SIAM J. NUMER. ANAL.
Vol. 34, No. 2, pp. 513–516, April 1997

005

A NOTE ON THE SUPERLINEAR CONVERGENCE OF GMRES*

IGOR MORET†

Abstract. In the paper it is shown how the rate of convergence of the GMRES method for solving a linear operator equation $(\lambda I + K)u = f$ in a Hilbert space is related to the degree of compactness of K measured by the products of its singular values.

Key words. linear equations, Krylov subspace methods, GMRES method, compact operators, singular values

AMS subject classification. 65F10

PII. S0036142993259792

1. Introduction.

Concerning Krylov subspace methods for solving the equation $(\lambda I + K)u = f$ (with $\lambda \neq 0$) in a Hilbert space, some authors have pointed out the dependence between the speed of convergence and the degree of compactness of K, measured in terms of the decay of its singular values. To our knowledge, the first results of this type were given by Vorobyev [9] and Winther [10] for the conjugate gradient (CG) method. In [10] it was shown that if K has p-summable singular values for some $p > 0$, then the sequence $\{r_n\}$ of the residuals in the CG satisfies

$$\|r_n\|^{1/n} = O(n^{-1/p}). \tag{1.1}$$

With the limitation $p \geq 2$, analogous estimates were proved in [2] for Broyden-like methods, in [4] for an iterative variant of the degenerate kernel method, and in [6] for the iterated Vorobyev method of moments. Concerning the GMRES method, Kerkhoven and Saad [3], dealing with diagonalizable operators, proved estimates which imply (1.1). Extensions of such results are contained in the recent book by Nevanlinna [5, Chap. 5]. All these analyses give a measure of the superlinear convergence in terms of the *arithmetic means* of powers of the singular values of K. In this paper, we show that actually the rate of convergence of GMRES is directly related to the products of such singular values. Further, relation (1.1) will follow as a straightforward consequence.

2. Superlinear convergence results for GMRES.

Let H be a complex separable Hilbert space, with scalar product $\langle\ ,\ \rangle$ and norm $\|\ \|$. This last notation will also be used for the norm of any bounded linear operator on H. Let $A : H \to H$ be a linear operator. Let us assume that A is invertible and consider the equation $Au = f$. For $n = 1, 2, \ldots$, we denote by $W^{(n)}$ the nth Krylov subspace generated by A and f, namely $W^{(n)} = \text{span}\{f, AF, \ldots, A^{n-1}f\}$. We set $N = \sup_n \dim(W^{(n)})$. The well-known Arnoldi's algorithm (see [8]) provides a possibly finite orthonormal system $\{t_1, t_2, \ldots, t_j, \ldots\}$ such that $W^{(n)} = \text{span}\{t_1, t_2, \ldots, t_n\}$ for each $n \leq N$. Namely, $t_1 = f/\|f\|$ and, for $j \geq 1$, $t_{j+1} = (I - P_j)At_j/\|(I - P_j)At_j\|$, where P_j is the orthogonal projection on $W^{(j)}$. If, for some j, $(I - P_j)At_j = 0$, then $N = j$ and we set $t_{N+1} = 0$. In this case $u \in W^{(N)}$, and it is the unique solution of $P_N Au = f$. Then,

*Received by the editors December 9, 1993; accepted for publication (in revised form) May 19, 1995.

http://www.siam.org/journals/sinum/34-2/25979.html

†Dipartimento di Scienze Matematiche, Università di Trieste, Piazzale Europa 1, 34127 Trieste, Italy (Moret@univ.trieste.it).

without restriction, in what follows we consider $n \le N$. Accordingly, we denote by $\{z_1, z_2, \ldots, z_j, \ldots\}$ an orthonormal system such that $AW^{(n)} = \text{span}\{z_1, z_2, \ldots, z_n\}$. Observe that $\|(I - P_n)At_n\| = \langle At_n, t_{n+1}\rangle$.

LEMMA 1. *Let $A - \lambda I : H \to H$ be compact. Then*

$$\lim_{n\to\infty} \|(I - P_n)At_n\| = \lim_{n\to\infty} \langle (A - \lambda I)t_n, t_{n+1}\rangle = 0. \tag{2.1}$$

Proof. Let N be infinite. Since $\{t_n\}, n = 1, 2, \ldots$, is an orthonormal system in H the compactness of $A - \lambda I$ implies (2.1). ∎

For simplicity, in what follows we set $A_n = P_n A_{|W^{(n)}}$; i.e., A_n denotes $P_n A$ as an operator from $W^{(n)}$ into itself.

LEMMA 2. *If A_n is singular, then $|\langle t_{n+1}, z_n\rangle| = 1$.*

Proof. If A_n is singular, then there is $w_n (\neq 0) \in AW^{(n)}$ such that

$$P_n w_n = 0.$$

Clearly $w_n \in \text{span}\{z_n\}$, then $z_n \in \text{span}\{t_{n+1}\}$. Hence the conclusion follows. ∎

LEMMA 3. *For $\lambda \neq 0$, let $A - \lambda I : H \to H$ be compact. Then A_n is eventually invertible.*

Proof. Clearly, we have

$$\begin{aligned} |\langle t_{n+1}, z_n\rangle| &= \left|\langle t_{n+1}, AP_nA^{-1}z_n\rangle\right| = \left|\langle t_{n+1}, At_n\rangle \ \langle t_n, A^{-1}z_n\rangle\right| \\ &\le \|(I - P_n)At_n\| \ \|A^{-1}\|. \end{aligned} \tag{2.2}$$

Hence by Lemma 1,

$$\lim_{n\to\infty} |\langle t_{n+1}, z_n\rangle| = 0. \tag{2.3}$$

Thus, by Lemma 2, for all n sufficiently large, A_n is invertible. ∎

Now, let us consider the GMRES algorithm. As is well known, it computes an approximation $u_n \in W^{(n)}$ of u such that $\|f - Au_n\|$ is minimized over all the elements of $W^{(n)}$ (see [8] for computational details) and thus $f - Au_n$ is orthogonal to $AW^{(n)}$. Here below the nth GMRES residual $f - Au_n$ is indicated by $r_n (r_0 = f)$.

LEMMA 4. *If A_n is singular, then $\|r_n\| = \|r_{n-1}\|$.*

Proof. Since $r_n - r_{n-1} (\in AW^{(n)})$ is orthogonal to $AW^{(n-1)}$, then $r_n - r_{n-1} \in \text{span}\{z_n\}$. In the proof of Lemma 2 we observed that if A_n is singular, then $P_n z_n = 0$. Thus $r_{n-1} = P_n r_n$ and the conclusion follows. ∎

LEMMA 5. *Assume that A_n is invertible. Let $y_n \in W^{(n)}$ be such that $A_n y_n = f$ and set $r'_n = f - Ay_n$. Then the nth GMRES approximation is*

$$u_n = \alpha_n y_n + (1 - \alpha_n)u_{n-1}, \tag{2.4}$$

where

$$\alpha_n = \|r_{n-1}\|^2 / (\|r'_n\|^2 + \|r_{n-1}\|^2).$$

Proof. Let u_{n-1} be the $(n-1)$th GMRES approximation ($u_0 = 0$). Let u_n be defined by (2.4) and $r_n = \alpha_n r'_n + (1 - \alpha_n)r_{n-1}$. In order to prove that u_n is the nth GMRES approximation we have to show that $\langle z_i, r_n\rangle = 0$ for $i = 1, \ldots, n$. Since r'_n is orthogonal to $W^{(n)}$, then $\langle z_i, r_n\rangle = 0$ for $i = 1, \ldots, n-1$. It remains to show that $\langle z_n, r_n\rangle = 0$. Since $r'_n - r_{n-1} \in AW^{(n)}$, this last requirement is equivalent to

$\langle r'_n - r_{n-1}, r_n\rangle = 0$. By the orthogonality between r'_n and r_{n-1}, the conclusion follows by straightforward computation. ∎

LEMMA 6. *The nth residual of GMRES satisfies*

$$\|r_n\| = |\langle t_{n+1}, z_n\rangle| \|r_{n-1}\| = |\langle t_{n+1}, (I - \lambda A^{-1}) z_n\rangle| \; \|r_{n-1}\| \tag{2.5}$$

for every parameter λ.

Proof. If A_n is singular, then Lemmas 2 and 4 yield (2.5) (taking into account that $\langle t_{n+1}, A^{-1} z_n\rangle = 0$). Otherwise, from (2.4) we have

$$\|r_n\|^2 = \|r'_n\|^2 \|r_{n-1}\|^2 / (\|r'_n\|^2 + \|r_{n-1}\|^2). \tag{2.6}$$

Since $r'_n - r_{n-1} \in \operatorname{span}\{z_n\}$ from the identity $\|r'_n\| = |\langle t_{n+1}, r'_n\rangle| = |\langle t_{n+1}, r'_n - r_{n-1}\rangle|$ we easily obtain

$$\|r'_n\| / (\|r'_n\|^2 + \|r_{n-1}\|^2)^{1/2} = |\langle t_{n+1}, z_n\rangle|.$$

Then (2.5) follows from (2.6). ∎

We notice that relation (2.6) was already proved, by different arguments, by Brown [1].

Now, given a bounded linear operator $T : H \to H$, for any positive integer j, we set

$$\sigma_j(T) = \{\inf \|T - T_j\| : T_j : H \to H, \ \operatorname{rank} T_j < j\}.$$

As is well known, the singular values can be defined in this way.

LEMMA 7 (See [7, p. 125]). *Let* $\{g_1, \ldots, g_n\}$ *and* $\{h_1, \ldots, h_n\}$ *be any pair of finite orthonormal systems in* H. *Then*

$$|\det[\langle h_i, T g_j\rangle]| \le \prod_{1 \le j \le n} \sigma_j(T).$$

THEOREM 1. *For* $\lambda \ne 0$, *let* $A - \lambda I : H \to H$ *be compact. Then* GMRES *converges q-superlinearly and*

$$\|r_n\| \le \left[\prod_{1 \le j \le n} \sigma_j(A - \lambda I) \sigma_j(A^{-1}) \right] \|f\|. \tag{2.7}$$

Moreover, if $\sum_{1 \le j < \infty} (\sigma_j(A - \lambda I))^p$ *is finite for* $p > 0$, *then* (1.1) *holds.*

Proof. The q-superlinear convergence, i.e., $\|r_n\| / \|r_{n-1}\| \to 0$, follows from (2.5) and (2.3). Moreover, from (2.5) we obtain

$$\|r_n\| = \left(\prod_{1 \le j \le n} |\langle r_{j+1}, (A - \lambda I) t_j\rangle \ \langle t_j, A^{-1} z_j\rangle| \right) \|f\|.$$

It is immediate to check that the matrices $[\langle t_{i+1}, (A - \lambda I) t_j\rangle]$ and $[\langle t_i, A^{-1} z_j\rangle], i, j = 1, \ldots, n$ are upper triangular. Therefore,

$$\prod_{1 \le j \le n} \langle t_{j+1}, (A - \lambda I) t_j\rangle = |\det[\langle t_{i+1}, (A - I) t_j\rangle]| \le \prod_{1 \le j \le n} \sigma_j(A - \lambda I)$$

and

$$\prod_{1\le j\le n} |\langle t_j, A^{-1}z_j\rangle| = |\det[\langle t_i, A^{-1}z_j\rangle]| \le \prod_{1\le j\le n} \sigma_j(A^{-1}).$$

Hence we get (2.7). Finally, if $\sum_{1\le j<\infty}(\sigma_j(A-\lambda I))^p$ is finite, then (1.1) follows by a straightforward application of the *geometric-arithmetic* means inequality to (2.7). $\square$

Acknowledgments. The author wishes to thank Pierpaolo Omari for the helpful discussions and an anonymous referee for the useful suggestions.

REFERENCES

[1] P. N. Brown, *A theoretical comparison of the Arnoldi and* GMRES *algorithms*, SIAM J. Sci. Statist. Comput., 12 (1991), pp. 58–78.

[2] A. Griewank, *The local convergence of Broyden-like methods on Lipschitzian problems in Hilbert spaces*, SIAM J. Numer. Anal., 24 (1987), pp. 684–705.

[3] T. Kerkhoven and Y. Saad, *On acceleration methods for coupled nonlinear elliptic system*, Numer. Math., 60 (1992), pp. 525–548.

[4] I. Moret and P. Omari, *An iterative variant of the degenerate kernel method for solving Fredholm integral equations*, Math. Nachr., 143 (1989), pp. 41–54.

[5] O. Nevanlinna, *Convergence of Iterations for Linear Equations*, Birkhäuser, Basel, Switzerland, 1993.

[6] P. Omari, *On the fast convergence of a Galerkin-like method for equations of the second kind*, Math. Z., 201 (1989), pp. 529–539.

[7] A. Pietsch, *Eigenvalues and s-numbers*, Cambridge University Press, London, 1987.

[8] 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.

[9] Y. V. Vorobyev, *Method of Moments in Applied Mathematics*, Gordon and Breach, New York, 1965.

[10] R. Winther, *Some superlinear convergence results for the conjugate gradient method*, SIAM J. Numer. Anal., 17 (1980), pp. 14–17.