Improving Eigenvectors in Arnoldi's Method

Zhongxiao Jia* Ludwig Elsner†

July 14, 1995

Abstract

The Ritz vectors obtained by Arnoldi's method may not be good approximations and even may not converge even if the corresponding Ritz values do. In order to improve the quality of Ritz vectors and enhance the efficiency of Arnoldi type algorithms, we propose a strategy that uses Ritz values obtained from an m-dimensional Krylov subspace but chooses modified approximate eigenvectors in an $(m+1)$-dimensional Krylov subspace. Residual norm of each new approximate eigenpair is minimal over the span of the Ritz vector and the $(m+1)$th basis vector, which is *available* when the m-step Arnoldi process is run. The resulting modified m-step Arnoldi method is better than the standard m-step one in theory and cheaper than the standard $(m+1)$-step one. Based on this strategy, we present a modified m-step restarted Arnoldi algorithm. Numerical examples show that the modified m-step restarted algorithm and its version with Chebyshev acceleration are often considerably more efficient than the standard $(m+1)$-step restarted ones.

Key words. Large unsymmetric, the m-step Arnoldi process, the m-step Arnoldi method, eigenvalue, Ritz value, eigenvector, Ritz vector, modified

AMS(MOS) subject classification. 65F15

1 Introduction

Arnoldi's method [1, 11] is suitable for computing a few selected eigenpairs of large unsymmetric matrices. It has been investigated since the 1980s; see, e.g., [3, 4, 6, 7, 8, 9, 10, 12, 13, 14].

*Department of Applied Mathematics, Dalian University of Technology, Dalian 116023, P.R. China

†Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, 33501 Bielefeld, Germany (elsner@mathematik.uni-bielefeld.de).

It is well known that the m-step Arnoldi process, as described in detail in Section 2, generates an orthonormal basis $\{v_j\}_{j=1}^m$ of the Krylov subspace $\mathcal{K}_m(v_1, A)$ spanned by $v_1, Av_1, \ldots, A^{m-1}v_1$. Here v_1 is an initial unit norm vector. The projected matrix of A onto $\mathcal{K}_m(v_1, A)$ is represented by an $m \times m$ upper Hessenberg matrix H_m. Then the m-step Arnoldi method uses the Ritz pairs $\lambda^{(m)}, \varphi^{(m)}$ of A in $\mathcal{K}_m(v_1, A)$ to approximate some eigenpairs λ, φ of A.

The convergence analysis in [4, 8, 12] gives a *necessary* condition for the convergence of Arnoldi's method. It states that the distance between the wanted φ and $\mathcal{K}_m(v_1, A)$ must approach zero as m increases. However, it is not *sufficient*, since it is proved in [4, 8] that Ritz vectors $\varphi^{(m)}$ may not be good approximations in theory for a general unsymmetric matrix A even though these distances approach zero and the corresponding $\lambda^{(m)}$ converge. For more details and numerical experiments we refer to [7].

This suggests that we seek new strategies to improve the quality of $\varphi^{(m)}$ and to enhance the efficiency of Arnoldi type algorithms. The first approach is to keep $\lambda^{(m)}$ but *completely discard* $\varphi^{(m)}$. Instead, the refined approximate eigenvector is chosen in $\mathcal{K}_m(v_1, A)$, such that it minimizes the norm of the residual formed with the Ritz value. It can be computed relatively cheaply by solving an $(m+1) \times m$ singular value decomposition problem. It is shown in [7, 8] that the refined vectors converge if the corresponding Ritz values $\lambda^{(m)}$ do, so the possible non-convergence of eigenvectors is ruled out in theory. Numerical experiments have demonstrated that the resulting refined algorithms are much more efficient than the standard counterparts.

In this note, we follow a different approach. We observe that the m-step Arnoldi process generates an orthonormal basis of $\mathcal{K}_{m+1}(v_1, A)$ *rather than* of $\mathcal{K}_m(v_1, A)$, that is, the $(m+1)$th basis vector v_{m+1} of $\mathcal{K}_{m+1}(v_1, A)$ is already *available* when the m-step Arnoldi process is performed. However, although the basis vectors $\{v_j\}_{j=1}^{m+1}$ of $\mathcal{K}_{m+1}(v_1, A)$ are available in this case, we have only obtained the $m \times m$ projected matrix H_m of A onto $\mathcal{K}_m(v_1, A)$. Therefore, v_{m+1} contributes *nothing* to the wanted eigenvectors and is wasted. It only indicates that residuals of $\lambda^{(m)}, \varphi^{(m)}$ are in its direction.

We want to exploit the available v_{m+1} and propose a new strategy that uses the Ritz values $\lambda^{(m)}$ obtained from $\mathcal{K}_m(v_1, A)$ but computes modified approximate eigenvectors $\psi^{(m)}$ in $\mathcal{K}_{m+1}(v_1, A)$. These new vectors are linear combinations of $\varphi^{(m)}$ and v_{m+1}, such that residual norm of each pair $\lambda^{(m)}, \psi^{(m)}$ is minimal over the span of $\varphi^{(m)}$ and v_{m+1}. They can be calculated by solving a two-dimensional minimization problem for each eigenvector. As a result, residual norms of the new approximate eigenpairs $\lambda^{(m)}, \psi^{(m)}$ are at least as small as those of $\lambda^{(m)}, \varphi^{(m)}$ and often dramatically smaller. The total extra cost of this strategy is one matrix-vector multiplication plus only $2r$ inner products, where r is the number of wanted eigenpairs and it is typically much smaller than m in practice. So the resulting modified m-step Arnoldi method is cheaper than the standard $(m+1)$-step one, which requires besides one matrix-vector multiplication at least $(m+1)$ inner products.