On Arbitrary Convergence Behavior of the Arnoldi Method

Jurjen Duintjer Tebbens and Gérard Meurant

Abstract

In the early days of the computer it was believed that the Lanczos method is of no practical use for solving algebraic Hermitian eigenvalue problems. The observed rapid loss of orthogonality between basis vectors in finite precision prevents the method, in general, from finding an acceptable approximation of the spectrum. The opinion on the value of the Lanczos method for practice changed with the appearance of the PhD thesis of Paige [8], who recommended its use for efficient approximation of only a few eigenvalues of a large sparse matrix. Paige showed not only that loss of orthogonality poses no problem for evaluating the quality of computed eigenvalue approximations (Ritz values), but that, on the contrary, convergence of Ritz values goes hand in hand with loss of orthogonality (see [8], or also [7]). An important result needed to prove this is the persistence theorem, which states that once a Ritz value θ is at a certain distance δ from an eigenvalue, all Lanczos iterations to come will produce a Ritz value at a distance of at most δ from θ. In the terminology of [8], the Ritz value θ is stabilized to within δ. The persistence theorem is a consequence, among others, of the fact that Ritz values produced at subsequent Lanczos iterations interlace, i.e. between every pair of Ritz values there lies a Ritz value of the previous iteration.

Generalizations of the Lanczos method for non-Hermitian matrices basically consist of two classes of methods. Methods of the Bi-Lanczos type exploit short recurrences to build a pair of bi-orthogonal bases for the involved Krylov subspaces and Arnoldi type methods compute a single, in exact arithmetics orthogonal, basis with long recurrences. Methods of both classes are nowadays widely used for the solution of large sparse eigenvalue problems, especially block or band variants in combination with a robust deflation strategy. Because of the long recurrences, Arnoldi methods are in practice restarted; a popular restart strategy is implicit restarting [11], which is the main feature of one of the most efficient software packages for large non-Hermitian eigenvalue problems, ARPACK [5].

Recent convergence results of the Arnoldi method focus on properties of restarted versions, in particular versions using implicit restarts [2]. In this talk we are interested in more basic, theoretical properties of the unrestarted Arnoldi method. We are mainly motivated by the questions: Seen the ability of the Arnoldi method to find accurate approximations of eigenvalues for a large variety of problems, can a result on stabilization of Ritz values as in the Lanczos method be proven? If not so, what is the worst behavior one can expect from the Arnoldi method?

For the Lanczos method, the worst behavior was described by Scott in [9]. Given a Hermitian matrix of order n with the eigenvalues $\lambda_1 < \lambda_2 < \ldots < \lambda_n$, he constructed a perverse starting vector v such that the Ritz values at the last Lanczos iteration are

$$\frac{\lambda_1 + \lambda_2}{2}, \frac{\lambda_2 + \lambda_3}{2}, \ldots, \frac{\lambda_{n-1} + \lambda_n}{2}.$$

That is, convergence is postponed until the very last iteration and the final Ritz values are as far from the eigenvalues as can be allowed by the interlacing property. For the formulation of an analogue of this result for the Arnoldi method one first has to find a generalized interlacing property for the non-Hermitian case. Some work on this topic has been done by both numerical analysts and scientists from other area's like Lie Algebra, see e.g. [10, 6]; in [6] one finds a geometrical interpretation of an interlacing property for the Ritz values generated with normal matrices. This interlacing property can be generalized to diagonalizable matrices with considerable effort but the

general case poses serious difficulties. We will explain this briefly in our talk and then we present our main result: Given *any* set of complex numbers

$$\theta_1^{(1)}, \theta_1^{(2)}, \theta_2^{(2)}, \ldots\ldots, \theta_1^{(n-1)}, \ldots, \theta_{n-1}^{(n-1)}, \theta_1^{(n)}, \ldots, \theta_n^{(n)},$$

there exists a whole class of matrices with starting vectors such that when the Arnoldi method is applied to members of the class, it generates at the kth iteration the Ritz values $\theta_1^{(k)}, \ldots, \theta_k^{(k)}$, $k = 1, \ldots, n$. In addition, we show how to construct such matrices and starting vectors by giving a full parametrization of the class that generates $n(n+1)/2$ prescribed Ritz values.

Thus we have shown that the Arnoldi method can exhibit arbitrary convergence behavior. As a consequence, a persistence theorem is not possible for the Arnoldi method and no stabilization of Ritz values can be guaranteed. It also follows that there is no interlacing property for the Ritz values generated by general non-Hermitian matrices. Our result is similar in spirit to the results of Arioli, Greenbaum, Pták and Strakoš showing that the GMRES method can generate any convergence curve with any spectrum [4, 3, 1]. While the GMRES method is closely related to the Arnoldi method, one may suspect our result to be a straightforward consequence of [4, 3, 1]. This is not the case, although we do exploit the theory developed in [4, 3, 1]. In fact, our result implies that GMRES can generate a prescribed convergence curve not only with any spectrum but even with any distribution of the $n(n+1)/2$ Ritz values produced during the iterative process, except for one Ritz value at each iteration.

References

[1] M. Arioli, V. Pták, and Z. Strakoš, *Krylov sequences of maximal length and convergence of GMRES*, BIT, 38 (1998), pp. 636–643.

[2] C. A. Beattie, M. Embree, and D. C. Sorensen, *Convergence of polynomial restart Krylov methods for eigenvalue computations*, SIAM Rev., 47 (2005), pp. 492–515 (electronic).

[3] 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.

[4] A. Greenbaum and Z. Strakoš, *Matrices that generate the same Krylov residual spaces*, in Recent advances in iterative methods, vol. 60 of IMA Vol. Math. Appl., Springer, New York, 1994, pp. 95–118.

[5] R. B. Lehoucq, D. C. Sorensen, and C. Yang, *ARPACK users' guide*, vol. 6 of Software, Environments, and Tools, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1998. Solution of large-scale eigenvalue problems with implicitly restarted Arnoldi methods.

[6] S. M. Malamud, *Inverse spectral problem for normal matrices and the Gauss-Lucas theorem*, Trans. Amer. Math. Soc., 357 (2005), pp. 4043–4064 (electronic).

[7] G. Meurant and Z. Strakoš, *The lanczos and conjugate gradient algorithms in finite precision arithmetic*, 15 (2006), pp. 471–542.

[8] C. C. Paige, *The computation of eigenvalues and eigenvectors of very large sparse matrices*, PhD thesis, University of London, 1971.

[9] D. S. Scott, *How to make the Lanczos algorithm converge slowly*, Math. Comp., 33 (1979), pp. 239–247.

[10] N. Shomron and B. N. Parlett, *Linear algebra meets Lie algebra: the Kostant-Wallach theory*, Linear Algebra Appl., 431 (2009), pp. 1745–1767.

[11] D. C. Sorensen, *Implicit application of polynomial filters in a k-step Arnoldi method*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 357–385.