On Arbitrary Convergence Behavior of the Arnoldi Method

Jurjen Duintjer Tebbens and Gérard Meurant

Institute of Computer Science,
Academy of Sciences of the Czech Republic (first author)
tebbens@cs.cas.cz,gerard.meurant@gmail.com

Abstract

This poster shows that arbitrary convergence behavior is possible for the Arnoldi method and gives two parametrizations of the class of matrices with initial Arnoldi vectors that generates prescribed Ritz values. The second parametrization enables to prove that any GMRES convergence curve is possible with any prescribed Ritz values, provided the stagnation case is treated appropriately.

1. Arbitrary convergence behavior of the Arnoldi method

Our results are inspired by the theory developed by Arioli, Greenbaum, Pták and Strakoš in [7], [8] and [3], which resulted in a parametrization of all matrices and right hand sides with prescribed spectrum of the matrix and prescribed convergence of the GMRES method [16] (see [3, Theorem 2.1,Corollary 2.4]). The GMRES method for linear systems and the Arnoldi method for eigenvalues [1] being closely related, a natural question is whether a result on arbitrary convergence behavior of the Arnoldi method can be proved. By arbitrary convergence behavior of the Arnoldi method, we mean the ability to prescribe *all* Ritz values from the very first until the very last iteration (we do not address convergence of Ritz vectors).

In a recent paper, Parlett and Strang proved there is a unique upper Hessenberg matrix with the entry one along the subdiagonal such that all leading principal submatrices have arbitrary prescribed eigenvalues (the Ritz values) [15, Theorem 3]. We here give a characterization of this unique matrix which shows how it is constructed from the prescribed Ritz values. In the sequel we will denote by

$$\begin{array}{rl} \mathcal{R} = \{ & \rho_1^{(1)}, \\ & (\rho_1^{(2)}, \rho_2^{(2)}), \\ & \cdots \\ & (\rho_1^{(n-1)}, \ldots, \rho_{n-1}^{(n-1)}), \\ & (\lambda_1, \ldots\ldots\ldots\ldots, \lambda_n) \quad \} \end{array} \tag{1}$$

a set of tuples of complex numbers representing $(n+1)n/2$ arbitrary Ritz values for a Hessenberg matrix of size n. By $C^{(k)}$ we denote the companion matrix of the polynomial with roots $\rho_1^{(k)}, \ldots, \rho_k^{(k)}$ (that is with roots $\lambda_1, \ldots, \lambda_n$ for $C^{(n)}$).

Theorem 1 *Given the arbitrary Ritz values in (1) with companion matrices $C^{(k)}$, define the unit upper triangular matrix*

$$U = I_n - \begin{bmatrix} 0 & C^{(1)}e_1 & \vdots & & \\ & & C^{(2)}e_2 & \cdots & \vdots \\ & & & & C^{(n-1)}e_{n-1} \\ & & & & 0 \end{bmatrix}. \tag{2}$$

Then the unique upper Hessenberg matrix $H(\mathcal{R})$ with the entry one along the subdiagonal such that the kth leading principal submatrix has eigenvalues $\rho_1^{(k)}, \ldots, \rho_k^{(k)}$ is

$$H(\mathcal{R}) = U^{-1}C^{(n)}U.$$

The similarity transformation

$$\text{diag}\,(1, \sigma_1, \ldots, \Pi_{j=1}^{n-1}\sigma_j)\, H(\mathcal{R}) \left(\text{diag}\,(1, \sigma_1, \ldots, \Pi_{j=1}^{n-1}\sigma_j)\right)^{-1}$$

for positive entries $\sigma_1, \sigma_2, \ldots, \sigma_{n-1}$ does not change the eigenvalues of the leading principal submatrices of $H(\mathcal{R})$, but gives a Hessenberg matrix with arbitrary positive subdiagonal $[\sigma_1, \sigma_2, \ldots, \sigma_{n-1}]$. This immediately gives

Corollary 2 *The Arnoldi method applied to the matrix A and the initial unit Arnoldi vector b generate the prescribed Ritz values $\mathcal{R}$ if and only if $b = Ve^1$ and A has the form*

$$V\text{diag}\,(1, \sigma_1, \ldots, \Pi_{j=1}^{n-1}\sigma_j) H(\mathcal{R}) \text{diag}\,(1, \sigma_1^{-1}, \ldots, \Pi_{j=1}^{n-1}\sigma_j^{-1}) V^*$$

for a unitary matrix V and positive numbers $\sigma_1, \sigma_2, \ldots, \sigma_{n-1}$.

Thus convergence behavior of Ritz values generated in the Arnoldi method can be arbitrary for non-normal matrices. The parametrization in Corollary 2 may be useful for convergence analysis of versions of Arnoldi used in practice, e.g. implicitly restarted Arnoldi with polynomial shifts [4]; in particular it may help to better understand (and avoid) cases where Arnoldi with exact shifts fails, see, e.g. [5]. It also shows there is no interlacing property for the Hessenberg matrices generated by the Arnoldi method. Without an interlacing property, important properties of the Lanczos method for Hermitian eigenproblems [9], like the persistence theorem or stabilization of Ritz values (see, e.g., [12, 13, 14] or [11]), need not hold anymore. More on the absence of interlacing with general matrices (Hessenberg or not) can be found in [15]; for interlacing properties of normal Hessenberg matrices see also [6], [2], [10].

We next give an alternative parametrization of arbitrary Arnoldi behavior which reveals the relation with the parametrization of arbitrary GMRES behavior in [3, Theorem 2.1,Corollary 2.4].

Theorem 3 *Consider the $(n+1)n/2$ arbitrary values in (1) with the condition that $(\lambda_1, \ldots, \lambda_n)$ contains no zero number. If A is a matrix of order n and b a unit n-dimensional vector, then the following assertions are equivalent:*

1. *The Hessenberg matrix generated by the Arnoldi process applied to A and initial Arnoldi vector b has eigenvalues $\lambda_1, \ldots, \lambda_n$, and $\rho_1^{(j)}, \ldots, \rho_j^{(j)}$ are the eigenvalues of its jth leading principal submatrix for all $j = 1, \ldots, n-1$.*
2. *The matrix A is of the form*

$$A = WYC^{(n)}Y^{-1}W^* \tag{3}$$

and $b = Wh$, where W is a unitary matrix and Y is of the form $Y = \begin{bmatrix} h & \begin{matrix} R \\ 0 \end{matrix} \end{bmatrix}$. R is the unique upper triangular matrix satisfying

$$R^*R = \hat{U}_{n-1}^*\hat{U}_{n-1} + \hat{u}\hat{u}^*, \quad e_j^T R^{-*}\hat{u} \geq 0, \quad j = 1, \ldots, n-1$$

for the partitioning

$$\text{diag}\,(1, \sigma_1, \ldots, \Pi_{j=1}^{n-1}\sigma_j)\, U^{-1} = \begin{bmatrix} 1 & \hat{u}^* \\ 0 & \hat{U}_{n-1} \end{bmatrix},$$

of the inverse of the matrix U in (2) scaled with positive numbers $\sigma_1, \ldots, \sigma_{n-1}$. The entries of $h = [\eta_1, \ldots, \eta_n]^T$ are

$$[\eta_1, \ldots, \eta_{n-1}]^T = R^{-*}\hat{u}, \qquad \eta_n = \sqrt{1 - \|R^{-*}\hat{u}\|^2}.$$

2. Arbitrary convergence behavior of the Arnoldi and the GMRES methods for the same pair $\{A, b\}$

The only freedom in the parametrization of Theorem 3 is in the unitary matrix W and the positive numbers $\sigma_1, \ldots, \sigma_{n-1}$. Can we choose these positive numbers such that we prescribe, in addition to Ritz values, also GMRES residual norms ? Recall that the parametrization of arbitrary GMRES behavior has the same form (3), except that the nonsingular upper triangular matrix R contained in Y is arbitrary. The vector h contains the prescribed convergence curve.

Theorem 4 *Let the $(n+1)n/2$ prescribed values in (1) be such that $(\lambda_1, \ldots, \lambda_n)$ contains no zero number and let n positive numbers*

$$1 = f(0) \geq f(1) \geq \cdots \geq f(n-1) > 0,$$

be such that $f(k-1) = f(k)$ if and only if the k-tuple $(\rho_1^{(k)}, \ldots, \rho_k^{(k)})$ contains a zero number. Let the matrix U in (2) be partitioned as

$$U = \begin{bmatrix} 1 & c_0^* \\ 0 & U_{n-1} \end{bmatrix}.$$

If A is a square matrix of size n and b is a unit n-dimensional vector, then the following assertions are equivalent:

1. *The GMRES method applied to A and right hand side b with zero initial guess yields residuals*

$$\|r^{(j)}\| = f(j), \quad j = 0, \ldots, n-1,$$

A has eigenvalues $\lambda_1, \ldots, \lambda_n$, and $\rho_1^{(j)}, \ldots, \rho_j^{(j)}$ are the eigenvalues of the jth leading principal submatrix of the generated Hessenberg matrix for all $j = 1, \ldots, n-1$.

2. *The matrix A is of the form*

$$A = WYC^{(n)}Y^{-1}W^*$$

and $b = Wh$, where W is any unitary matrix and Y is given by $Y = \begin{bmatrix} h & \begin{matrix} R \\ 0 \end{matrix} \end{bmatrix}$, $h = [\eta_1, \ldots, \eta_n]^T$ being the vector

$$\eta_j = (f(j-1)^2 - f(j)^2)^{1/2}, \quad j < n, \quad \eta_n = f(n-1),$$

and R being the nonsingular upper triangular matrix

$$R = R_h^{-1}D_u U_{n-1}^{-1},$$

where R_h is the upper triangular factor of the Cholesky decomposition $I - [\eta_1, \ldots, \eta_{n-1}]^T[\eta_1, \ldots, \eta_{n-1}] = R_h R_h^*$, and D_u is a nonsingular diagonal matrix such that

$$D_u^* R_h^{-T}\hat{h} = -c_0.$$

Therefore, in general, converging Ritz values need not imply accelerated convergence speed in the GMRES method. The only restriction Ritz values put on GMRES is that a zero Ritz value implies stagnation in the corresponding iteration. Note, however, that for particular matrices, e.g. matrices close to normal, the bounds derived by Van der Vorst and Vuik in [17] suggest that as soon as eigenvalues of such matrices are sufficiently well approximated by Ritz values, GMRES from then on converges at least as fast as for a related system in which these eigenvalues are missing.

Future work includes investigation of what happens to Theorems 3 and 4 when A is normal, the question whether only a few Ritz values can be prescribed and investigation whether our theory gives some insight on what is a good Arnoldi starting vector b for non-normal matrices.

Acknowledgements: The authors are indebted to Zdeněk Strakoš for initiating this work. The work of the first author is a part of the Institutional Research Plan AV0Z10300504 and it was supported by the project M100300901 of the institutional support of the ASCR.

References

[1] W. E. Arnoldi, *The principle of minimized iteration in the solution of the matrix eigenvalue problem*, Quart. Appl. Math., 9(1951), pp. 17–29.
[2] G. S. Ammar and Ch. Y. He, *On an inverse eigenvalue problem for unitary Hessenberg matrices*, Linear Algebra Appl., 218(1995), pp. 263–271.
[3] M. Arioli, V. Pták and Z. Strakoš, *Krylov Sequences of maximal length and convergence of GMRES*, B.I.T., 38(1998), pp. 636–643.
[4] Ch. A. Beattie, M. Embree and D. C. Sorensen *Convergence of polynomial restart Krylov methods for eigenvalue computations*, SIAM Rev., 47(2005), pp. 492–515.
[5] M. Embree, *The Arnoldi eigenvalue iteration with exact shifts can fail*, SIAM J. Matrix Anal. Appl., 31(2009), pp. 1–10.
[6] K. Fan and G. Pall, *Imbedding conditions for Hermitian and normal matrices*, Canad. J. Math., 9(1957), pp. 298–304.
[7] A. Greenbaum and Z. Strakoš, *Matrices that generate the same Krylov Residual Spaces*, in Recent Advances in Iterative Methods, G.H. Golub, A. Greenbaum and M. Luskin, eds, 50(1994), pp. 95–118.
[8] 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.
[9] C. Lanczos, *An iteration method for the solution of the eigenvalue problem of linear differential and integral operators*, J. Research Nat. Bur. Standards, 45(1950), pp. 255–282.
[10] S. M. Malamud, *Inverse spectral problem for normal matrices and the Gauss-Lucas theorem*, Trans. Amer. Math. Soc., 357(2005), pp. 4043–4064.
[11] G. Meurant and Z. Strakoš, *The Lanczos and conjugate gradient algorithms in finite precision arithmetic*, Acta Numerica, 15(2006), pp. 471–542.
[12] C. C. Paige, *The computation of eigenvalues and eigenvectors of very large sparse matrices*, PhD thesis, University of London, 1976.
[13] C. C. Paige, *Error analysis of the Lanczos algorithm for tridiagonalizing a symmetric matrix*, J. Inst. Math. Appl., 18(1976), pp. 341–349.
[14] C. C. Paige, *Accuracy and effectiveness of the Lanczos algorithm for the symmetric eigenproblem*, Linear Algebra Appl., 34(1980), pp. 235–258.
[15] B. N. Parlett and G. Strang, *Matrices with prescribed Ritz values*, Linear Algebra Appl., 428(2008), pp. 1725–1739.
[16] Y. Saad, M. H. Schultz, *GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems*, SIAM J. Sci. Stat. Comput., 7(1986), pp. 856–869.
[17] H. A. Van der Vorst, C. Vuik, *The superlinear convergence behaviour of GMRES*, J. Comput. Appl. Math., 48(1993), pp. 327–341.