) < k$ for some k, then $\mathcal{K}_j = \mathcal{K}_k$ for all $j \geq k$ and no further improvement is possible, even if subsequent $z_j \in \mathcal{K}_j$ are well-defined in some way.

For perspective, we recall that Proposition 2, p. 865, of [8] ensures that, if A is nonsingular, then GMRES does not break down until the solution of (1.1) has been found. Breakdown in [8, Prop. 2, p. 865] is associated specifically with breakdown of the Arnoldi process used in the GMRES implementation in [8], but the statement remains true with our definition.

In contrast to the nonsingular case, anything may happen when A is singular. Example 1.1 below shows that GMRES may break down before getting anywhere at all, or it may determine a least-squares solution[1] or the pseudo-inverse solution[2] without breaking down. Example 1.2 shows that even if a least-squares solution or the pseudo-inverse solution is reached, this may not be evident from the behavior of GMRES; indeed, GMRES may continue for a number of additional steps without breakdown (or further progress).

Example 1.1. Suppose

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad b = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad x_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Then $r_0 = (1,0)^T$ and $Ar_0 = (0,0)^T$, and GMRES breaks down at the first step. Note that x_0 is not a least-squares solution. If A is changed to

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix},$$

then, for the same b and x_0, we have $r_0 = (1,0)^T = Ar_0$, and GMRES determines without breakdown $x_1 = (1,0)^T$, which is a least-squares solution but not the pseudo-inverse solution. If we also change b to $b = (1,1)^T$, then, for the same x_0, we have $r_0 = (1,1)^T$ and $Ar_0 = (2,0)^T$, and GMRES determines without breakdown $x_1 = (1/2, 1/2)^T$, which is the pseudo-inverse solution. Note that $\dim A(\mathcal{K}_2) = 1$ in these last two cases, so GMRES breaks down at the step after the least-squares or pseudo-inverse solution has been found.

Example 1.2. For arbitrary n, let A be the "shift" operator with ones on the first subdiagonal and zeroes elsewhere. Then for $b = (1, 0, \cdots, 0)^T$ and $x_0 = (0, \cdots, 0)^T$, x_0 itself is the pseudo-inverse solution, but GMRES proceeds without breakdown (or progress) until the nth step, at which point it breaks down with $\dim A(K_n) = n - 1$.

In §2 below, we explore the theoretical behavior of GMRES and, in particular, determine circumstances in which the GMRES iterates converge without breakdown to a least-squares solution or the pseudo-inverse solution of (1.1). We also discuss the conditioning of the least-squares problem (1.2) prior to breakdown. The results in §2 apply not only to GMRES but also to any mathematically equivalent method, i.e., any method that takes steps characterized by the residual minimizing property (1.2). (See [6, §2.4] for a discussion of mathematically equivalent methods.) Thus in §2, one can think of GMRES as a *generic* minimal residual method that characterizes corrections by (1.2). In §3, we discuss further how ill-conditioning can appear in GMRES and affect its practical performance. We outline an efficient and reliable way of detecting and handling singularity or ill-conditioning when it threatens to cause breakdown or otherwise degrade the performance of the method. In §4, we outline and discuss several numerical experiments.

[1] An $x \in R^n$ for which $\|b - Ax\|_2$ is minimal.

[2] The least-squares solution x such that $\|x\|_2$ is minimal.

In the following, we denote the null-space and range of A by $\mathcal{N}(A)$ and $\mathcal{R}(A)$, respectively. We say (1.1) is *consistent* if $b \in \mathcal{R}(A)$ and denote $r_k = b - Ax_k$ for each k. As a convention, we always regard x_0 as determined without breakdown at the "0th" step and define $\mathcal{K}_0 \equiv \{0\}$.

2. Theoretical discussion. Our interest is primarily in (1.1) when A is singular, but the results also apply, as appropriate, when A is nonsingular. We note again that they are valid not only for GMRES but also for any method that determines corrections by (1.2).

The questions of interest are the following:

- Will GMRES determine a least-squares solution without breakdown?
- When has a least-squares solution been reached?
- When is a least-squares solution the pseudo-inverse solution?
- How ill-conditioned can the least-squares problem (1.2) be?

We begin with several general results.

LEMMA 2.1. *Suppose that* $\dim \mathcal{K}_k = k$ *for some* k. *Then exactly one of the following holds:*

(i) $\dim A(\mathcal{K}_k) = k - 1$;

(ii) $\dim A(\mathcal{K}_k) = k$, $\dim \mathcal{K}_{k+1} = k$, x_k *is uniquely defined and is a solution of* (1.1);

(iii) $\dim A(\mathcal{K}_k) = k$, $\dim \mathcal{K}_{k+1} = k+1$, x_k *is uniquely defined but is not a solution of* (1.1).

Proof. First, note that if $\dim \mathcal{K}_k = k$ for $k > 0$, then $\dim A(\mathcal{K}_{k-1}) = k - 1$. Indeed, in this case $r_0, Ar_0, \ldots, A^{k-1}r_0$ constitute a basis of $\mathcal{K}_k$ and, therefore, $Ar_0, \ldots, A^{k-1}r_0$ constitute a basis of $A(\mathcal{K}_{k-1})$. With this observation and the fact that $A(\mathcal{K}_{k-1}) \subseteq A(\mathcal{K}_k)$ for $k > 0$, it is clear that the assumption $\dim \mathcal{K}_k = k$ implies $k - 1 \le \dim A(\mathcal{K}_k) \le k$ for all $k \ge 0$. If $\dim A(\mathcal{K}_k) = k - 1$, then (only) conclusion (i) holds.

Suppose that $\dim A(\mathcal{K}_k) = k$. Then x_k is uniquely defined; furthermore, since $A(\mathcal{K}_k) \subseteq \mathcal{K}_{k+1}$, we have $k = \dim A(\mathcal{K}_k) \le \dim \mathcal{K}_{k+1} \le k + 1$. If $\dim \mathcal{K}_{k+1} = k$, then we must have $A(\mathcal{K}_k) = \mathcal{K}_{k+1}$ and, hence, $r_0 \in A(\mathcal{K}_k)$. It follows from (1.2) that $r_k = 0$ and x_k is a solution of (1.1); thus (only) conclusion (ii) holds. If $\dim \mathcal{K}_{k+1} = k + 1$, then $r_0 \notin A(\mathcal{K}_k)$, $r_k \neq 0$, x_k is not a solution of (1.1), and (only) conclusion (iii) holds. □

This lemma implies the following result:

THEOREM 2.2. *Apply* GMRES *to* (1.1). *Then either*

(a) *at some step,* GMRES *breaks down with rank deficiency of the least-squares problem* (1.2), *or*

(b) *equation* (1.1) *is consistent and* GMRES *determines a solution without breakdown at some step, in which case it breaks down at the next step through degeneracy of the dimension of the Krylov subspace.*

Proof. We have $\dim \mathcal{K}_0 = 0$. Assume that, for some $k \ge 0$, GMRES has proceeded to the kth step with $\dim \mathcal{K}_k = k$. If $\dim A(\mathcal{K}_k) = k - 1$, then GMRES breaks down with rank deficiency of the least-squares problem (1.2). So, assume that $\dim A(\mathcal{K}_k) = k$. If $\dim \mathcal{K}_{k+1} = k$, then (1.1) is consistent, x_k is a solution of (1.1), and GMRES breaks down through degeneracy of the dimension of the Krylov subspace at the next step. If $\dim \mathcal{K}_{k+1} = k + 1$, then x_k is not a solution and the iteration continues to the next step. One concludes that either (a) or (b) must hold. □

The alternatives of this theorem give useful insights into the eventual outcome of applying GMRES to (1.1). For example, if (1.1) is not consistent, then breakdown through rank deficiency of (1.2) will eventually occur; in practice, this may be preceded by dangerous ill-conditioning, as discussed further below. Conversely, breakdown through degeneracy of the dimension of the Krylov subspace occurs if and only if (1.1) is consistent and the solution has been found. Also, these results imply the result in [8, Prop. 2, p. 865] cited earlier: If A is nonsingular, then GMRES does not break down until the solution of (1.1) has been found. Indeed, if A is nonsingular, then GMRES cannot break down through rank deficiency of (1.2), and the second alternative must hold. However, the reader is cautioned to make inferences carefully; e.g., Example 1.1 above shows that there can be breakdown through rank deficiency in the consistent case before the solution is found.

The next result characterizes circumstances in which a least-squares solution has been reached.

LEMMA 2.3. *At the kth step,* GMRES *determines a least-squares solution of* (1.1) *without breakdown if and only if*

$$\dim A^T(\mathcal{K}_{k+1}) = \dim A(\mathcal{K}_k) = k. \tag{2.1}$$

Proof. By definition, GMRES does not break down at the kth step if and only if $\dim A(\mathcal{K}_k) = k$. Thus we need only show that x_k is a least-squares solution of (1.1) if and only if $\dim A^T(\mathcal{K}_{k+1}) = \dim A(\mathcal{K}_k)$.

From (1.2), we have that x_k is a least-squares solution of (1.1) if and only if it is possible to reach a least-squares solution of (1.1) through *some* correction in $\mathcal{K}_k$, i.e., if and only if there is some $z \in \mathcal{K}_k$ such that

$$0 = A^T[b - A(x_0 + z)] = A^T(r_0 - Az). \tag{2.2}$$

But (2.2) holds for some $z \in \mathcal{K}_k$ if and only if $A^T r_0 \in A^T A(\mathcal{K}_k)$, which is equivalent to $A^T(\mathcal{K}_{k+1}) = A^T A(\mathcal{K}_k)$. To complete the proof, we note that $\dim A^T A(\mathcal{K}_k) = \dim A(\mathcal{K}_k)$. Indeed, we clearly have $\dim A^T A(\mathcal{K}_k) \leq \dim A(\mathcal{K}_k)$. If $\dim A^T A(\mathcal{K}_k) < \dim A(\mathcal{K}_k)$, then there is a $w \in \mathcal{K}_k$ such that $Aw \neq 0$ and $A^T Aw = 0$. But then $0 = w^T A^T Aw = \|Aw\|_2^2$, which is a contradiction. □

With Lemma 2.1, one can easily extend Lemma 2.3 to conclude additionally that if (2.1) holds, then (1.1) is consistent if and only if $\dim \mathcal{K}_{k+1} = k$, i.e., GMRES breaks down at step $k+1$ through degeneracy of the dimension of the Krylov subspace.

We use Lemma 2.3 to characterize the property of A that yields the most satisfactory answers to the questions posed at the beginning of this section. This property is $\mathcal{N}(A) = \mathcal{N}(A^T)$, equivalently, $\mathcal{N}(A) = \mathcal{R}(A)^\perp$, which holds, e.g., when A is (skew) symmetric, normal, or, of course, nonsingular. It may also hold in other important cases.

THEOREM 2.4. GMRES *determines a least-squares solution of* (1.1) *without breakdown for all* b *and* x_0 *if and only if* $\mathcal{N}(A) = \mathcal{N}(A^T)$. *If* $\mathcal{N}(A) = \mathcal{N}(A^T)$ *and a least-squares solution is reached at step* k, *then* GMRES *breaks down at step* $k+1$. *Furthermore, if* (1.1) *is consistent and if* $x_0 = 0$, *then the solution reached is the pseudo-inverse solution.*

Proof. First, suppose $\mathcal{N}(A) \neq \mathcal{N}(A^T)$. One can choose b and x_0 such that $r_0 \in \mathcal{N}(A)$ and $A^T r_0 \neq 0$. Then x_0 is not a least-squares solution. Furthermore, $\dim A(\mathcal{K}_1) = 0$, so GMRES breaks down at the first step before reaching a least-squares solution.

Now assume $\mathcal{N}(A) = \mathcal{N}(A^T)$. Then for each k, we have $\dim A^T(\mathcal{K}_{k+1}) = \dim A(\mathcal{K}_{k+1})$, and (2.1) becomes

$$\dim A(\mathcal{K}_{k+1}) = \dim A(\mathcal{K}_k) = k. \tag{2.3}$$

This condition must hold for some k, $0 \leq k \leq n$, and it follows from Lemma 2.3 that GMRES determines a least-squares solution without breakdown at the kth step. Furthermore, since $\dim A(\mathcal{K}_{k+1}) = k$, GMRES breaks down at step $k+1$. If the system is consistent and $x_0 = 0$, then $x_k = z_k \in \mathcal{K}_k \subseteq \mathcal{R}(A) = \mathcal{N}(A)^\perp$; hence, x_k is the pseudo-inverse solution. □

If it is known that $\mathcal{N}(A) = \mathcal{N}(A^T)$, then Theorem 2.4 provides theoretical assurance not only that GMRES will determine a least-squares solution of (1.1) without breakdown but also that reaching it will be indicated by breakdown at the next step. If (1.1) is consistent as well, then choosing $x_0 = 0$ will yield the pseudo-inverse solution without breakdown, and reaching it will be indicated by zero residual norm.

If $\mathcal{N}(A) = \mathcal{N}(A^T)$ and (1.1) is consistent, then the least-squares problem (1.2) will remain as well-conditioned as the nature of A will allow until a solution of (1.1) is reached. Indeed, if we denote the restriction of A to $\mathcal{K}_k$ by A_k, then the appropriate condition number for (1.2) is $\kappa_2(A_k)$, which satisfies

$$\kappa_2(A_k) \equiv \frac{\|A_k\|_2}{\min\limits_{z \in \mathcal{K}_k,\ z \neq 0} \|A_k z\|_2 \big/ \|z\|_2} \leq \frac{\|A\|_2}{\min\limits_{z \in \mathcal{R}(A),\ z \neq 0} \|Az\|_2 \big/ \|z\|_2} \tag{2.4}$$

since $\mathcal{K}_k \subseteq \mathcal{R}(A)$ in the consistent case. Note that the denominator of the rightmost term in (2.4) is the smallest non-zero singular value of A, and the term itself is the condition number of the restriction of A to $\mathcal{R}(A)$. Also, recall from above that, in the consistent case, if a solution is reached at step k, then $\dim \mathcal{K}_{k+1} = k$; in particular, breakdown of GMRES at step $k+1$ occurs because of degeneracy of the dimension of the Krylov subspace, and not because of rank deficiency of the least-squares problem (1.2). These reassuring results are to be expected, for if $\mathcal{N}(A) = \mathcal{N}(A^T)$ and (1.1) is consistent, then everything reduces to the nonsingular case on $\mathcal{R}(A)$.

If $\mathcal{N}(A) = \mathcal{N}(A^T)$ but (1.1) is not consistent, then, despite the theoretical guarantee of Theorem 2.4 that GMRES will not break down, the least-squares problem (1.2) may necessarily become dangerously ill-conditioned before a least-squares solution of (1.1) is reached, regardless of the conditioning of the restriction of A to $\mathcal{R}(A)$. This is shown by Theorem 2.5 below. It is, perhaps, not surprising, because if a least-squares solution is reached at step k, then, in the inconsistent case, breakdown at step $k+1$ occurs because of rank deficiency of the least-squares problem (1.2), rather than degeneracy of the dimension of the Krylov subspace.

THEOREM 2.5. *Suppose $\mathcal{N}(A) = \mathcal{N}(A^T)$, and denote the least-squares residual for* (1.1) *by r_*. If $r_{k-1} \neq r_*$ for some k, then*

$$\kappa_2(A_k) \geq \frac{\|A_k\|_2}{\|A\|_2} \cdot \frac{\|r_{k-1}\|_2}{\sqrt{\|r_{k-1}\|_2^2 - \|r_*\|_2^2}}. \tag{2.5}$$

Proof. Note that $r_* \in \mathcal{R}(A)^\perp = \mathcal{N}(A)$ and $(r_{k-1} - r_*) \in \mathcal{R}(A) = \mathcal{N}(A)^\perp$. Then

$$\begin{aligned} \|A_k r_{k-1}\|_2 &= \|A(r_{k-1} - r_* + r_*)\|_2 = \|A(r_{k-1} - r_*)\|_2 \\ &\leq \|A\|_2 \cdot \|r_{k-1} - r_*\|_2 = \|A\|_2 \cdot \sqrt{\|r_{k-1}\|_2^2 - \|r_*\|_2^2}, \end{aligned}$$

whence

$$\frac{\|A_k r_{k-1}\|_2}{\|r_{k-1}\|_2} \leq \|A\|_2 \cdot \frac{\sqrt{\|r_{k-1}\|_2^2 - \|r_*\|_2^2}}{\|r_{k-1}\|_2}. \tag{2.6}$$

Since $r_{k-1} \in \mathcal{K}_k$, (2.5) follows from (2.6) and the definition of $\kappa_2(A_k)$ (see (2.4)). □

It is evident from (2.5) that, for an unfortunate choice of b and x_0, the least-squares problem (1.2) will become so ill-conditioned before breakdown that little or no accuracy can be expected in a solution computed in finite-precision arithmetic. Indeed, in view of (2.5), one would expect that, in many cases, the computed residual will first decrease in norm for a number of iterations and then lose accuracy and perhaps increase as a least-squares solution is approached and accuracy is degraded by increasing ill-conditioning. (This is seen in Experiment 4.2 below.) In such cases, it would clearly be desirable to terminate the iterations when approximately optimal accuracy has been reached.

We show how (2.5) can be used to derive a heuristic guideline for terminating the iterations at an approximately optimal point in finite-precision arithmetic. We make two assumptions that are reasonable but by no means the only possible assumptions; our main purpose is to demonstrate the method of derivation. The first assumption is that $\kappa_2(A_k)$ is about as small as possible, given the lower bound (2.5), i.e., that

$$\kappa_2(A_k) \approx \frac{\|A_k\|_2}{\|A\|_2} \cdot \frac{\|r_{k-1}\|_2}{\sqrt{\|r_{k-1}\|_2^2 - \|r_*\|_2^2}}.$$

The second assumption is that the computed value of r_k, denoted by $\hat{r}_k$, satisfies

$$\frac{\|\hat{r}_k - r_k\|_2}{\|r_0\|_2} \approx \mathbf{u}\kappa_2(A_k),$$

where $\mathbf{u}$ is unit rounding error. A rigorous worst-case bound on $\|\hat{r}_k - r_k\|_2/\|r_0\|_2$ would require $\mathbf{u}\kappa_2(A_k)$ multiplied by a polynomial of low degree in n and k (see [7, Ch. 5]), but this is not necessary here. With these assumptions, we have

$$\begin{aligned} \frac{\|\hat{r}_k - r_*\|_2}{\|r_0\|_2} &\leq \frac{\|\hat{r}_k - r_k\|_2}{\|r_0\|_2} + \frac{\|r_k - r_*\|_2}{\|r_0\|_2} \\ &\approx \mathbf{u}\kappa_2(A_k) + \frac{\sqrt{\|r_k\|_2^2 - \|r_*\|_2^2}}{\|r_0\|_2} \\ &\leq \mathbf{u}\kappa_2(A_k) + \frac{\sqrt{\|r_{k-1}\|_2^2 - \|r_*\|_2^2}}{\|r_0\|_2} \\ &\approx \mathbf{u}\kappa_2(A_k) + \frac{\|A_k\|_2}{\|A\|_2} \cdot \frac{\|r_{k-1}\|_2}{\|r_0\|_2} \cdot \frac{1}{\kappa_2(A_k)} \\ &= B(\kappa_2(A_k)), \end{aligned} \tag{2.7}$$

where

$$B(\kappa) \equiv \mathbf{u}\kappa + \frac{\|A_k\|_2}{\|A\|_2} \cdot \frac{\|r_{k-1}\|_2}{\|r_0\|_2} \cdot \frac{1}{\kappa}.$$

It is easily seen that B is minimized when

$$\kappa = \kappa_{\min} \equiv \sqrt{\frac{\|A_k\|_2}{\|A\|_2} \cdot \frac{\|r_{k-1}\|_2}{\|r_0\|_2} \cdot \frac{1}{\mathbf{u}}}, \tag{2.8}$$

which suggests a heuristic guideline as follows: If the iterations are terminated with $\kappa_2(A_k) \approx \kappa_{\min}$ given by (2.8), then (2.7) gives an approximate minimal bound

$$\frac{\|\hat{r}_k - r_*\|_2}{\|r_0\|_2} \leq B(\kappa_{\min}) = 2\sqrt{\frac{\|A_k\|_2}{\|A\|_2} \cdot \frac{\|r_{k-1}\|_2}{\|r_0\|_2} \cdot \mathbf{u}}. \tag{2.9}$$

This can be simplified for practical purposes by assuming that $\|A_k\|_2/\|A\|_2 \approx 1$ and $\|r_{k-1}\|_2 \approx \|\hat{r}_{k-1}\|_2$. We discuss how to monitor $\kappa_2(A_k)$ efficiently in practice in §3.

If $\mathcal{N}(A) \neq \mathcal{N}(A^T)$, then it follows from Theorem 2.4 that, for *some* b and x_0, GMRES will break down before determining a least-squares solution of (1.1). However, there is an important special case in which GMRES still reliably determines a least-squares solution, viz., that in which $\mathcal{N}(A) \cap \mathcal{R}(A) = \{0\}$ and (1.1) is consistent. This occurs, e.g., in Experiment 4.3 below.

THEOREM 2.6. *Suppose* $\mathcal{N}(A) \cap \mathcal{R}(A) = \{0\}$. *If* (1.1) *is consistent, then* GMRES *determines a solution without breakdown. If a solution is reached at step* k, *then* GMRES *breaks down at step* $k+1$ *with* $\dim \mathcal{K}_{k+1} = k$.

Proof. Since (1.1) is consistent, $r_0 \in \mathcal{R}(A)$ and $\mathcal{K}_k \subseteq \mathcal{R}(A)$ for each k. Since $\mathcal{N}(A) \cap \mathcal{R}(A) = \{0\}$, this implies that $\dim A(\mathcal{K}_k) = \dim \mathcal{K}_k$ for each k, and the theorem follows from Lemma 2.1. □

If $\mathcal{N}(A) \cap \mathcal{R}(A) = \{0\}$ and (1.1) is consistent, then $\kappa_2(A_k)$ satisfies (2.4). Note that the denominator of the rightmost term in (2.4) may be less than the smallest non-zero singular value of A if $\mathcal{N}(A) \neq \mathcal{N}(A^T)$. In any event, though, the least-squares problem (1.2) is as well-conditioned as the nature of A will allow and cannot become arbitrarily ill-conditioned through an unfortunate choice of b and x_0 before a solution is determined by GMRES. This is not surprising, since GMRES breakdown occurs because of degeneracy of the dimension of the Krylov subspace, rather than rank deficiency of the least-squares problem (1.2). When (1.1) is not consistent, breakdown must occur because of rank deficiency of (1.2), and in general we cannot expect (1.2) to remain well-conditioned, whether or not a least-squares solution is reached.

We conclude this section by noting that, in some applications, one can easily project b onto $\mathcal{R}(A)$. For example, in each of Experiments 4.2 and 4.3 below, $\mathcal{N}(A^T)$ is one-dimensional, and it is not difficult to determine a unit vector in $\mathcal{N}(A^T)$ and then to project b onto $\mathcal{N}(A^T)^\perp = \mathcal{R}(A)$. In such an application, if GMRES can be expected to behave well on a consistent system, e.g., if $\mathcal{N}(A) = \mathcal{N}(A)^T$ or $\mathcal{N}(A) \cap \mathcal{R}(A) = \{0\}$, then it is clearly desirable to project b onto $\mathcal{R}(A)$ before starting GMRES. By doing this, one can determine a least-squares solution for the original b without risking the dangerous ill-conditioning that may precede GMRES breakdown with rank deficiency of (1.2). In addition, if $\mathcal{N}(A) = \mathcal{N}(A)^T$, then one can determine the pseudo-inverse solution by taking $x_0 = 0$.

3. Practical handling of (near) singularity.

In §2, we consider the conditioning of the least-squares problem (1.2) and how it might be affected by A and perhaps b and x_0. In this section, we look further into how ill-conditioning can arise in GMRES and discuss how conditioning can be monitored efficiently in practice.

Recall from §1 that, prior to breakdown, an implementation of GMRES generates a basis of $\mathcal{K}_k$ for each k. We denote the matrix having the basis vectors as columns by $B_k \in \mathbb{R}^{n \times k}$. The kth GMRES correction z_k, which is the solution of (1.2), is not computed for each k, but when desired, it is determined by first finding y_k that solves

$$\min_{y \in R^k} \|r_0 - AB_k y\|_2. \tag{3.1}$$

and then forming $z_k = B_k y_k$. Thus ill-conditioning or singularity is a concern in GMRES only if it becomes manifested in ill-conditioning or rank deficiency of AB_k or B_k.

Sound GMRES implementations are designed so that, as much as possible, each B_k is well-conditioned regardless of the conditioning of A. For example, the standard implementation of [8] and Householder variants in [10] determine ideally conditioned B_k such that $B_k^T B_k = I_k$ (in exact arithmetic). Other implementations in [2] and [11] generate B_k that are usually well-conditioned, if not ideally conditioned. In any event, in well-constructed GMRES implementations, the conditioning of B_k does not suffer directly from ill-conditioning of A; furthermore, ill-conditioning of B_k seems likely to be reflected in ill-conditioning of AB_k. Therefore, we focus on the conditioning of AB_k here.

In practice, a reasonable course is to monitor the conditioning of AB_k and terminate the GMRES iterations if excessive ill-conditioning or rank deficiency appears. Typically, the solution of (3.1) is computed using a factorization $AB_k = Q_k R_k$, where $Q_k \in \mathbb{R}^{n\times k}$ has orthonormal columns and $R_k \in \mathbb{R}^{k\times k}$ is upper triangular. Each Q_k may be only implicitly determined, as in the implementations of [8] and [10], but each R_k is always produced explicitly. Since the conditioning of AB_k is just that of R_k, it suffices to monitor the conditioning of R_k and terminate the iterations if excessive ill-conditioning or singularity appears.

A very effective means of monitoring the conditioning of R_k is provided by *incremental condition estimation* (ICE) [3], [4]. This determines estimates of the largest and smallest singular values of each R_k in $O(k)$ arithmetic operations, given estimates of the largest and smallest singular values of R_{k-1}. Thus one can begin with $k = 1$ and use ICE to estimate incrementally the condition number of each successive R_k as k increases. Over a cycle of m GMRES steps, the total cost of estimating the condition number of each R_k, $1 \le k \le m$, is $O(m^2)$ arithmetic operations, which is negligible in most applications. A well-developed Fortran implementation of ICE is provided by auxiliary routine xLAIC1 of LAPACK [1], where x = S for single precision or x = D for double precision. This implementation was used in all of the numerical experiments reported in §4.

4. Numerical experiments. In this section, we discuss several numerical experiments that illustrate the theoretical and practical points brought out above. A standard modified Gram-Schmidt GMRES implementation, as originally outlined in [8], was used in all experiments. Recall that with this implementation, the basis matrix B_k is ideally conditioned, with $B_k^T B_k = I_k$. This implementation was augmented with routine DLAIC1 of LAPACK for monitoring conditioning of the triangular factor of AB_k as discussed above. In all experiments, we took took the zero vector to be the initial approximate solution and specified a stopping tolerance *tol* so that the GMRES iterations would terminate when $\|r_k\|_2 \le tol\|b\|_2$. Of course, there was no expectation of stopping on the basis of such a test in cases in which (1.1) was not consistent; in these cases, termination was based on other criteria noted below. All computing was done in double precision Fortran on Sun Microsytems Sparc architectures.

Experiment 4.1. This experiment, which involves a contrived problem, points up the danger of not monitoring the conditioning of AB_k and terminating when excessive ill-conditioning appears. The matrix A is from the example in [5, §6],

$$A = \begin{pmatrix} 0 & 1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ & & -1 & 0 \end{pmatrix}.$$

We assume that n is odd, in which case A is singular with

$$\mathcal{N}(A) = \text{span}\{(1,0,1,0,\cdots,0,1)^T\}.$$

Since A is skew symmetric, the conclusions of Theorem 2.4 hold, at least in exact arithmetic, and GMRES should find a least-squares solution of (1.1) without breakdown and then exhibit breakdown at the next step. In floating point arithmetic, however, GMRES produced misleading results.

We took $n = 49$, $tol = 10^{-6}$ and first ran GMRES with $b = (1/\sqrt{2}, 0, \cdots, 0, -1/\sqrt{2})^T$, for which (1.1) is consistent. GMRES safely terminated with a computed residual norm of 1.57×10^{-16} when the pseudo-inverse solution was reached at the 24th step; the largest observed condition number estimate was 12.7. We then ran GMRES with $b = (1/\sqrt{2}, 0, \cdots, 0, 1/\sqrt{2})^T$, for which (1.1) is not consistent; the least-squares residual is $\sqrt{2}/5$. In exact arithmetic, a least-squares solution would have been obtained at the 24th step, and this would have been indicated by breakdown at the 25th step in the form of rank deficiency in the least-squares problems (1.2) and (3.1). Because of rounding error, exact breakdown did not occur, nor were any arithmetic exceptions such as overflow observed. However, the condition number estimate went from 12.7 at the 24th step to 1.47×10^{16} and 1.79×10^{30} at the 25th and 26th steps, respectively. We allowed GMRES to continue, restarting every 49 steps, until it declared *successful* termination at the 185th step with a computed residual norm of 6.68×10^{-7}. Of course, this was the value of the residual norm maintained recursively by GMRES and not the true residual norm, which was 9.14×10^{12} on termination!

We also note that the GMRES implementation used in these experiments did not re-evaluate the residual and its norm "from scratch" at each restart, i.e., it did not multiply the current approximate solution by A and subtract the result from b. Instead, it updated the residual at each restart by forming a certain linear combination of the Arnoldi basis vectors generated at the previous cycle of steps. Such updating saves an A-product at each restart and is almost always a safe thing to do, unless extreme residual norm reduction is desired. In this example, however, it was not safe, and re-evaluating the residual "from scratch" at restarts would have indicated that GMRES had gone astray.

The next two experiments involve discretizations of boundary value problems for the partial differential equation

$$\Delta u + d\frac{\partial u}{\partial x_1} = f(x), \quad x = (x_1, x_2) \in \Omega \equiv [0,1] \times [0,1], \tag{4.1}$$

where d is a constant and f is a given function. In the experiments reported here, we discretized (4.1) with the usual second-order centered differences on a 100×100 mesh of equally spaced discretization points, so that the resulting linear systems were of dimension 10,000. We took $d = 10$ and preconditioned the discretized problems on the right with a fast Poisson solver from FISHPACK [9]. This preconditioner is very effective for this fairly small value of d. We took $tol = 10^{-10}$ in order to see how GMRES behaved with a tight stopping tolerance. We also stopped the iterations when the condition number estimate became greater than $1/(50\mathbf{u}) \approx 10^{14}$. In the trials outlined below, there was no need to restart GMRES; in each case, there was termination because of either sufficient residual norm reduction or excessive ill-conditioning before the maximum allowable number of iterations (50) had been reached.

In each of these two experiments, it is possible to give a simple characterization of $\mathcal{N}(A^T)$. In each, then, we first consider a b for which (1.1) is not consistent and then project it onto $\mathcal{R}(A)$ to obtain a consistent system that is effectively solved by GMRES. The result is both an approximate solution of the consistent system and an approximate least-squares solution of the original inconsistent system.

Experiment 4.2. In this experiment, we imposed periodic boundary conditions: $u(x_1, 0) = u(x_1, 1)$ and $u(0, x_2) = u(1, x_2)$ for $0 \le x_1, x_2 \le 1$. The matrix A is given as follows:

$$A = \frac{1}{h^2}\begin{pmatrix} T_m & I_m & & I_m \\ I_m & \ddots & \ddots & \\ & \ddots & \ddots & I_m \\ I_m & & I_m & T_m \end{pmatrix}, \quad T_m = \begin{pmatrix} -4 & \alpha_+ & & \alpha_- \\ \alpha_- & \ddots & \ddots & \\ & \ddots & \ddots & \alpha_+ \\ \alpha_+ & & \alpha_- & -4 \end{pmatrix} \in \mathbb{R}^{m \times m}, \tag{4.2}$$

and $m = \sqrt{n} = 100$, $h = 1/m$, and $\alpha_\pm = 1 \pm dh/2$. It is easy to verify that A is normal and that

$$\mathcal{N}(A) = \mathcal{N}(A^T) = \text{span}\{(1, 1, \cdots, 1)^T\}; \tag{4.3}$$

then Theorems 2.4 and 2.5 are applicable.

We first took b to be a discretization of $f(x) = x_1 + x_2$. For this b, (1.1) is not consistent; the least-squares residual norm is 99. GMRES began with an initial residual norm of 107.1 and

Iteration No.	GMRES Recursive Residual Norm	Computed Residual Norm	Condition No. Estimate
9	99.000000080681	99.000000080680	7.80×10^3
10	99.000000005202	99.000000005201	4.17×10^4
11	99.000000000146	99.000000000145	1.65×10^5
12	99.000000000008	99.000000000007	9.97×10^5
13	99.000000000002	99.000000000000	4.71×10^6
14	99.000000000002	99.000000000000	3.20×10^7
15	99.000000000001	99.000000000001	1.76×10^8
16	98.999999999935	99.000000000068	1.33×10^9
17	98.999999997323	99.000000002679	8.41×10^9
18	98.999999811806	99.000000188196	7.05×10^{10}
19	98.999990468226	99.000009534599	5.02×10^{11}

TABLE 1

GMRES *iterations* 9–19 *on problem* (4.1) *with periodic boundary conditions.*

terminated after 21 iterations with a condition number estimate greater than the termination value $1/(50\mathbf{u}) \approx 10^{14}$. A subset of the iterations is shown in Table 1, which gives to 14-digit accuracy both the residual norm values maintained recursively by GMRES and the residual norms computed "from scratch", as well as the condition number estimates. Note that the two norm values agree well and decrease toward the least-squares residual norm through iteration 15, but then the computed norms begin to increase while the recursive norm values continue erroneously to decrease below the least-squares residual norm. Since $\mathbf{u} \approx 2.2 \times 10^{-16}$ here, the heuristic guideline developed in §2 would have called for termination when the condition number estimate was about 10^8. Table 1 shows that this would have been a very good point at which to terminate: The computed residual norm would have been near its minimum value, and the recursive residual norm value would have still been accurate. Note the pessimism of the bound (2.9) in this case.

Using the characterization of $\mathcal{N}(A)^T$ in (4.3), we next projected the above b onto $\mathcal{R}(A)$ to obtain a consistent system. The initial residual norm was 40.82. After 17 iterations, GMRES successfully met the termination test based on $tol = 10^{-10}$ and terminated with a residual norm of 2.441×10^{-9}. No major inaccuracy was observed; the recursive residual norm value agreed with the residual norm computed "from scratch" to five significant digits. Since $\mathcal{N}(A) = \mathcal{N}(A^T)$ and the initial guess was zero, the final iterate was an approximate pseudo-inverse solution of not only the consistent system but also the inconsistent system with the original b.

Experiment 4.3. In this experiment, we imposed Neumann boundary conditions: $\partial u(x)/\partial \nu = 0$ for $x \in \partial\Omega$. The matrix A is now given by

$$(4.4)\quad A = \frac{1}{h^2}\begin{pmatrix} T_m & 2I_m & & & \\ I_m & T_m & I_m & & \\ & \ddots & \ddots & \ddots & \\ & & I_m & T_m & I_m \\ & & & 2I_m & T_m \end{pmatrix}, \quad T_m = \begin{pmatrix} -4 & 2 & & & \\ \alpha_- & -4 & \alpha_+ & & \\ & \ddots & \ddots & \ddots & \\ & & \alpha_- & -4 & \alpha_+ \\ & & & 2 & -4 \end{pmatrix} \in \mathbb{R}^{m\times m},$$

and m, h, and $\alpha_\pm$ are as in Experiment 4.2. We have $\mathcal{N}(A) = \text{span}\{(1,1,\cdots,1)^T\}$ as before, but now $\mathcal{N}(A^T) \neq \mathcal{N}(A)$. Indeed, we determine $\mathcal{N}(A^T)$ as follows: Set

$$D_m \equiv \text{diag}\,(1, 2/\alpha_-, 2\alpha_+/\alpha_-^2, \cdots, 2\alpha_+^{m-3}/\alpha_-^{m-2}, \alpha_+^{m-2}/\alpha_-^{m-2}) \in \mathbb{R}^{m\times m},$$

and define a block-diagonal matrix $D = \text{diag}(D_m, 2D_m, \cdots, 2D_m, D_m) \in I\!R^{n\times n}$. Then one can verify that DA is symmetric, and it follows that $\mathcal{N}(A^T) = \text{span}\{D(1,1,\cdots,1)^T\}$. With this characterization of $\mathcal{N}(A^T)$, one sees that $\mathcal{N}(A)\cap\mathcal{R}(A) = \{0\}$; then Theorem 2.6 applies when (1.1) is consistent.

The procedures and observations in this experiment were much like those in Experiment 4.2. We first took b to be a discretization of $f(x) = x_1 + x_2 + \sin 10x_1 \cos 10x_2 + e^{10x_1x_2}$. This gave somewhat more dramatic results than the choice of f in Experiment 4.2. For this b, (1.1) is not consistent; the least-squares residual is 5.302×10^4. GMRES began with an initial residual norm of 1.232×10^5 and terminated after 30 iterations with a condition number estimate greater than $1/(50\mathbf{u}) \approx 10^{14}$. The final computed residual norm was 6.305×10^4, which suggests that the GMRES iterates were not converging to a least-squares solution (at least not in any practical sense, given the very large condition number). We next used the characterization $\mathcal{N}(A^T) = \text{span}\{D(1,1,\cdots,1)^T\}$ to project this b onto $\mathcal{R}(A)$ and obtain a consistent system. The initial residual norm was 1.112×10^5. After 23 iterations, GMRES successfully met the termination test based on $tol = 10^{-10}$ and terminated with a residual norm of 8.716×10^{-6}. No major inaccuracy was observed; the recursive residual norm agreed with the residual norm computed "from scratch" to three significant digits. In this case, the final iterate was not a pseudo-inverse solution of either the consistent system or the inconsistent system with the original b.

5. Summary discussion. We have addressed the performance of GMRES on a linear system $Ax = b$ when A is singular or ill-conditioned. In §2, we outline theoretical results; these hold not only for GMRES but also for any mathematically equivalent method. The most extensive results hold when $\mathcal{N}(A) = \mathcal{N}(A^T)$. This condition is necessary and sufficient for GMRES to determine a least-squares solution without breakdown for *all* b and x_0. If $\mathcal{N}(A) = \mathcal{N}(A^T)$ and the system is consistent, then the condition number of the least-squares problem (1.2) remains bounded by the ratio of $\|A\|_2$ to the smallest non-zero singular value of A; if $x_0 = 0$ as well, then the solution determined by GMRES is the pseudo-inverse solution. If $\mathcal{N}(A) = \mathcal{N}(A^T)$ and the system is not consistent, then, for some b and x_0, the least-squares problem (1.2) will necessarily become dangerously ill-conditioned before a least-squares solution is reached, despite the theoretical guarantee of no breakdown. However, one may be able to use the condition number for (1.2) to determine when to terminate with nearly the best obtainable accuracy. If $\mathcal{N}(A)\cap\mathcal{R}(A) = \{0\}$ and the system is consistent, then GMRES will produce a solution without breakdown, even if $\mathcal{N}(A) \neq \mathcal{N}(A^T)$. Furthermore, in this case, the least-squares problem (1.2) remains as well-conditioned as the nature of A will allow, although the conditioning may be worse than the ratio of $\|A\|_2$ to the smallest non-zero singular value of A; in particular, the condition number for (1.2) cannot become arbitrarily large for an unfortunate choice of b and x_0.

In §3, we further discuss how ill-conditioning can arise in GMRES. In practice, the kth GMRES step is obtained by reducing (1.2) to an unconstrained k-dimensional least-squares problem, which is solved through QR factorization. In numerically sound GMRES implementations, ill-conditioning is a concern only if it becomes manifested in ill-conditioning of the upper-triangular factors. The condition numbers of these factors can be estimated very efficiently using incremental condition estimation (ICE) [3], [4].

In §4, we describe several illustrative numerical experiments.

REFERENCES

[1] E. ANDERSON, Z. BAI, C. BISCHOF, J. DEMMEL, J. DONGARRA, J. DU CROZ, A. GREENBAUM, S. HAMMARLING, A. MCKENNEY, S. OSTROUCHOV, AND D. SORENSEN, LAPACK *User's Guide*, Society for Industrial and Applied Mathematics, Philadelphia, 1992.

[2] Z. Bai, D. Hu, and L. Reichel, *A Newton basis* GMRES *implementation*, Tech. Rep. 91-03, Department of Mathematics, University of Kentucky, April 1991.

[3] C. H. Bischof, *Incremental condition estimation*, SIAM J. Matrix Anal. Appl., 11 (312-322), p. 1990.

[4] C. H. Bischof and P. T. P. Tang, *Robust incremental condition estimation*, Tech. Rep. CS-91-133, LAPACK Working Note 33, Computer Science Department, University of Tennessee, May 1991.

[5] P. N. Brown, *A theoretical comparison of the Arnoldi and* GMRES *algorithms*, SIAM J. Sci. Stat. Comput., 12 (58-78), p. 1991.

[6] R. W. Freund, G. H. Golub, and N. M. Nachtigal, *Iterative solution of linear systems*, Acta Numerica 1992, 1 (1992), pp. 57–100.

[7] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, second ed., 1989.

[8] Y. Saad and M. H. Schultz, GMRES*: a generalized minimal residual method for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 856–869.

[9] P. N. Swarztrauber and R. A. Sweet, *Efficient* fortran *subprograms for the solution of elliptic partial differential equations*, ACM Trans. Math. Soft., 5 (1979), pp. 352–364.

[10] H. F. Walker, *Implementation of the* GMRES *method using Householder transformations*, SIAM J. Sci. Stat. Comput., 9 (1988), pp. 152–163.

[11] H. F. Walker and L. Zhou, *A simpler* GMRES, Tech. Rep. 1/92/54, Mathematics and Statistics Department, Utah State University, January 1992. to appear in J. Numer. Lin. Alg. Appl.