University of Maryland College Park

Institute for Advanced Computer Studies TR–2001–90
Department of Computer Science TR–4316

Addendum to "A Krylov–Schur Algorithm for Large Eigenproblems"*

G. W. Stewart†

December 2001

ABSTRACT

In this addendum to an earlier paper by the author, it is shown how to compute a Krylov decomposition corresponding to an arbitrary Rayleigh-Quotient. This decomposition can be used to restart an Arnoldi process, with a selection of the Ritz vectors corresponding to the Rayleigh quotient.

*This report is available by anonymous ftp from `thales.cs.umd.edu` in the directory `pub/reports` or on the web at `http://www.cs.umd.edu/~stewart/`.

†Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742 (`stewart@cs.umd.edu`).

Addendum to "A Krylov–Schur Algorithm for Large Eigenproblems"

G. W. Stewart

ABSTRACT

In this addendum to an earlier paper by the author, it is shown how to compute a Krylov decomposition corresponding to an arbitrary Rayleigh-Quotient. This decomposition can be used to restart an Arnoldi process, with a selection of the Ritz vectors corresponding to the Rayleigh quotient.

In [3] the author introduced a decomposition of the form

$$AU = UB + ub^{\mathrm{H}}, \tag{1}$$

where A is a matrix of order n and $(U\ \ u)$ has full column rank. It was shown that the column space of $(U\ \ u)$ (called the *subspace of the decomposition*) is a (possibly restarted) Krylov subspace of A and conversely that every Krylov subspace has such a representation, so that the *Krylov decomposition* (1) is a characterization of Krylov subspaces. Arnoldi and Lanczos decompositions are special cases of Krylov decompositions.

The advantage of working with Krylov decompositions is that their subspaces remain invariant under two classes of transformations. The first, called a *similarity*, transforms the decomposition into

$$A(UW^{-1}) = (UW^{-1})(WBW^{-1}) + u(b^{\mathrm{H}}W) \equiv A\tilde{U} = \tilde{B} + u\tilde{b}^{\mathrm{H}},$$

where W is any nonsingular matrix. The second, called a *translation,* transforms the decomposition to the form

$$AU = U\tilde{B} + \tilde{u}b^{\mathrm{H}},$$

where

$$\tilde{B} = B + gb^{\mathrm{H}}, \quad \tilde{u} = \frac{u - Ug}{\gamma}, \quad \text{and} \quad \tilde{b}^{\mathrm{H}} = \gamma b^{\mathrm{H}},$$

for any vector g and any scalar $\gamma \neq 0$.

The computational algorithms in [3] were based on similarities. Translations were used primarily in the derivation of the properties of Krylov decompositions. The purpose of this note is to show that translations have a computational role to play in restarting an Arnoldi process with a selection of Rayleigh–Ritz approximations to a set of eigenvectors.

The Rayleigh–Ritz method for producing these approximations does not depend on whether the subspace in question is a Krylov subspace. It can be presented in different ways. The one we give here leads most directly to the main result of this note. Let U be a basis for the subspace $\mathcal{U}$ in question and let V be such that $V^{\mathrm{H}}U$ is nonsingular. Then the matrix

$$B = (V^{\mathrm{H}}U)^{-1}V^{\mathrm{H}}AU \tag{2}$$

has the property that (μ, Uw) is an eigenpair of A, then (μ, w) is an eigenpair of B. Specifically,

$$Bw = (V^{\mathrm{H}}U)^{-1}V^{\mathrm{H}}AUw = \mu(V^{\mathrm{H}}U)^{-1}V^{\mathrm{H}}Uw = \mu w.$$

By continuity one might expect that if $\mathcal{U}$ contains an approximate eigenvector of A, then it can be found by computing an appropriate eigenpair (μ, w) of B and forming Uw. This is the essence of the Rayleigh–Ritz method (for an analysis of the method see [1]). The matrix B is called a *Rayleigh quotient* (with respect to U and V) because (2) is a generalization of the ordinary Rayleigh quotient $v^{\mathrm{H}}Au/v^{\mathrm{H}}u$.

It was observed in [3] that the matrix B in the Krylov decomposition (1) is a Rayleigh quotient. Specifically, let $(V\ \ v)^{\mathrm{H}}$ be a left inverse of $(U\ \ u)$. Then $V^{\mathrm{H}}U = I$ and $V^{\mathrm{H}}u = 0$. It follows from (1) that $B = V^{\mathrm{H}}AU$ is a Rayleigh quotient, which can be used in the Rayleigh–Ritz procedure.

In some cases, however, we may not have the freedom to choose V. For example, in the harmonic Rayleigh–Ritz method, which has superior properties for approximating interior eigenvalues [2], [4, pp. 292–294], we must take $V = (A - \kappa I)U$, where κ is near the eigenvalues of interest. The following theorem shows that although B in (1) need not be the Rayleigh quotient with respect to V there is a translated Krylov decomposition whose Rayleigh quotient is.

Theorem 1. *In the Krylov decomposition* (1) *let $V^{\mathrm{H}}U$ be nonsingular. Then with $g = (V^{\mathrm{H}}U)^{-1}V^{\mathrm{H}}u$, we have*

$$B + gb^{\mathrm{H}} = (V^{\mathrm{H}}U)^{-1}V^{\mathrm{H}}AU. \tag{3}$$

Proof. By translation,

$$AU = U(B + gb^{\mathrm{H}}) + (u - Ug)b^{\mathrm{H}} \tag{4}$$

is a Krylov decomposition. Moreover, by the definition of g, we have $V^{\mathrm{H}}(u - Ug) = 0$. Hence (3) follows on multiplying (4) by $(V^{\mathrm{H}}U)^{-1}V^{\mathrm{H}}$. ■

We have shown in effect that if B is the Rayleigh quotient of a Krylov decomposition, all other Rayleigh quotients with respect to U are rank-one modifications of B. For harmonic Ritz vectors the formulas simplify, since from (1)

$$V = (A - \kappa I)U = U(B - \kappa I) + ub^{\mathrm{H}}.$$

In practical implementations of Krylov method, $(U\ \ u)$ will be orthogonal, so that

$$V^{\mathrm{H}}U = (B-\kappa I)^{\mathrm{H}} \quad \text{and} \quad V^{\mathrm{H}}u = b.$$

Hence, $g = (B-\kappa I)^{-\mathrm{H}}b$ and the Rayleigh quotient is

$$B + (B-\kappa I)^{-\mathrm{H}}bb^{\mathrm{H}}.$$

Thus there is no need to form V explicitly. For the symmetric case this formula is due to Morgan [2].

The importance of this result, however, is not in the fact that it provides formulas for Rayleigh quotients. That could be done from the original decomposition. Instead the key fact is that Rayleigh quotient is part of the Krylov decomposition (4). This means that we can use the decomposition to restart the an Arnoldi process with selected Ritz vectors via the Krylov–Schur method described in the original paper.

Specifically, suppose that in the decomposition (1) the matrix (Uu) is orthonormal. We compute the partitioned Schur decomposition

$$\begin{pmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{pmatrix} = \begin{pmatrix} W_1^{\mathrm{H}} \\ W_2^{\mathrm{H}} \end{pmatrix} (B + gb^{\mathrm{H}})(W_1\ \ W_2),$$

where T_{11} contains the Ritz values corresponding to the Ritz vectors we wish to retain. It then follows that the Krylov decomposition

$$A(UW_1) = (UW_1)T_{11} + (u - Ug)b^{\mathrm{H}}W_1$$

is a Krylov decomposition containing those Ritz vectors. The matrix UW_1 is orthonormal, but the the vector $u - Ug$ is not orthogonal to the columns of UW_1. However, by a second translation we can orthogonalize it. The resulting decomposition is an orthogonal Krylov decomposition, which can be extended by the Arnoldi process in the usual way.

Acknowledgement

I am indebted to the Mathematical and Computational Sciences Division of the National Institute of Standards and Technology for the use of their research facilities.

References

[1] Z. Jia and G. W. Stewart. An analysis of the Rayleigh–Ritz method for approximating eigenspaces. *Mathematics of Computation*, 70:637–647, 2001.

[2] R. B. Morgan. Computing interior eigenvalues of large matrices. *Linear Algebra and Its Applications*, 154–156:289–309, 1991.

[3] G. W. Stewart. A Krylov–Schur algorithm for large eigenproblems. Technical Report TR–4127, Department of Computer Science, University of Maryland, College Park, 2000. To appear in the SIAM Journal on Matrix Analysis and Applications.

[4] G. W. Stewart. *Matrix Algorithms II: Eigensystems*. SIAM, Philadelphia, 2001.