národní
úložiště
šedé
literatury

Any Nonicreasing Convergence Curve is Possible for GMRES

Greenbaum, A.
1994

Dostupný z http://www.nusl.cz/ntk/nusl-33550

Dílo je chráněno podle autorského zákona č. 121/2000 Sb.

Tento dokument byl stažen z Národního úložiště šedé literatury (NUŠL).

Datum stažení: 21.05.2024

Další dokumenty můžete najít prostřednictvím vyhledávacího rozhraní nusl.cz .

Any Nonincreasing Convergence Curve is Possible for GMRES

Anne Greenbaum* Vlastimil Pták Zdeněk Strakoš†

September 22, 1994

Abstract

Given a nonincreasing positive sequence, $f(0) \geq f(1) \geq \ldots \geq f(n-1) > 0$, it is shown that there exists an n by n matrix A and a vector r^0 with $\|r^0\| = f(0)$ such that $f(k) = \|r^k\|$, $k = 1, \ldots, n-1$, where r^k is the residual at step k of the GMRES algorithm applied to the linear system $Ax = b$, with initial residual $r^0 = b - Ax^0$. Moreover, the matrix A can be chosen to have any desired eigenvalues.

1 Introduction

The GMRES algorithm [2] is a popular iterative technique for solving large sparse nonsymmetric (non-Hermitian) linear systems. Let A be an n by n nonsingular matrix and b an n-dimensional vector (both may be complex). To solve a linear system $Ax = b$, given an initial guess x^0 for the solution, the algorithm constructs successive approximations x^k, $k = 1, 2, \ldots$, from the affine spaces

$$x^0 + \text{span}\{r^0, Ar^0, \ldots, A^{k-1}r^0\}, \tag{1}$$

*Courant Institute of Mathematical Sciences, 251 Mercer St., New York, NY 10012. This work was supported by NSF contract INT 9218024. Part of this work was performed while visiting the Institute of Computer Science, Czech Academy of Sciences.

†Institute of Computer Science, Czech Academy of Sciences, Pod vod. věží 2, 18207 Praha 8, Czech Republic.

where $r^0 \equiv b - Ax^0$ is the initial residual. The approximations are chosen to minimize the Euclidean norm of the residual vector $r^k \equiv b - Ax^k$, i.e,

$$\|r^k\| = \min_{u \in AK_k(A,r^0)} \|r^0 - u\|, \tag{2}$$

where $K_k(A, r^0) = \text{span}\{r^0, Ar^0, \ldots, A^{k-1}r^0\}$ is the k-th Krylov subspace generated by A and r^0. We call $AK_k(A, r^0)$ the k-th Krylov residual subspace.

In a previous paper [1] it was shown that any convergence curve that can be generated by the GMRES algorithm can be generated by the algorithm applied to a matrix having any desired eigenvalues. This is in marked contrast to the situation for normal matrices, where the eigenvalues of the matrix, together with the initial residual, completely determine the GMRES convergence curve. This dramatically illustrates the fact that when highly nonnormal matrices are allowed, eigenvalue information alone cannot guarantee fast convergence of GMRES.

The residual norms of successive GMRES approximations are nonincreasing, since the residuals are being minimized over a set of expanding subspaces. The question arises, however, as to whether every nonincreasing sequence of residual norms is possible for the GMRES algorithm applied to some linear system. Considering the result from [1] mentioned above, the question is formulated in the following way: Given a nonincreasing positive sequence $f(0) \geq f(1) \geq \ldots \geq f(n-1) > 0$, and a set of nonzero complex numbers $\{\lambda_1, \ldots, \lambda_n\}$, is there an n by n matrix A having eigenvalues $\lambda_1, \ldots, \lambda_n$, and an initial residual r^0 with $\|r^0\| = f(0)$, such that the GMRES algorithm applied to the linear system $Ax = b$, with initial residual r^0, generates approximations x^k such that $\|r^k\| = f(k)$, $k = 1, \ldots, n-1$? In this paper we answer this question affirmatively and show how to construct such a matrix and initial residual. Moreover, for a given convergence behavior, we characterize all the matrices and initial residuals for which GMRES generates the prescribed sequence of residual norms.

Note that the assumption $f(n-1) > 0$ means that the related GMRES porcedure does not converge to the exact solution until the step n and the dimensions of both $K_n(A, r^0)$ and $AK_n(A, r^0)$ are equal to n. Using that assumption will simplify the notation; the modification of the results to the general case is straightforward.

Throughout the paper we assume exact arithmetic.

2 Constructing a Problem with a Given Convergence Curve and any Prescribed Nonzero Eigenvalues

In this section, we construct a matrix A and a right hand side b, solving the question formulated in the introduction, without using the results from [1]. In the next section we give an alternative characterization of the solutions, which is based on [1].

We start with a simple analysis of some properties of the desired solution. Since the residual vectors generated by the GMRES algorithm applied to a linear system $Ax = b$, with initial guess x^0, are completely determined by the matrix A and the initial residual r^0, we can assume without loss of generality that the initial guess x^0 is zero and the right-hand side vector b is the initial residual. We will refer to this procedure as GMRES(A, b). Suppose that A and b represent the unknown matrix and right hand side. Let $\mathcal{W} = \{w^1, \ldots, w^n\}$ be an orthonormal basis for the Krylov residual space $AK_n(A,b)$ such that $\text{span}\{w^1, \ldots, w^j\} = AK_j(A,b)$, $j = 1, 2, ..., n$, and let W be the matrix with the orthonormal columns $(w^1, \ldots, w^n)$. From the minimization property (2) it is clear that b can be expanded as

$$b = \sum_{j=1}^{n} \langle b, w^j \rangle w^j, \tag{3}$$

where $|\langle b, w^j\rangle| = \sqrt{\|r^{j-1}\|^2 - \|r^j\|^2}$, $r^0 = b$, $\|r^n\| = 0$. Given a nonincreasing sequence $f(0) \geq f(1) \geq \ldots \geq f(n-1) > 0$, define the differences $g(k)$ by

$$g(k) = \sqrt{(f(k-1))^2 - (f(k))^2}, \quad k = 1, \ldots, n-1. \tag{4}$$

The conditions $\|b\| = f(0), \|r^j\| = f(j), j = 1, 2, ..., n-1$, will then be satisfied if the coordinates of b in the basis $\mathcal{W}$ are determined by the prescribed sequence of residual norms,

$$W^* b = (g(1), \ldots, g(n-1), f(n-1))^T. \tag{5}$$

Let $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$, $\lambda_j \neq 0$, $j = 1, 2, \ldots, n$, be a set of nonzero points in the complex plane. Consider the monic polynomial

$$z^n - \sum_{j=0}^{n-1} \alpha_j z^j = (z-\lambda_1)(z-\lambda_2)\dots(z-\lambda_n). \tag{6}$$

Clearly, $\alpha_0 \neq 0$.

Construction of the matrix A and the right hand side b is straightforward. The idea is the following. Matrix A can be considered as a linear operator on the n-dimensional Hilbert space C^n. We denote this operator by $\mathcal{A}$; its matrix representation in the standard basis $\mathcal{E} = \{e_1, \dots, e_n\}$ gives the desired matrix A,

$$\mathcal{A}^{\mathcal{E}} = A.$$

$\mathcal{A}$ is uniquely determined by its values on any set of basis vectors.

Let $\mathcal{V} = \{v^1, \dots, v^n\}$ be *any* orthonormal basis in C^n, and let V be the matrix with the orthonormal columns $(v^1, \dots, v^n)$. Let b satisfy

$$V^* b = (g(1), \dots, g(n-1), f(n-1))^T. \tag{7}$$

Since $f(n-1)$ is nonzero, the set of vectors $\mathcal{B} = \{b, v^1, \dots, v^{n-1}\}$ is linearly independent and also forms a basis for C^n. Let B be the matrix with columns $(b, v^1, \dots, v^{n-1})$. Then, the operator $\mathcal{A}$ is simply determined by the equations

$$\begin{aligned} \mathcal{A}b &\overset{def}{=} v^1 \\ \mathcal{A}v^1 &\overset{def}{=} v^2 \\ &\vdots \\ \mathcal{A}v^{n-2} &\overset{def}{=} v^{n-1} \\ \mathcal{A}v^{n-1} &\overset{def}{=} \alpha_0 b + \alpha_1 v^1 + \dots + \alpha_{n-1} v^{n-1}. \end{aligned} \tag{8}$$

Its matrix representation in the basis $\mathcal{B}$ is

$$\mathcal{A}^{\mathcal{B}} = \begin{pmatrix} 0 & \dots & 0 & \alpha_0 \\ 1 & & 0 & \alpha_1 \\ & \ddots & \vdots & \vdots \\ & & 1 & \alpha_{n-1} \end{pmatrix}, \tag{9}$$

which is the companion matrix corresponding to the set of eigenvalues Λ. Finally, the matrix A is given by

$$A = \mathcal{A}^{\mathcal{E}} = B\mathcal{A}^{\mathcal{B}}B^{-1}. \qquad (10)$$

Summarizing, we have proved the following theorem:

Theorem 2.1.

Given a nonincreasing positive sequence $f(0) \geq f(1) \geq \ldots \geq f(n-1) > 0$ and a set of nonzero complex numbers $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$, there exists a matrix A with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and a right hand side b with $\|b\| = f(0)$ such the residual vectors r^k at each step of GMRES(A, b) satisfy $\|r^k\| = f(k)$, $k = 1, 2, \ldots, n-1$.

It is obvious that the whole subject can be formulated in terms of linear operators and operator equations on a finite dimensional Hilbert space.

For any chosen orthonormal basis $\mathcal{V}$, the matrix A and the right hand side b can be constructed via (6), (9), (10) and (4), (7). In the next section, we characterize *all the matrices and right hand sides* for which GMRES generates the prescribed sequence of residual norms.

3 An Alternative Characterization of the Solution

In [1] it was shown that many different matrices can generate the same Krylov residual spaces. We start with a slightly generalized formulation of the theorem from [1].

Theorem 3.1.

Let $E_1 \subset E_2 \subset \ldots \subset E_n$ be a sequence of subspaces of C^n, where E_j is of dimension j, $j = 1, 2, \ldots, n$, and let b be any n-dimensional vector. By $\mathcal{W} = \{w^1, \ldots, w^n\}$ we denote an orthonormal basis of E_n such that *span* $\{w_1, \ldots, w_j\} = E_j$, $j = 1, 2, \ldots, n$, by W the matrix with orthonormal columns $(w^1, \ldots, w^n)$. Let $\mathcal{A}$ be any nonsingular linear operator on E_n represented by its matrix A in the standard basis $\mathcal{E}$, $A = \mathcal{A}^{\mathcal{E}}$.

Then $AK_j(A, b) = E_j$, $j = 1, 2, \ldots, n$, if and only if $\langle b, w^n \rangle \neq 0$ and the operator $\mathcal{A}$ has in the basis $\mathcal{W}$ matrix

$$\mathcal{A}^{\mathcal{W}} = R\hat{H},$$

where R is any nonsingular upper triangular matrix and

$$\hat{H} = \begin{pmatrix} 0 & \dots & 0 & 1/\langle b, w^n \rangle \\ 1 & & 0 & -\langle b, w^1 \rangle / \langle b, w^n \rangle \\ & \ddots & \vdots & \vdots \\ 0 & \dots & 1 & -\langle b, w^{n-1} \rangle / \langle b, w^n \rangle \end{pmatrix}. \tag{11}$$

Proof: The condition $AK_j(A, b) = E_j$, $j = 1, 2, \dots, n$, is equivalent to

$$A(b, w^1, w^2, \dots, w^{n-1}) = WR \tag{12}$$

for some nonsingular upper triangular matrix R. We can also write

$$A(b, w^1, \dots, w^{n-1}) = AW \begin{pmatrix} \langle b, w^1 \rangle & 1 & \dots & 0 \\ \langle b, w^2 \rangle & & \ddots & \vdots \\ \vdots & & & 1 \\ \langle b, w^n \rangle & 0 & \dots & 0 \end{pmatrix}. \tag{13}$$

Substituting this expression into (12) and solving for A finishes the proof. Note that $\hat{H}$ is the inverse of the righmost matrix in (13). □

Prior to applying Theorem 3.1. to our problem, we prove the following useful Lemma.

Lemma 3.2
For any nonsingular matrix A and orthonormal matrix Q, GMRES(QAQ^*, b) generates the same sequence of residual norms as GMRES(A, Q^*b).

Proof. Denoting by $\bar{r}^k$ the k-th residual for GMRES(QAQ^*, b), and by $\hat{r}^k$ the k-th residual for GMRES(A, Q^*b), and using (2), we can write

$$\begin{aligned}\|\bar{r}^k\| &= \min_y \|b - \Big((QAQ^*)b, \ldots, (QAQ^*)^k b\Big)\, y\| = \\ &= \min_y \|Q^*b - \Big(A(Q^*b), \ldots, A^k(Q^*b)\Big)\, y\| = \\ &= \|\hat{r}^k\|.\end{aligned}$$

□

Using Theorem 3.1. and Lemma 3.2, the solution of our question formulated in the Introduction is again very simple.

Given a sequence $f(0) \geq f(1) \geq \ldots f(n-1) > 0$, the right hand side vector b must satisfy $|\langle b, w^j\rangle| = g(j)$, $j = 1, \ldots, n-1$, where $g(j)$ is given by (4), and $|\langle b, w^n\rangle| = f(n-1)$, for some orthonormal set $\{w^1, \ldots, w^n\}$. Equivalently, if the vectors $w^1, \ldots w^n$ are scaled by the appropriate complex units, then the right hand side vector must satisfy (5). In order that $AK_j(A, b)$ span the same space as $\{w^1, \ldots, w^j\}$ for all $j = 1, \ldots, n$, it follows from Theorem 3.1 that A must be of the form $WR\hat{H}W^*$, where $\hat{H}$ is given by (11) and R is some nonsingular upper triangular matrix. Thus, *all* matrices A and right hand side vectors b for which GMRES(A, b) generates the required residual norms must be such that A is of the form $WR\hat{H}W^*$, where $\hat{H}$ is given by (11) and b satisfies (5), for some orthonormal matrix W. It follows from Lemma 3.2 that for all matrix-vector pairs A, b of this form, GMRES(A, b) does indeed generate residual vectors with the required norms.

If we take, using the notation from (4), (6)

$$R = \begin{pmatrix} 1 & 0 & \ldots & 0 & \alpha_1 + \alpha_0 g(1) \\ 0 & 1 & & 0 & \alpha_2 + \alpha_0 g(2) \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & & & 1 & \alpha_{n-1} + \alpha_0 g(n-1) \\ 0 & 0 & \ldots & 0 & \alpha_0 f(n-1) \end{pmatrix}, \tag{14}$$

then $\hat{H}R$ is a companion matrix corresponding to the eigenvalues $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$. Since the matrix $\hat{H}R$ is similar to $R\hat{H}$, it follows that, with this choice of R, the matrix $A = WR\hat{H}W^*$ has any desired eigenvalues.

Note that for the simplest choice $W = I$, $b = (g(1), g(2), \ldots, g(n-1), f(n-1))^T$, the matrices $\hat{H}$ (11) resp. R (14) are identical to the matrices B^{-1} resp. $B\mathcal{A}^{\mathcal{B}}$ from the previous section,

$$B^{-1} \equiv \hat{H} = \begin{pmatrix} 0 & 0 & \dots & 0 & 1/f(n-1) \\ 1 & 0 & \dots & 0 & -g(1)/f(n-1) \\ & \ddots & & \vdots & \vdots \\ & & 1 & 0 & -g(n-2)/f(n-1) \\ & & & 1 & -g(n-1)/f(n-1) \end{pmatrix}, \qquad (15)$$

and A is given by $R\hat{H}$. Emphasizing the fact that *any* nonincreasing convergence curve can be considered, these simple formulas form a useful tool for constructing numerical experiments.

4 Conclusions and Open Questions

The results of this paper and [1] demonstrate clearly that eigenvalues are **not** the relevant quantities in determining the behavior of GMRES for nonnormal matrices. Any nonincreasing convergence curve can be obtained with GMRES applied to a matrix having any desired eigenvalues. Different quantities on which to base a convergence analysis have been suggested by others, for example, [5], [4]. It remains an open problem to determine the most appropriate set of system parameters for describing the behavior of GMRES. Another open problem is to determine what convergence curves are possible for the *envelope* of GMRES [3]. That is, if one does not consider a particular initial residual but instead considers the worst possible initial residual for each step k, $\max_{\|r^0\|=1} \|r^k\|$, $k = 1, \dots, n-1$, where the vectors r^k are generated by GMRES(A,r^0), then the sequence of norms must again be nonincreasing, but not every nonincreasing sequence is possible. It remains an open problem to characterize the possible sequences.

References

[1] A. Greenbaum and Z. Strakoš, "Matrices that Generate the Same Krylov Residual Spaces," in *Recent Advances in Iterative Methods*, G. Golub, A. Greenbaum, and M. Luskin, eds., Springer-Verlag, New York, 1994.

[2] Y. Saad and M. H. Schultz, "GMRES: A Generalized Minimum Residual Algorithm for Solving Nonsymmetric Linear Systems," SIAM J. Sci. Stat. Comput. 7 (1986), pp. 856-869.

[3] A. Greenbaum and L.N. Trefethen, "GMRES\CR and Arnoldi\Lanczos as Matrix Approximation Problems", SIAM J. Sci. Stat. Comput., to appear.

[4] A. Greenbaum, "Norms of Functions of Matrices", Technical Report 645, NYU, 1993.

[5] L.N. Trefethen, "Pseudospectra of Matrices", D.F. Griffiths and G.A. Watson, eds., Numerical Analysis 1991, Longman, 1992.