



národní
úložiště
šedé
literatury

On the Ritz values that can be generated by the Arnoldi method

Duintjer Tebbens, Jurjen
2010

Dostupný z <http://www.nusl.cz/ntk/nusl-41602>

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

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

Datum stažení: 06.05.2024

Další dokumenty můžete najít prostřednictvím vyhledávacího rozhraní nusl.cz.

On the Ritz values that can be generated by the Arnoldi method

J. Duintjer Tebbens, G. Meurant

Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague

1 Introduction

The Arnoldi method generates approximate eigenvalues of a complex $n \times n$ matrix A by considering a starting unit vector $v \in \mathbb{C}^n$ and a decomposition

$$AV_k = V_k H_k, \quad V_k e_1 = v,$$

where $V_k^* V_k = I$ and H_k is upper Hessenberg with a positive real lower sub-diagonal. The approximate eigenvalues found in the k th iteration of the Arnoldi method, called Ritz values, are the eigenvalues of H_k . In case A is Hermitian, the method generates a matrix H_k which is tridiagonal and the method is called Lanczos method.

In his 1979 paper, Scott showed that the Lanczos method may converge very slowly in pathologic cases [4]. More precisely, given a Hermitian positive definite matrix A with the eigenvalues

$$\lambda_1 < \lambda_2 < \dots < \lambda_n,$$

he constructed a perverse starting vector v such that the eigenvalues of H_{n-1} are

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

That is, convergence may be postponed until the very last iteration.

This extended abstract deals with generalizations of Scott's result to the Arnoldi algorithm. In the case where A is normal but not Hermitian, we can easily exploit a procedure due to Ericsson to obtain the desired generalization. It turns out that in the next to last iteration one may generate any distribution of Ritz values as long as it satisfies a generalized interlacing property with respect to the spectrum of A . This is done in the next section. The last section presents further generalization for non-normal, diagonalisable matrices and discusses several other issues related to generating prescribed Ritz values in the Arnoldi method.

2 The normal case

The procedure described on page 10 of [1] leads to a method to compute a normal upper Hessenberg matrix $H \in \mathbb{C}^{n \times n}$ with given distinct eigenvalues and given spectrum of its leading principal submatrix. This spectrum, μ_1, \dots, μ_{n-1} , must satisfy what is called a generalized interlacing property in [1], namely

$$\Pi^{(r)} \equiv \frac{\prod_{j=1}^{n-1} (\lambda_r - \mu_j)}{\prod_{j=1, j \neq r}^n (\lambda_r - \lambda_j)} > 0, \quad 1 \leq r \leq n, \quad (1)$$

where $\lambda_1, \dots, \lambda_n$ are the distinct eigenvalues of H .

The following theorem shows how the procedure on page 10 of [1] can be used to construct for a given normal matrix with distinct eigenvalues an initial Arnoldi vector such that the Arnoldi method applied to the normal matrix with the initial Arnoldi vector yields prescribed Ritz values in the next to last step.

Theorem. Consider a normal matrix A with spectral decomposition $A = W \text{diag}(\lambda_1, \dots, \lambda_n) W^H$ where the eigenvalues $\lambda_1, \dots, \lambda_n$ are distinct and consider $n-1$ values μ_i such that the generalized interlacing property (1) is satisfied. Let $z \in \mathbb{C}^n$ be any vector satisfying

$$|z_r|^2 = \Pi^{(r)}, \quad 1 \leq r \leq n,$$

and let

$$\Lambda Z = Z \hat{H}$$

be the Arnoldi decomposition generated by the matrix $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ and initial vector $Ze_1 = z$, i.e. $Z^H Z = I$ and \hat{H} is upper Hessenberg with a positive real lower subdiagonal. Then the Arnoldi algorithm applied to A with initial vector

$$v \equiv W \bar{Z} e_n$$

generates in the $(n-1)$ st iteration the Ritz values μ_1, \dots, μ_{n-1} .

P r o o f : Let H be an upper Hessenberg matrix with eigenvalues $(\lambda_1, \dots, \lambda_n)$ and leading principal submatrix whose spectrum consists of the values μ_1, \dots, μ_{n-1} . Let $HX = X\Lambda$ be the spectral decomposition of H . Paige showed in [2] that

$$|X_{n,r}|^2 = \Pi^{(r)}, \quad 1 \leq r \leq n,$$

see also [3, 1, 5]. Hence there holds $X_{n,r} = e^{i\phi_r} z_r$ for values ϕ_r , $0 \leq \phi_r \leq 2\pi$, and for $1 \leq r \leq n$. If $D = \text{diag}(\phi_1, \dots, \phi_n)$ this means that we have

$$X^T e_n = Dz. \tag{2}$$

Let $P = (e_n, \dots, e_1)$ be the permutation matrix containing the columns of the identity matrix in reversed order and let \tilde{H} be the upper Hessenberg matrix defined by $\tilde{H} \equiv PH^T P$. Then from $HX = X\Lambda$ we have

$$\Lambda X^T P = X^T P \tilde{H}.$$

This is an Arnoldi decomposition generated by the matrix Λ and initial vector $X^T P e_1 = X^T e_n = Dz$. On the other hand, from $\Lambda Z = Z \hat{H}$ we obtain

$$D\Lambda Z = \Lambda DZ = DZ \hat{H},$$

i.e. an Arnoldi decomposition generated by the matrix Λ and initial vector $DZe_1 = Dz$. It follows from the uniqueness of the Arnoldi decomposition that the two decompositions are identical and $\hat{H} = \tilde{H}$. Then we obtain from $\Lambda Z = Z \hat{H}$, subsequently,

$$\begin{aligned} \Lambda Z P &= Z P P^T \tilde{H} P = Z P H^T, \\ P Z^T \Lambda &= H P Z^T, \\ P Z^T W^H W \Lambda W^H &= H P Z^T W^H, \\ A W (P Z^T)^H &= W (P Z^T)^H H. \end{aligned}$$

Because $W(PZ^T)^H$ is unitary, the last equation represents the Arnoldi decomposition generated by the matrix A and initial vector $W \bar{Z} P e_1 = W \bar{Z} e_n$. Its Hessenberg matrix has the desired Ritz values. \square

3 Further generalizations

The previous theorem can be generalized to the case where A is diagonalisable but not necessarily normal. The eigenvalues must be distinct and the Ritz values must satisfy a modification of (1).

Theorem. *Let A be diagonalisable with spectral decomposition $A = W \text{diag}(\lambda_1, \dots, \lambda_n) W^{-1}$ where the eigenvalues $\lambda_1, \dots, \lambda_n$ are distinct and consider $n-1$ values μ_i such that the nonlinear system of equations in the complex variables z_1, \dots, z_n ,*

$$\text{diag}(z_1, \dots, z_n)(W^H W)^{-1} \bar{z} = \begin{pmatrix} \Pi^{(1)} \\ \dots \\ \Pi^{(n)} \end{pmatrix} \quad (3)$$

has a solution. Let $z \in \mathbb{C}^n$ be any vector satisfying (3) and let

$$\bar{\Lambda} Z = Z \hat{H}$$

be the $(W^H W)^{-1}$ -orthogonal Arnoldi decomposition generated by the complex conjugate of the matrix $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ and initial vector $Z e_1 = \bar{z}$, i.e. $Z^H (W^H W)^{-1} Z = I$ and \hat{H} is upper Hessenberg with a positive real lower subdiagonal. Then the Arnoldi algorithm applied to A with initial vector

$$v \equiv W Z^{-H} e_n$$

generates in the $(n-1)$ st iteration the Ritz values μ_1, \dots, μ_{n-1} .

In our talk we plan to address the proof of this theorem and the geometric meaning of the interlacing properties (1) and (3). We also envisage to discuss generating prescribed Ritz values in arbitrary iteration numbers smaller than $n-1$. Finally, we will mention the problem of generating prescribed Ritz values when A is not given but, as the starting vector v , is constructed. In this case it is possible to prescribe the Ritz values of more than one iteration number.

Acknowledgements: The authors thank Zdeněk Strakoš for initiating work on this topic. The work of the first author was supported by project number KJB100300703 of the Grant Agency of the Academy of Sciences of the Czech Republic.

References

- [1] T. Ericsson: *On the eigenvalues and eigenvectors of Hessenberg matrices*. Technical report, 1990.
- [2] C.C. Paige: *The Computation of Eigenvalues and Eigenvectors of Very Large Sparse Matrices*. Ph.D. Thesis, University of London, 1971.
- [3] B.N. Parlett: *The symmetric eigenvalue problem*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1980. Prentice-Hall Series in Computational Mathematics.
- [4] D.S. Scott: *How to make the Lanczos algorithm converge slowly*. Math. Comp., 33(145), 239–247, 1979.
- [5] J.-P.M. Zemke: *Hessenberg eigenvalue-eigenmatrix relations*. Linear Algebra Appl., 414(2-3), 589–606, 2006.