Journal of Computational Mathematics, Vol.17, No.3, 1999, 257–274.

ARNOLDI TYPE ALGORITHMS FOR LARGE UNSYMMETRIC MULTIPLE EIGENVALUE PROBLEMS*[1)]

Zhong-xiao Jia
(*Department of Applied Mathematics, Dalian University of Technology, Dalian 116024, China*)

Abstract

As is well known, solving matrix multiple eigenvalue problems is a very difficult topic. In this paper, Arnoldi type algorithms are proposed for large unsymmetric multiple eigenvalue problems when the matrix A involved is diagonalizable. The theoretical background is established, in which lower and upper error bounds for eigenvectors are new for both Arnoldi's method and a general perturbation problem, and furthermore these bounds are shown to be optimal and they generalize a classical perturbation bound due to W. Kahan in 1967 for A symmetric. The algorithms can adaptively determine the multiplicity of an eigenvalue and a basis of the associated eigenspace. Numerical experiments show reliability of the algorithms.

Key words: Arnoldi's process, Large unsymmetric matrix, Multiple eigenvalue, Diagonalizable, Error bounds

1. Introduction

The Lanczos algorithm[20] is a very powerful tool for extracting a few extreme eigenvalues and associated eigenvectors of large symmetric matrices[4,5,22]. Since the 1980's, considerable attention has been paid to generalizing it to large unsymmetric problems. One of its generalizations is Arnoldi's method[1,25]. It can be used to compute outer part of the spectrum and corresponding eigenvectors[10,11,24,25,26,28]. In order to improve overall performance, Saad[27] suggested to use it in conjunction with the Chebyshev iteration. There are other variants available; see, e.g. [12, 13, 16, 17, 19, 24, 28].

To apply Arnoldi's algorithm and its variants to practical problems, one must account for the following difficulty[2,3,6,8]:

Difficulty* *Multiple eigenvalues are a common occurrence.*

In the symmetric case, Parlett and Scott[21] used the Lanczos algorithm with selective orthogonalization to solve Difficulty*. Their algorithm maintains the semi-orthogonality among the Lanczos vectors so as to avoid the occurrence of spurious

* Received April 27, 1995.

1) Supported by the China State Major Key Project for Basic Researches, the National Natural Science Foundation of China, the Doctoral Program of the China State Educational Commission, the Foundation for Excellent Young Scholars of the Ministry of Education, the Foundation of Returned Scholars of China and the Natural Science Foundation of Liaoning Province.

eigenvalues and determines the multiplicities of the required eigenvalues and the associated eigenspaces by restarting. The key idea is that, before restarting, a new initial vector is orthogonalized with respect to all the converged eigenvectors until the eigenspace associated with a multiple eigenvalue is found.

In the unsymmetric case, the situation becomes much more complicated. The strategy of restarting[21] cannot solve Difficulty* since the eigenvectors of unsymmetric matrices are, in general, not mutually orthogonal just as those of symmetric matrices are. The mutual orthogonality of eigenvectors forms the basis of the algorithm in [12]. Theoretically speaking, a simple simulation of the idea used in [21] suggests that before restarting we use Arnoldi's method with a new initial vector orthogonal to all the left eigenvectors of the matrix A associated with all the converged right eigenvectors. Proceeding in such a way, we can find the multiplicities of the required eigenvalues and determine the associated eigenspaces. However, an easy analysis[23] shows that Arnoldi's method is inefficient for computing the left eigenvectors of A. Of course, one can apply Arnoldi's method to A^H, the conjugate transpose of A, to get the left eigenvectors of A, while this doubles the amount of computation.

In order to deal with Difficulty*, generalized block Lanczos methods are studied in [10, 14]. They can be used to compute outer part of the spectrum and corresponding eigenvectors, up to a multiplicity equal to block size when A is diagonalizable. However, if the multiplicities of the required eigenvalues are bigger than block size, the block algorithms themselves are not able to determine the multiplicity of an eigenvalue and the associated eigenspace. Therefore, to be able to detect the multiplicity, the block algorithms have to combine with other techniques in practice.

In this paper, we design Arnoldi type algorithms for solving Difficulty* when A is diagonalizable. As is seen from [10, 14], the proposed idea is important not only in its own right but also indispensable for the block Arnoldi method when block size is smaller than or equal to the multiplicities of the required eigenvalues.

In Section 2, we introduce the notation used and go through the underlying Arnoldi algorithm; in Section 3, assuming that A is diagonalizable, we present the theoretical background of the Arnoldi type algorithms to be proposed in Section 4. Some of the results, i.e. theoretical error bounds for eigenvectors, are new for both Arnoldi's method and a general perturbation problem; in Section 4 we present two Arnoldi type algorithms to solve Difficulty*; in Section 5, we discuss some implementations of the algorithms; in Section 6, we report three numerical examples to show reliability of the algorithms, followed by some concluding remarks in Section 7.

2. The Underlying Arnoldi Algorithm

2.1. Notation

Throughout the paper, assume that A is an $N \times N$ real diagonalizable matrix, $N \gg 1$ and it has M distinct eigenvalues λ_i, where the multiplicities of λ_i are d_i, $i = 1, 2, \cdots, M$. Under this assumption let $\mathcal{P}_i$ be the d_i-dimensional eigenspace associated with λ_i and the columns of $\Phi_{id_i} = (\varphi_{i1}, \varphi_{i2}, \cdots, \varphi_{id_i})$ form a basis of $\mathcal{P}_i$, where $\|\varphi_{ij}\| = 1$

and $\|\cdot\|$ denotes the 2-norm. Let $\Psi_{id_i} = (\psi_{i1}, \psi_{i2}, \cdots, \psi_{id_i})$, where ψ_{ij} are the left eigenvectors associated with λ_i such that $\psi_{ik}^H \varphi_{ij} = \delta_{kj}$, $k, j = 1, 2, \cdots, d_i$. Here the superscript H denotes the conjugate transpose of a matrix, a vector and a scalar and δ_{kj} the Kronecker delta. We want to compute a few, say r, special eigenvalues, e.g. those λ_i with largest (smallest) real parts or largest moduli, and determine the corresponding multiplicities d_i and bases of $\mathcal{P}_i$, $i = 1, 2, \cdots, r$.

We denote by $\mathcal{K}_m(v, A)$ the Krylov subspace spanned by $v, Av, \cdots, A^{m-1}v$, by $\pi_m(v)$ the orthogonal projector onto $\mathcal{K}_m(v, A)$ and by $\theta(x, y)$ the acute angle between two nonzero vectors x and y. Let $P_i = \sum_{j=1}^{d_i} \varphi_{ij} \psi_{ij}^H$ be the eigenprojectors associated with λ_i, $i = 1, 2, \cdots, M$.

2.2. The Underlying Arnoldi Algorithm

A basic Arnoldi process can be described as follows.

Algorithm 1. Arnoldi's process

1. Start: Choose a real initial vector v_1, $\|v_1\| = 1$, and the steps m of Arnoldi's process.

2. Iterate: For $l = 1, 2, \cdots, m$ do

2.1. $w = Av_l$.

2.2. For $j = 1, 2, \cdots, l$ do

$h_{jl} = v_j^H A v_l$,

$w = w - h_{jl} v_j$.

2.3. $h_{l+1l} = \|w\|$.

2.4. $v_{l+1} = w / h_{l+1l}$.

This algorithm generates an orthonormal basis $\{v_l\}_1^m$ of $\mathcal{K}_m(v_1, A)$. Define the matrix $V_m = (v_1, v_2, \cdots, v_m)$. In the basis $\{v_l\}_1^m$, the restriction of A to $\mathcal{K}_m(v_1, A)$ is represented by an upper Hessenberg matrix $H_m = V_m^H A V_m$ with the entries h_{jl} computed by Algorithm 1. The m eigenvalues $\lambda_i^{(m)}$, called the Ritz values of A in $\mathcal{K}_m(v_1, A)$, of H_m are used to approximate m eigenvalues of A, and the corresponding approximate eigenvectors $\varphi_i^{(m)}$, called the Ritz vectors of A in $\mathcal{K}_m(v_1, A)$, are computed by

$$\varphi_i^{(m)} = V_m y_i^{(m)}, \tag{1}$$

where $y_i^{(m)}$ are eigenvectors of H_m associated with $\lambda_i^{(m)}$.

How good some approximations are can be measured in terms of an a-posteriori bound

$$\|r_i^{(m)}\| = \|(A - \lambda_i^{(m)} I)\varphi_i^{(m)}\| = h_{m+1m} \mid e_m^H y_i^{(m)} \mid, \tag{2}$$

in which $e_m = (0, 0, \cdots, 0, 1)^H$. (2) can be used as a stopping criterion which checks cheaply the size of the residual without computing $\varphi_i^{(m)}$ by (1).

In terms of the a-priori theoretical analysis[11,12,14,25,26], $\|r_i^{(m)}\|$ has been proved to converge to zero as m increases if the behavior of Ritz pairs $\lambda_i^{(m)}, \varphi_i^{(m)}$, $i = 1, 2, \cdots, m$ is

not too bad[17], and it usually tends to zero first for the right-most and left-most eigenvalues and the corresponding eigenvectors. If the eigenproblem of A is ill conditioned, m might be quite large in order to make Algorithm converge.

Since Algorithm 1 has to save all the vectors v_l generated previously and its amount of computation increases quadratically with steps, the above Arnoldi algorithm usually has to be used iteratively in practice. In order to improve efficiency, an iterative Arnoldi algorithm can be accelerated by the Chebyshev iteration[9,27]. The purpose of using the Chebyshev iteration is to amplify the components of an initial vector in the directions of the required eigenvectors. The resulting algorithm is called the Arnoldi-Chebyshev algorithm.

3. Theoretical Background

We now establish the theoretical background of Arnoldi type algorithms for solving Difficulty* when A is diagonalizable.

Let us choose a set of vectors $v_1^{(1)}, v_1^{(2)}, \cdots, v_1^{(k)}$. Then each $v_1^{(j)}, 1 \le j \le k$, can be expanded in the eigenspaces $\mathcal{P}_i$, $i = 1, 2, \cdots, M$ as

$$v_1^{(j)} = b_{j1}\varphi_{i1} + b_{j2}\varphi_{i2} + \cdots + b_{jd_i}\varphi_{id_i} + u_i^{(j)}, \tag{3}$$

$$u_i^{(j)} \in \mathcal{P}_1 \oplus \cdots \oplus \mathcal{P}_{i-1} \oplus \mathcal{P}_{i+1} \oplus \cdots \oplus \mathcal{P}_M, \ 1 \le j \le k, \tag{4}$$

where $\oplus$ denotes the direct sum.

Let

$$B_k = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1d_i} \\ b_{21} & b_{22} & \cdots & b_{2d_i} \\ \vdots & \vdots & \cdots & \vdots \\ b_{k1} & b_{k2} & \cdots & b_{kd_i} \end{pmatrix} \tag{5}$$

Assume that the matrix B_k is row full rank for $k \le d_i$. Obviously, B_k is row rank deficient when $k > d_i$. We rewrite the above $v_1^{(j)}$ as

$$v_1^{(j)} = \beta_j \tilde{\varphi}_{ij} + u_i^{(j)}, \ u_i^{(j)} \in \mathcal{P}_1 \oplus \cdots \oplus \mathcal{P}_{i-1} \oplus \mathcal{P}_{i+1} \oplus \cdots \oplus \mathcal{P}_M, \tag{6}$$

where $\tilde{\varphi}_{ij}$, $j = 1, 2, \cdots, k$ are also unit norm eigenvectors associated with λ_i and β_j normalizing factors. Under the assumption on B_k, just as $\{\varphi_{ij}\}_{j=1}^{d_i}$, $\{\tilde{\varphi}_{ij}\}_{j=1}^{d_i}$ is also a basis of $\mathcal{P}_i$, and for $k > d_i$, $\tilde{\varphi}_{ij}$, $j = d_i + 1, \cdots, k$ belong to the span of $\{\tilde{\varphi}_{ij}\}_{j=1}^{d_i}$.

Define $\tilde{\Phi}_{ik} = (\tilde{\varphi}_{i1}, \tilde{\varphi}_{i2}, \cdots, \tilde{\varphi}_{ik})$ and $\tilde{\Psi}_{ik} = (\tilde{\psi}_{i1}, \tilde{\psi}_{i2}, \cdots, \tilde{\psi}_{ik})$, where $\tilde{\psi}_{ij}$ are the left eigenvectors associated with λ_i such that $\tilde{\psi}_{ik}^H \tilde{\varphi}_{ij} = \delta_{kj}$.

Following a result of [25], we have the following a-priori residual bound.

Theorem 1. *Let*

$$\gamma_f m(v_1^{(j)}) = \|\pi_m(v_1^{(j)}) A (I - \pi_m(v_1^{(j)}))\|.$$

Then for the eigenpairs λ_i, $\tilde{\varphi}_{ij}$, $1 \le i \le M$, $1 \le j \le k$, we have

$$\|(A_m(v_1^{(j)}) - \lambda_i I)\tilde{\varphi}_{ij}\| \le \sqrt{\gamma_m^2(v_1^{(j)}) + |\lambda_i|^2}\,\|(I - \pi_m(v_1^{(j)}))\tilde{\varphi}_{ij}\|, \tag{7}$$

where $A_m(v_1^{(j)}) = \pi_m(v_1^{(j)})A\pi_m(v_1^{(j)})$.

Therefore, whether some of the eigenpairs $\lambda_{ij}^{(m)}, \tilde{\varphi}_{ij}^{(m)}$, $i = 1, 2, \cdots, m$, $j = 1, 2, \cdots, k$ of $A_m(v_1^{(j)})$ are good approximations to $\lambda_i, \tilde{\varphi}_{ij}$ or not heavily depends on the behavior of $\|(I - \pi_m(v_1^{(j)}))\tilde{\varphi}_{ij}\|$. It is shown[10,14,25,26] that the right-hand side of (7) approaches zero as m increases, usually first for outer part of the spectrum. Furthermore, it is shown[10,14,17] that some Ritz pairs obtained by Arnoldi's method starting with $v_1^{(j)}$, $1 \leq j \leq k$ converge to outer part of the spectrum and corresponding eigenvectors provided the eigenproblem of $A_m(v_1^{(j)})$ is not too ill conditioned and the right-hand side of (7) tends to zero. According to the previous statement, we can find a basis of $\mathcal{P}_i$ approximately provided d_i initial vectors $v_1^{(j)}$ are chosen such that the assumption on B_k of (5) is satisfied. For $k > d_i$, $\tilde{\varphi}_{ij}^{(m)}$, $j = d_i + 1, \cdots, k$ will approximately belong to the span of $\{\tilde{\varphi}_{ij}^{(m)}\}_{j=1}^{d_i}$.

More concisely, we can determine the multiplicity d_i and a basis of $\mathcal{P}_i$ based on the following strategy: First, choose an initial vector $v_1^{(1)}$, and use Arnoldi's method to compute a converged eigenpair $\lambda_{i1}^{(m)}, \tilde{\varphi}_{i1}^{(m)}$ to $\lambda_i, \tilde{\varphi}_{i1}$. Then, start with a second initial vector $v_1^{(2)}$, and use Arnoldi's method to compute a converged eigenpair $\lambda_{i2}^{(m)}, \tilde{\varphi}_{i2}^{(m)}$ to $\lambda_i, \tilde{\varphi}_{i2}$. Having $\tilde{\varphi}_{i1}^{(m)}, \tilde{\varphi}_{i2}^{(m)}$, we decide whether or not the matrix $\tilde{\Phi}_{i2}^{(m)} = (\tilde{\varphi}_{i1}^{(m)}, \tilde{\varphi}_{i2}^{(m)})$ is column rank approximately deficient. If not, this shows that $d_i \geq 2$, and we take the columns of $\tilde{\Phi}_{i2}^{(m)}$ as a basis of $\mathcal{P}_i$; otherwise, $d_i = 1$ and $\tilde{\varphi}_{i1}^{(m)}$ is a basis of $\mathcal{P}_i$. We stop, else continue restarting with a new initial vector $v_1^{(3)}$. Proceeding this way until at some $(k+1)$th restarting, we have that the matrix $\tilde{\Phi}_{ik}^{(m)} = (\tilde{\varphi}_{i1}^{(m)}, \tilde{\varphi}_{i2}^{(m)}, \cdots, \tilde{\varphi}_{ik}^{(m)})$ is column full rank, but $\tilde{\Phi}_{ik+1}^{(m)} = (\tilde{\varphi}_{i1}^{(m)}, \tilde{\varphi}_{i2}^{(m)}, \cdots, \tilde{\varphi}_{ik+1}^{(m)})$ is approximately column rank deficient. This means that $d_i = k$ and the columns of $\tilde{\Phi}_{ik}^{(m)}$ are an approximate basis of $\mathcal{P}_i$.

To give a quantitative analysis for the above assertion, we have to study error bounds for the approximate eigenpairs $\lambda_{ij}^{(m)}, \tilde{\varphi}_{ij}^{(m)}$, $i = 1, 2, \cdots, r$, $j = 1, 2, \cdots, k$, in terms of the a-posteriori computable residual norms $\|\tilde{r}_{ij}^{(m)}\| = \|(A - \lambda_{ij}^{(m)} I)\tilde{\varphi}_{ij}^{(m)}\|$. They can be rewritten as

$$(A - \tilde{r}_{ij}^{(m)} \tilde{\varphi}_{ij}^{(m)H})\tilde{\varphi}_{ij}^{(m)} = \lambda_{ij}^{(m)} \tilde{\varphi}_{ij}^{(m)}.$$

Thus, the approximate eigenpairs $\lambda_{ij}^{(m)}, \tilde{\varphi}_{ij}^{(m)}, i = 1, 2, \cdots, r$, $j = 1, 2, \cdots, k$ are the exact eigenpairs of the matrices $(A - \tilde{r}_{ij}^{(m)} \tilde{\varphi}_{ij}^{(m)H})$, in which the perturbation matrices are $\tilde{r}_{ij}^{(m)} \tilde{\varphi}_{ij}^{(m)H}$. We then have the following result [30, p.69].

Theorem 2. *Assume* $\|\tilde{r}_{ij}^{(m)}\|$, $1 \leq i \leq r$, $1 \leq j \leq k$, *to be small enough. Then*

$$|\lambda_i - \lambda_{ij}^{(m)}| \leq \|\tilde{\psi}_{ij}\| \cdot \|\tilde{r}_{ij}^{(m)}\| + O(\|\tilde{r}_{ij}^{(m)}\|^2). \tag{8}$$

From Theorem 2, we see that if λ_i is not very ill conditioned, i.e. $\|\tilde{\psi}_{ij}\|$ is not very large, then $\lambda_{ij}^{(m)}$ is a good approximation to λ_i provided that $\|\tilde{r}_{ij}^{(m)}\|$ is small, so that $\lambda_{ij}^{(m)}$ obtained by different $v_1^{(j)}$ are numerically equal for the same i.

Before deriving error bounds for approximate eigenvectors $\tilde{\varphi}_{ij}^{(m)}$, we need the following lemma[10,11].

Lemma 1. *Let $x_1, x_2, \cdots, x_s$ be s vectors and $\alpha_1, \alpha_2, \cdots, \alpha_s$ be s scalars, and define the matrix $X = (x_1, x_2, \cdots, x_s)$. Then*

$$\left\| \sum_{j=1}^{s} \alpha_j x_j \right\| \geq \frac{1}{\inf_D \ diag. \ \kappa(XD)} \min_{1 \leq j \leq s} \mid \alpha_j \mid \left\| \sum_{j=1}^{s} x_j \right\|, \tag{9}$$

where D's are $s \times s$ nonsingular diagonal matrices and $\kappa(XD)$ denotes the condition number of XD, which equals the ratio of the largest and smallest singular values of XD.

Theorem 3. *Let $\lambda_{ij}^{(m)}, \tilde{\varphi}_{ij}^{(m)}$ be a Ritz pair of A in $\mathcal{K}_m(v_1^{(j)}, A)$ and $g_{i,m}^{(j)} = \min_{l \neq i} \mid \lambda_l - \lambda_{ij}^{(m)} \mid$. Define the matrix*

$$X_{ij}^{(m)} = (P_1 \tilde{\varphi}_{ij}^{(m)}, \cdots, P_{i-1} \tilde{\varphi}_{ij}^{(m)}, P_{i+1} \tilde{\varphi}_{ij}^{(m)}, \cdots, P_M \tilde{\varphi}_{ij}^{(m)}).$$

Then

$$\sin \theta(\tilde{\varphi}_{ij}, \tilde{\varphi}_{ij}^{(m)}) \leq \|(I - P_i)\tilde{\varphi}_{ij}^{(m)}\| \leq \frac{\inf\limits_{D \text{diag.}} \kappa(X_{ij}^{(m)} D)(1 + \|P_i\|)}{g_{i,m}^{(j)}} \|\tilde{r}_{ij}^{(m)}\|. \tag{10}$$

If A is symmetric, then

$$\sin \theta(\tilde{\varphi}_{ij}, \tilde{\varphi}_{ij}^{(m)}) = \|(I - P_i)\tilde{\varphi}_{ij}^{(m)}\| \leq \frac{\|\tilde{r}_{ij}^{(m)}\|}{g_{i,m}^{(j)}}. \tag{11}$$

Proof. It is well known that

$$P_l P_i = \delta_{il} P_l, \tag{12}$$

$$\sum_{l=1}^{M} P_l = I. \tag{13}$$

Hence we have

$$(A - \lambda_{ij}^{(m)} I)\tilde{\varphi}_{ij}^{(m)} = (A - \lambda_{ij}^{(m)} I) \sum_{l=1}^{M} P_l \tilde{\varphi}_{ij}^{(m)} = \sum_{l=1}^{M} (\lambda_l - \lambda_{ij}^{(m)}) P_l \tilde{\varphi}_{ij}^{(m)}.$$

Premultiplying the two hand sides of the above relation by $I - P_i$, we obtain

$$(I - P_i)(A - \lambda_{ij}^{(m)} I)\tilde{\varphi}_{ij}^{(m)} = \sum_{l \neq i} (\lambda_l - \lambda_{ij}^{(m)}) P_l \tilde{\varphi}_{ij}^{(m)}.$$

From Lemma 1, we have

$$\|(I - P_i)(A - \lambda_{ij}^{(m)} I)\tilde{\varphi}_{ij}^{(m)}\| \geq \frac{g_{i,m}^{(j)}}{\inf\limits_{D \text{ diag.}} \kappa(X_{ij}^{(m)} D)} \left\| \sum_{l \neq i} P_l \tilde{\varphi}_{ij}^{(m)} \right\|.$$

On the other hand,

$$\|(I-P_i)(A-\lambda_{ij}^{(m)}I)\tilde{\varphi}_{ij}^{(m)}\| \le \|I-P_i\|\|(A-\lambda_{ij}^{(m)}I)\tilde{\varphi}_{ij}^{(m)}\| \le (1+\|P_i\|)\|\tilde{r}_{ij}^{(m)}\|.$$

By (13), we obtain

$$\|(I-P_i)\tilde{\varphi}_{ij}^{(m)}\| = \left\|\sum_{l\neq i} P_l\tilde{\varphi}_{ij}^{(m)}\right\|.$$

Therefore, combining the above three relations gives

$$\|(I-P_i)\tilde{\varphi}_{ij}^{(m)}\| \le \frac{\inf\limits_{D \text{ diag.}} \kappa(X_{ij}^{(m)}D)(1+\|P_i\|)}{g_{i,m}^{(j)}}\|\tilde{r}_{ij}^{(m)}\|.$$

According to the definition of $\theta(x,y)$, we have

$$\sin\theta(\tilde{\varphi}_{ij},\tilde{\varphi}_{ij}^{(m)}) = \min_{\alpha}\|\tilde{\varphi}_{ij}^{(m)}-\alpha\tilde{\varphi}_{ij}\| \le \|\tilde{\varphi}_{ij}^{(m)}-P_i\tilde{\varphi}_{ij}^{(m)}\| = \|(I-P_i)\tilde{\varphi}_{ij}^{(m)}\|.$$

Thus, (10) holds.

If A is symmetric, its eigenvectors are mutually orthogonal. In this case, $\|I-P_i\| = 1$, $\inf\limits_{D \text{ diag.}} \kappa(X_{ij}^{(m)}D) = 1$ and $\sin\theta(\tilde{\varphi}_{ij},\tilde{\varphi}_{ij}^{(m)}) = \|(I-P_i)\tilde{\varphi}_{ij}^{(m)}\|$. Thus, (11) holds. □

We comment that (11) is a well known result; see [22, Ch.11, p.222] and [29, Ch.5, p.250] for more general formulations. We have now generalized it to (10) for the unsymmetric case.

From Theorem 3, the sensitivity of $\tilde{\varphi}_{ij}$ depends on the sensitivity of the eigenvalue problem of A, the gap $\|\tilde{r}_{ij}^{(m)}\|$ and $\|\tilde{r}_{ij}^{(m)}\|$. If $\|\tilde{r}_{ij}^{(m)}\|$ is small, $\tilde{\varphi}_{ij}^{(m)}$ is a good approximation to $\tilde{\varphi}_{ij}$ provided the eigenvalue problem of A is not too ill conditioned and $g_{i,m}^{(j)}$ is not small, while if A is symmetric, the sensitivity of $\tilde{\varphi}_{ij}$ depends mainly on $g_{i,m}^{(j)}$ and $\|\tilde{r}_{ij}^{(m)}\|$. For example, if $\|\tilde{r}_{ij}^{(m)}\| = 10^{-8}$, $\|\tilde{r}_{ij}^{(m)}\| = 1$ and $\inf\limits_{D \text{ diag.}} \kappa(X_{ij}^{(m)}D)(1+\|P_i\|) = 1000$, then $\sin\theta(\tilde{\varphi}_{ij},\tilde{\varphi}_{ij}^{(m)})$ is no more than 10^{-5}. Such a $\tilde{\varphi}_{ij}^{(m)}$) can be considered to be a good approximation to $\tilde{\varphi}_{ij}$.

Now we deviate from the topic of Arnoldi's method for a while and study a general eigenvalue problem that will be posed immediately. To this end we temporarily change the notation used. The problem is stated as follows:

Assume that A is an $N\times N$ diagonalizable matrix and it has M distinct eigenvalues λ_j, $j=1,2,\cdots,M$, and assume that a perturbed matrix $A+E$ is also diagonalizable and has the eigenpairs $\tilde{\lambda}_i, \tilde{\varphi}_i$, $i=1,2,\cdots,N$ with $\|\tilde{\varphi}_i\|=1$, where E is a small perturbation matrix. Suppose that $\tilde{\lambda}_i$, $i=1,2,\cdots,N$ are divided into M groups, in which each group approximates one eigenvalue of A. Then if $\tilde{\lambda}_i$ is used to approximate λ_j, how well does $\tilde{\lambda}_i$ approximate some eigenvector associated with λ_j?

As was pointed out by Stewart and Sun [29, Ch.5, p.229] the problem of assessing the accuracy of an approximate eigenvector in terms of a residual is very closely related

to the usual perturbation problem, and the latter can be formulated easily and derived trivially from the former. So Theorem 3 can be directly exploited to establish the following result.

Theorem 4. *Under the above assumptions and notations, let*

$$d_{i,j} = \min_{k \neq j} \mid \tilde{\lambda}_i - \lambda_k \mid, \quad i = 1, 2, \cdots, N$$

(mathematically speaking, by a continuity argument, $d_{i,j} \neq 0$ can be guaranteed once E is sufficiently small), and define the matrix X_j to be the matrix whose columns consist of the eigenvectors associated with those $\lambda_k \neq \lambda_j$. Then there exists a unit norm eigenvector, for brevity, say φ_j, associated with λ_j such that

$$\sin\theta(\varphi_j, \tilde{\varphi}_i) \leq \|(I - P_j)\tilde{\varphi}_i\| \leq \frac{\inf\limits_{D \ diag.} \kappa(X_j D)(1 + \|P_j\|)}{d_{i,j}} \|E\|. \tag{14}$$

If A is symmetric, then

$$\sin\theta(\varphi_j, \tilde{\varphi}_i) = \|(I - P_j)\tilde{\varphi}_i\| \leq \frac{\|E\|}{d_{i,j}}. \tag{15}$$

From the proof of Lemma 1, it is seen that (9) cannot be improved, so the bounds in Theorems 3–4 are optimal. Although it is not easy to construct such a concrete example to show that these upper bounds are attainable, we are indeed able to present an example to illustrate that the upper bound of (15) are almost so, that is, there are A and E such that the ratio of the right-hand side and the left-hand side of (15)

$$\frac{\dfrac{\|E\|}{d_{i,j}}}{\sin\theta(\varphi_j, \tilde{\varphi}_i)} = 1 + O(\|E\|), \tag{16}$$

as shown below:

Construct the following symmetric matrix A and perturbation matrix E

$$A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, E = \begin{pmatrix} 0 & \epsilon \\ \epsilon & 0 \end{pmatrix},$$

where a, b are real numbers with $a > b$, ϵ is a positive number and $\epsilon \ll a - b$. Clearly, $\|E\| = \epsilon$, and for the eigenvalue a and the corresponding eigenvector e_1 of A, its perturbed eigenvalue and the corresponding unnormalized eigenvector are $\left(a + b + \sqrt{(a-b)^2 + 4\epsilon^2}\right)/2$ and $\left(1, -2\epsilon/\left(a - b + \sqrt{(a-b)^2 + 4\epsilon^2}\right)\right)^H$, respectively. It is easily verified that the left-hand side and the right-hand side of (15) are

$$\frac{2\epsilon}{\sqrt{\left(a - b + \sqrt{(a-b)^2 + 4\epsilon^2}\right)^2 + 4\epsilon^2}}, \quad \frac{2\epsilon}{a - b + \sqrt{(a-b)^2 + 4\epsilon^2}},$$

respectively. Then it is readily justified that for this example the left-hand side of (16) is bigger than one and less than $1+\epsilon/(a-b) = 1+\|E\|/(a-b) = 1+O(\|E\|)$.

Besides, we point out that the optimal lower bounds for $\sin\theta$ in Theorems 3–4 are zero, namely, there are A and a nonzero perturbation matrix E such that $\sin\theta = 0$ always holds in Theorems 3–4. For example, assume that $A = X\Lambda X^{-1}$ is the eigendecomposition of a diagonalizable matrix A, and construct a perturbation matrix $E = X\delta\Lambda X^{-1}$, where Λ is a diagonal matrix with the diagonal entries λ_i and $\delta\Lambda$ is a nonzero diagonal matrix with the diagonal entries $\delta\lambda_i$, $i = 1, 2, \cdots, N$. Obviously, $A+E = X(\Lambda+\delta\Lambda)X^{-1}$ is the eigendecomposition of $A+E$, whose eigenvalues $\tilde{\lambda}_i$ are $\lambda_i+\delta\lambda_i$, $i = 1, 2, \cdots, N$. Therefore, λ_i and $\tilde{\lambda}_i$ have the same eigenvectors no matter how $\delta\Lambda$ changes, namely, $\sin\theta = 0$ always holds in Theorem 4 no matter how $\delta\Lambda$ changes.

From now on we restore the original notation used.

Theorem 5. *Let $\sigma_{\min}(\tilde{\Phi}_{ik}^{(m)})$ and $\sigma_{\min}(\tilde{\Phi}_{ik})$ be the smallest singular values of the matrices $\tilde{\Phi}_{ik}^{(m)}$, $\tilde{\Phi}_{ik}$, respectively. Then*

$$\sigma_{\min}(\tilde{\Phi}_{ik}^{(m)}) \le \sigma_{\min}(\tilde{\Phi}_{ik}) + \sqrt{k}\max_{1\le j\le k}\|\tilde{\varphi}_{ij} - \tilde{\varphi}_{ij}^{(m)}\|. \tag{17}$$

In particular, if $k > d_i$, then

$$\sigma_{\min}(\tilde{\Phi}_{ik}^{(m)}) \le \sqrt{k}\max_{1\le j\le k}\|\tilde{\varphi}_{ij} - \tilde{\varphi}_{ij}^{(m)}\| \tag{18}$$

$$\approx \frac{\sqrt{k}\cdot C_{ik}}{g_{i,m}}\max_{1\le j\le k}\|\tilde{r}_{ij}^{(m)}\| \quad \textit{for small} \quad \|\tilde{r}_{ij}^{(m)}\| \tag{19}$$

where

$$C_{ik} = \max_{1\le j\le k}\inf_{D\ \textit{diag.}}\kappa(X_{ij}^{(m)}D)(1+\|P_i\|), \quad g_{i,m} = \min_{1\le j\le k} g_{i,m}^{(j)},$$

and $C_{ik} = 1$ when A is symmetric.

Proof. Let us decompose

$$\tilde{\Phi}_{ik}^{(m)} = \tilde{\Phi}_{ik} + \tilde{\Phi}_{ik}^{(m)} - \tilde{\Phi}_{ik}.$$

Then in terms of [30, p. 101–102], we have

$$\begin{aligned}\sigma_{\min}(\tilde{\Phi}_{ik}^{(m)}) &\le \sigma_{\min}(\tilde{\Phi}_{ik}) + \|\tilde{\Phi}_{ik}^{(m)} - \tilde{\Phi}_{ik}\| \\ &\le \sigma_{\min}(\tilde{\Phi}_{ik}) + \sqrt{k}\max_{1\le j\le k}\|\tilde{\varphi}_{ij} - \tilde{\varphi}_{ij}^{(m)}\|.\end{aligned}$$

If $k > d_i$, $\sigma_{\min}(\tilde{\Phi}_{ik}) = 0$. Note from Theorem 3 that for small $\|\tilde{r}_{ij}^{(m)}\|$

$$\|\tilde{\varphi}_{ij} - \tilde{\varphi}_{ij}^{(m)}\| = 2\sin\frac{\theta(\tilde{\varphi}_{ij}, \tilde{\varphi}_{ij}^{(m)})}{2} \approx \sin\theta(\tilde{\varphi}_{ij}, \tilde{\varphi}_{ij}^{(m)}).$$

Then we get (18) and (19). □

Based on this theorem, we can decide if $\tilde{\Phi}_{ik}^{(m)}$, $i = 1, 2, \cdots, r$, are approximately column rank deficient in the sense of (19) and thus detect the multiplicities d_i and compute approximate bases of $\mathcal{P}_i$.

4. Two Arnoldi Type Algorithms

Based on the previous analysis, we can present Arnoldi type algorithms for determining λ_i, d_i and bases of $\mathcal{P}_i$, $i = 1, 2, \cdots, r$ when A is diagonalizable.

Algorithm 2

1. Set $k = 1$, define the set $\mathcal{S} = \{1, 2, \cdots, r\}$ and give a tolerance *tol*.
2. Start: Choose a real initial vector $v_1^{(k)}$ of norm one and $m > r$, where m is the steps of Arnoldi's process.
3. Perform m steps of Arnoldi's process starting with $v_1^{(k)}$ and compute the m eigenvalues of the resulting Hessenberg matrix $H_m^{(k)}$. Then select certain $\lambda_{1k}^{(m)}, \cdots, \lambda_{rk}^{(m)}$ as approximations to the required $\lambda_1, \cdots, \lambda_r$.
4. Test convergence of r approximating eigenpairs $\lambda_{ik}^{(m)}$, $\tilde{\varphi}_{ik}^{(m)}$ using (2). If they all drop below *tol*, then go to Step 5, else go to Step 6.
5. For all $i \in \mathcal{S}$, set $\tilde{\Phi}_{ik}^{(m)} = (\tilde{\varphi}_{i1}^{(m)}, \cdots, \tilde{\varphi}_{ik}^{(m)})$. For $k > 1$, decide if $\tilde{\Phi}_{ik}^{(m)}$ for all $i \in \mathcal{S}$ are column rank deficient in the sense of (19) (how to use it sees later discussions in Section 5). If yes, set $d_i = k - 1$ and update $\mathcal{S} = \mathcal{S} - \{i\}$. If $\mathcal{S} = \emptyset$, stop; otherwise, assign $k = k + 1$ and go to Step 2.
6. Construct a new initial vector $v_1^{(k)}$ from $\tilde{\varphi}_{ik}^{(m)}$, $i = 1, 2, \cdots, r$ obtained from Step 4 by (1), then go to Step 3.

As was stated in Section 2, to enhance the efficiency of Algorithm 2, we suggest to use it in conjunction with the Chebyshev iteration, and the resulting Arnoldi-Chebyshev algorithm works for the right-most eigenpairs of A.

Algorithm 3

1. Set $k = 1$, define the set $\mathcal{S} = \{1, 2, \cdots, r\}$ and give a tolerance *tol*.
2. Start: Choose a real initial vector $v_1^{(k)}$ of norm one, the steps m of Arnoldi's process and the steps n of the Chebyshev iteration.
3. Perform m steps of Arnoldi's process starting with $v_1^{(k)}$ and compute the m eigenvalues of the resulting Hessenberg matrix $H_m^{(k)}$. Of them select $\lambda_{1k}^{(m)}, \cdots, \lambda_{rk}^{(m)}$ with largest real parts as approximations to $\lambda_1, \cdots, \lambda_r$, and set $R_{r,k}^{(m)} = \{\lambda_{r+1k}^{(m)}, \cdots, \lambda_{mk}^{(m)}\}$.
4. Test convergence of r approximating eigenpairs $\lambda_{ik}^{(m)}, \tilde{\varphi}_{ik}^{(m)}$ using (2). If they are all below *tol*, then go to Step 8, else go to Step 5.

5. From $R_{r,k}^{(m)}$, identify an ellipse containing $R_{r,k}^{(m)}$ but $\lambda_{1k}^{(m)}, \cdots, \lambda_{rk}^{(m)}$.

6. Generate an initial vector for the Chebyshev iteration from $\tilde{\varphi}_{ik}^{(m)}, i = 1, 2, \cdots, r$ obtained from Step 4 by (1).

7. Perform n steps of the Chebyshev iteration to obtain a vector z_n, take $v_1^{(k)} = z_n/\|z_n\|$, and return to Step 3.

8. For all $i \in \mathcal{S}$, set $\tilde{\Phi}_{ik}^{(m)} = (\tilde{\varphi}_{i1}^{(m)}, \cdots, \tilde{\varphi}_{ik}^{(m)})$. For $k > 1$, decide if $\tilde{\Phi}_{ik}^{(m)}$ for all $i \in \mathcal{S}$ are column rank deficient in the sense of (19). If yes, set $d_i = k - 1$ and update $\mathcal{S} = \mathcal{S} - \{i\}$. If $\mathcal{S} = \emptyset$, stop; otherwise, assign $k = k + 1$ and go to Step 2.

We refer to [9, 27] for details on Steps 5–7.

We point out that if A is symmetric then Algorithms 2–3 naturally work by reducing to the corresponding symmetric Lanczos and Lanczos-Chebyshev algorithms with considerable savings in generating an orthonormal basis of a given Krylov subspace because of the three term recurion formulas.

5. Implementations

Algorithms 2–3 are good in exact arithmetic. In finite precision, however, the computation of w in Algorithm 1 can undergo a severe cancellation, so that the resulting system $\{v_1, v_2, \cdots, v_m\}$ might be far from orthonormal. To maintain the mutually orthogonality of $\{v_l\}_1^m$, Saad[25] suggested to use the Gram-Schmidt method with iterative refinement developed in [7] as an effective remedy for loss of orthogonality. It performs reorthogonalization only when a cancellation occurs, and carries on reorthogonalization as long as cancellation persists. By adopting the strategy of reorthogonalization, we avoid the occurrence of spurious eigenvalues in Algorithms 2–3.

In implementations, Algorithm 1 is subject to breakdowns, that is, at some lth step, $l < m$, we have $h_{l+1l} = 0$. In fact, if some initial vector v_1 lies exactly in an invariant subspace of dimension l and not in any invariant subspace of smaller dimension, such a phenomenon will occur. However, it is shown [27] that in this case $\mathcal{K}_l(v_1, A)$ will be invariant, which implies, in particular, that the l Ritz pairs $\lambda_i^{(l)}, \varphi_i^{(l)}$, $i = 1, 2, \cdots, l$ are exact. Therefore, such breakdowns are lucky. In finite precision, exact breakdowns are rare, but near breakdowns are possible. At this time, the right-hand sides of (2) are very small. This suggests that we stop Arnoldi's process before the mth step once such a case occurs. If $l < r$, we continue Arnoldi's process and seek more eigenvalues, keeping those l ones that are already obtained.

For $v_1^{(k)}, k = 1, 2, \cdots$ in Step 2 of Algorithms 2–3, we choose them randomly in a uniform distribution. In such a way the assumption on B_k of (5) is satisfied in practice.

For Step 6 in Algorithms 2–3, we refer to [9, 25, 27].

Another important point is how to implement Step 5 of Algorithm 2 and Step 8 of Algorithm 3.

From Theorem 5, when $\tilde{\Phi}_{ik}^{(m)}$ is approximately column rank deficient, $\sigma_{\min}(\tilde{\Phi}_{ik}^{(m)})$ will be some small number depending on $\|\tilde{r}_{ij}^{(m)}\|$, $1 \le j \le k$. Note that C_{ik} and $g_{i,m}$ are unknown a-priori in practice. But according to Theorem 5, we may adopt the following criterion:

If restarting proceeds until the inequality

$$\sigma_{\min}(\tilde{\Phi}_{ik}^{(m)}) \le \frac{\sqrt{k} \cdot C'_{ik}}{g'_{i,m}} \max_{1 \le j \le k} \|\tilde{r}_{ij}^{(m)}\| \tag{20}$$

holds, then λ_i is $(k-1)$ multiple. Here C'_{ik} is a moderate factor, say, no bigger than 1000, which means that the eigenvalue problem of A is not too ill conditioned, and $g'_{i,m} = \min\limits_{l \ne i} \mid \lambda_{l1}^{(m)} - \lambda_{i1}^{(m)} \mid$. Note that if A is symmetric, we then take $C'_{ik} = 1$, $g'_{i,m} = \min(\lambda_{i-11}^{(m)} - \lambda_{i1}^{(m)}, \lambda_{i1}^{(m)} - \lambda_{i+11}^{(m)})$.

Finally, we point out that if $\lambda_{i+1} = \bar{\lambda}_i$ then $\tilde{\Phi}_{ik}^{(m)}$ and $\tilde{\Phi}_{i+1k}^{(m)}$ have the same rank as A is assumed to be real. In this case, it is only necessary to determine the rank of $\tilde{\Phi}_{ik}^{(m)}$.

6. Numerical Experiments

We report three numerical examples to show reliability of the algorithms. They are performed using MATLAB4.0 on an Intel Pentium 100MHZ with 40 MegaBytes primary memory and double precision $eps \approx 2.22 \times 10^{-16}$. The efficiency of the algorithms is dominated by the number of matrix-vector multiplications, indicated by $m.v$.

Example 1. Consider the Chuck matrices from [2, 6]. The matrices have several double multiple eigenvalues. The objective is to compute a few dominant eigenvalues with magnitudes greater than one and bases of corresponding eigenspaces. In the experiments to be reported below, we test the matrix CK656 which has largest order 656 among this family of matrices, and we want to compute the four dominant eigenvalues of A and determine their multiplicities. The dominant eigenvalues of A are equal to those right-most ones, so Algorithm 3 also works. We run Algorithms 2–3 for CK656, and they stop as soon as all actual residual norms drop below $tol = 10^{-8}$ and k satisfies condition (20), where we take all $C_{ik} = 1000$. Tables 1–2 show the results and processes of determining λ_i, d_i, $i = 1, 2, 3, 4$ obtained by Algorithms 2–3, respectively, where "*it*" denotes the number of iterations and $svd(X)$ the set of all singular values of a matrix X. The converged four eigenvalues are, e.g. for $m = 15$,

$$\lambda_1 \approx 5.50237837887538, \lambda_2 \approx 1.59397169682838,$$
$$\lambda_3 \approx 1.41904261708523, \lambda_4 \approx 1.41195129689298.$$

It can be seen from Tables 1–2 that the algorithms solve the problem efficiently and reliably. For this problem, the Chebyshev technique does not gain much since Algorithm 2 itself is really fast already.

Example 2. We construct a 1000×1000 matrix $A = X\Lambda X^{-1}$, where

$$\Lambda = \text{diag}\left(\begin{bmatrix} 1.9 & 0.5 \\ -2 & 1.9 \end{bmatrix}, \begin{bmatrix} 1.9 & 0.5 \\ -2 & 1.9 \end{bmatrix}, \begin{bmatrix} 1.9 & 0.5 \\ -2 & 1.9 \end{bmatrix}, 1.8, 1.6, 1.4, 1-(j-1)/1000 \right)$$
$$j = 10, 11, \cdots, 1000,$$

and X is generated randomly in a uniform distribution, $\kappa(X) \approx 148607$. Therefore, the eigenvalue problem of A is quite ill conditioned. The matrix A has two three multiple eigenvalues $\lambda_1 = 1.9 + i, \lambda_2 = 1.9 - i$ and the rest eigenvalues are simple.

Table 1. The process of determining the multiplicities of the four dominant eigenvalues of Example 1 by Algorithm 2.

m	it	$m.v$	Residual norms			
			1	2	3	4
15	8	120	$1.7D-15$	$2.8D-10$	$1.1D-9$	$2.4D-9$
20	3	60	$1.4D-9$	$1.5D-11$	$4.7D-9$	$8.2D-11$
25	3	75	$3.7D-12$	$9.7D-11$	$9.3D-10$	$4.7D-9$

m	k	$svd(\tilde{\Phi}_{1k}^{(m)})$	$svd(\tilde{\Phi}_{2k}^{(m)})$	$svd(\tilde{\Phi}_{3k}^{(m)})$	$svd(\tilde{\Phi}_{4k}^{(m)})$
15	1	1	1	1	1
20	2	1.41316204	1.34519400	1.41325908	1.40831909
		0.05452577	0.43640932	0.05194975	0.12898585
25	3	1.70722567	1.48993911	1.72952700	1.65897486
		0.29219943	0.88322220	0.09346846	0.49779755
		$1.8D-10$	$1.7D-10$	$6.5D-9$	$7.2D-8$

	λ_1	λ_2	λ_3	λ_4
Multiplicity	2	2	2	2

Table 2. The process of determining the multiplicities of the four dominant eigenvalues of Example 1 by Algorithm 3

m	n	it	$m.v$	Residual norms			
				1	2	3	4
15	10	3	65	$3.1D-15$	$1.3D-11$	$8.3D-11$	$1.5D-10$
20	10	2	50	eps	$4.3D-10$	$4.2D-9$	$5.4D-9$
25	10	2	50	eps	$2.7D-15$	$1.1D-13$	$1.5D-13$

m	k	$svd(\tilde{\Phi}_{1k}^{(m)})$	$svd(\tilde{\Phi}_{2k}^{(m)})$	$svd(\tilde{\Phi}_{3k}^{(m)})$	$svd(\tilde{\Phi}_{4k}^{(m)})$
15	1	1	1	1	1
20	2	1.40437873	1.06366280	1.00067142	1.01730241
		0.16649437	0.93199863	0.99932814	0.98239290
25	3	1.71678759	1.42597502	1.41421484	1.41507268
		0.22943490	0.98315575	0.99999820	0.99878391
		$8.6D-16$	$1.6D-8$	$1.9D-8$	$5.8D-9$

	λ_1	λ_2	λ_3	λ_4
Multiplicity	2	2	2	2

Both algorithms are run for A. We want to find the five right-most eigenvalues and determine their multiplicities. The stopping criterion is as in Example 1. Tables 3–4

show the results and processes of determining the required λ_i, d_i obtained by Algorithms 2–3, respectively. The computed eigenvalues are, e.g. for $m = 20$,

Table 3. The process of determining the multiplicities of the five right-most eigenvalues of Example 2 by Algorithm 2.

m	it	$m.v$	Residual norms				
			1	2	3	4	5
15	3	45	$1.7D-13$	$1.7D-13$	$9.5D-11$	$4.D-10$	$2.6D-10$
20	2	40	$1.4D-12$	$1.4D-12$	$9.5D-11$	$3.5D-11$	$1.2D-11$
30	2	60	$2.4D-13$	$2.4D-13$	$5.3D-11$	$3.7D-11$	$8.9D-11$
40	1	40	$1.5D-13$	$1.5D-13$	$3.6D-11$	$7.6D-13$	$9.1D-13$

m	k	$svd(\tilde{\Phi}_{1k}^{(m)})$	$svd(\tilde{\Phi}_{2k}^{(m)})$	$svd(\tilde{\Phi}_{3k}^{(m)})$	$svd(\tilde{\Phi}_{4k}^{(m)})$	$svd(\tilde{\Phi}_{5k}^{(m)})$
15	1	1	1	1	1	1
20	2	1.29495985	1.29495985	1.41421356	1.41421356	1.41421356
		0.56840036	0.56840036	$6.8D-11$	$3.3D-10$	$3.D-10$
30	3	1.50060992	1.50060992			
		0.76550906	0.76550906			
		0.40269809	0.40269809			
40	4	1.52352794	1.52352794			
		1.20323886	1.20323886			
		0.48070664	0.48070664			
		$2.6D-10$	$2.6D-10$			

	λ_1	λ_2	λ_3	λ_4	λ_5
Multiplicity	3	3	1	1	1

Table 4 The process of determining the multiplicities of the five right-most eigenvalues of Example 2 by Algorithm 3

m	n	it	$m.v$	Residual norms				
				1	2	3	4	5
15	20	2	50	$2.D-14$	$2.D-14$	$2.2D-9$	$2.2D-9$	$7.1D-10$
20	15	2	55	$1.D-11$	$1.D-11$	$2.8D-10$	$5.3D-10$	$2.3D-10$
10	20	6	110	$6.1D-11$	$6.1D-11$	$2.3D-9$	$4.5D-9$	$1.9D-9$
25	20	2	70	$6.5D-13$	$6.5D-13$	$5.4D-13$	$3.7D-12$	$7.6D-12$

m	k	$svd(\tilde{\Phi}_{1k}^{(m)})$	$svd(\tilde{\Phi}_{2k}^{(m)})$	$svd(\tilde{\Phi}_{3k}^{(m)})$	$svd(\tilde{\Phi}_{4k}^{(m)})$	$svd(\tilde{\Phi}_{5k}^{(m)})$
15	1	1	1	1	1	1
20	2	1.30118434	1.30118434	1.41421356	1.41421356	1.41421356
		0.55400299	0.55400299	$2.2D-10$	$1.4D-9$	$1.D-9$
10	3	1.51934966	1.51934966			
		0.77946537	0.77946537			
		0.28984632	0.28984632			
25	4	1.71536589	1.71536589			
		0.97752801	0.97752801			
		0.31930995	0.31930995			
		$1.6D-11$	$1.6D-11$			

	λ_1	λ_2	λ_3	λ_4	λ_5
Multiplicity	3	3	1	1	1

$$\lambda_{1,2} \approx 1.90000000000006 \pm 0.99999999999969i, \lambda_3 \approx 1.79999999999678,$$
$$\lambda_4 \approx 1.59999999999615, \lambda_5 \approx 1.40000000000096.$$

We see from Tables 3–4 that Algorithms 2–3 have found λ_i, $d_i, i = 1, 2, 3, 4, 5$ efficiently and reliably. Again, for this example, the Chebyshev technique gains little.

Example 3. We construct a 1000×1000 matrix $A = X\Lambda X^{-1}$, where

$$\Lambda = \text{diag}(1.66, 1.66, 1.62, 1.62, 1.3, 1, -j),$$
$$j = 7, 8, \cdots, 1000,$$

X is generated randomly in a uniform distribution and $\kappa(X) \approx 162636$. Therefore, the eigenvalue problem of A is quite ill conditioned. The matrix A has two double multiple eigenvalues $\lambda_1 = 1.66, \lambda_2 = 1.62$ which are quite clustered, and the rest eigenvalues are simple.

Algorithms 2–3 are run for this matrix. We want to find the four right-most eigenvalues and detect their multiplicities. The stopping criterion is as Examples 1–2. Tables 5–6 list the results obtained. The computed eigenvalues are, e.g. for $m = 70$ and $n = 60$,

$$\lambda_1 \approx 1.66000000028226, \lambda_2 \approx 1.61999999883579,$$
$$\lambda_3 \approx 1.30000000011923, \lambda_4 \approx 1.00000000003169.$$

Table 5. The process of determining the multiplicities of the four right-most eigenvalues of Example 3 by Algorithm 2

m	it	$m.v$	Residual norms			
			1	2	3	4
100	14	1400	$4.8D-10$	$1.6D-9$	$8.5D-10$	$3.7D-10$
110	10	1100	$1.1D-10$	$4.9D-10$	$1.4D-9$	$2.5D-10$
120	6	720	$3.6D-11$	$9.3D-11$	$5.8D-12$	$2.9D-10$

m	k	$svd(\tilde{\Phi}_{1k}^{(m)})$	$svd(\tilde{\Phi}_{2k}^{(m)})$	$svd(\tilde{\Phi}_{3k}^{(m)})$	$svd(\tilde{\Phi}_{4k}^{(m)})$
100	1	1	1	1	1
110	2	1.12394010	1.41395284	1.41421356	1.41421356
		0.85834646	0.02715461	$2.2D-10$	$8.7D-11$
120	3	1.31793692	1.73097850		
		1.12385154	0.06093795		
		$6.5D-10$	$1.6D-9$		

	λ_1	λ_2	λ_3	λ_4
Multiplicity	2	2	1	1

As in Examples 1–2, the algorithms solved this multiple eigenvalue problem reliably. However, Algorithm 2 converged quite slowly and it used many matrix-vector multiplications to achieve the desired accuracy. This is not surprising because the required λ_1 and λ_2 are not well separated and the eigenvalue problem of A is quite ill conditioned, while Arnoldi's method is less efficient in this case; see [10, 11, 14, 25, 26,

27] for a theoretical analysis. In contrast, Algorithm 3 was much better than Algorithm 2. This shows that the Chebyshev iteration may have a strong effect on the efficiency of Arnoldi's method.

A number of other experiments have been run, and they have shown reliability of the algorithms provided Arnoldi's method and the Arnoldi-Chebyshev method work well.

Table 6. The process of determining the multiplicities of the four right-most eigenvalues of Example 3 by Algorithm 3

m	n	it	$m.v$	Residual norms			
				1	2	3	4
60	40	12	1160	$2.9D-9$	$3.5D-9$	$9.4D-10$	$3.D-10$
70	60	4	460	$6.6D-9$	$8.1D-9$	$2.2D-9$	$7.5D-10$
80	60	4	500	$2.2D-12$	$2.7D-12$	$7.6D-13$	$2.6D-13$

m	k	$svd(\tilde{\Phi}_{1k}^{(m)})$	$svd(\tilde{\Phi}_{2k}^{(m)})$	$svd(\tilde{\Phi}_{3k}^{(m)})$	$svd(\tilde{\Phi}_{4k}^{(m)})$
60	1	1	1	1	1
70	2	1.29817559	1.41359726	1.41421356	1.41421356
		0.56101705	0.04174660	$8.1D-10$	$2.1D-10$
80	3	1.48380523	1.73089008		
		0.89348869	0.06339966		
		$5.D-9$	$1.3D-9$		

	λ_1	λ_2	λ_3	λ_4
Multiplicity	2	2	1	1

7. Concluding Remarks

We have proposed Arnoldi type algorithms for solving large unsymmetric multiple eigenvalue problems when the matrix is diagonalizable, supported by the theoretical background. Some of the results are new for Arnoldi's method and the perturbation analysis of a general eigenvalue problem. Numerical experiments have shown reliability of the proposed algorithms.

The idea used can be easily generalized to some other methods, e.g. the biorthogonalization Lanczos method[20], other variants of Arnoldi's method, e.g. [13, 24, 28] and the power method[30]. At the same time, we point out that this idea is essential for the block Arnoldi methods when block size is smaller than or equal to the multiplicities of the required eigenvalues[10,14], and it should work in other block methods, e.g.[15].

Acknowledgements A preliminary version of this paper was [18] that is part of [10]. I thank my superviser Professor L. Elsner for his constant guidance. I am grateful to Professor F. Chatelin for useful comments on it. Meanwhile, thanks go to Dr. S. Godet-Thobie for his making the Fortran code of finding the optimal ellipse developed by Dr. D. Ho available.

References

[1] W.E. Arnoldi, The principle of minimized iteration in the solution of the matrix eigenvalue

problem, *Quart. Appl. Math.,* **9** (1951), 11–29.
[2] Z. Bai, R. Barrett, D. Day, J. Demmel, J. Dongarra, Test matrix collection (non-Hermitian eigenvalue problems), *Tech. Rep., Math. Dept.,* University of Kentucky, 1996.
[3] F. Chatelin, S. Godet-Thobie, Stability Analysis in Aeronautical Industries, Proceedings of the 2nd Symposium on High Performance Computing, Eds. M. Durand and F.El. Dabaghi, Elsevier/North-Holland, (1991), 415–422.
[4] F. Chatelin, Eigenvalues of Matrices, John Wiley and Sons, 1993.
[5] J. Cullum, R.A. Willoughby, Lanczos Algorithms for Large Symmetric Eigenvalue Computations, Vol.1, Theory, Vol.2, Programs, Birkhäuser, Basel, 1985.
[6] J. Cullum, R.A. Willoughby, A Practical Procedure for Computing Eigenvalues of Large Sparse Nonsymmetric Matrices, in *Large Scale Eigenvalue Problem*, Eds. J. Cullum and R.A. Willoughby, Elsevier/North-Holland, 1986, 199–240.
[7] G.W. Daniel, G.B. Gragg, L. Kaufmann, G.W. Stewart, Reorthogonalization and stable algorithms for updating the Gram-Schmidt QR factorization, *Math. Comput.,* **30** (1976), 772–795.
[8] R.G. Grimes, J.G. Lewis, H.D. Simon, A shifted block Lanczos algorithm for solving sparse symmetric generalized eigenproblems, *SIAM J. Matrix Anal. Appl.,* **15** (1994), 228–272.
[9] D. Ho, F. Chatelin, M. Bennani, Arnoldi-Tchebychev procedure for large nonsymmetric matrices, *Math. Mod. Numer. Anal.,* **24** (1990), 53–65.
[10] Z. Jia, Some Numerical Methods for Large Unsymmetric Eigenproblems, Ph.D. thesis, Math. Dept., University of Bielefeld, Germany, 1994.
[11] Z. Jia, The convergence of generalized Lanczos methods for large unsymmetric eigenproblems, *SIAM J. Matrix Anal. Appl.,* **16** (1995), 843–862.
[12] Z. Jia, A block incomplete orthogonalization method for large unsymmetric eigenproblem, *BIT,* **35** (1995), 516–539.
[13] Z. Jia, Refined iterative algorithms based on Arnoldi's process for large unsymmetric eigenproblems, *Linear Algebra Appl.,* **259** (1997), 1–23.
[14] Z. Jia, Generalized block Lanczos methods for large unsymmetric eigenproblems, *Numer. Math.,* **80** (1998), 239–266.
[15] Z. Jia, A refined iterative algorithm based on the block Arnoldi process for large unsymmetric eigenproblems, *Linear Algebra Appl.,* **270** (1998), 171–189.
[16] Z. Jia, Polynomial characterizations of the approximate eigenvectors by the refined Arnoldi method and an implicitly restarted refined Arnoldi algorithm, *Linear Algebra Appl.,* **287** (1999), 191–214.
[17] Z. Jia, Composite orthogonal projection methods for large matrix eigenproblems, *Science in China* (*Series A*), **42** (1999), 577-585.
[18] Z. Jia, Arnoldi type algorithms for large unsymmetric multiple eigenvalue problems, Tech. Rep. 95–026, Math. Dept., University of Bielefeld, Germany, 1995.
[19] Z. Jia, L. Elsner, Improving eigenvectors in Arnoldi's method, *J. Comput. Math.,* to appear.
[20] C. Lanczos, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, *J. Res. Nat. Bur. Stand.,* **45** (1950), 225–280.
[21] B.N. Parlett, D.C. Scott, The Lanczos algorithm with selective orthogonalization, *Math. Comput.,* **33** (1979), 217–238.
[22] B.N. Parlett, The Symmetric Eigenvalue Problem, Prentice-Hall, Englewood Cliffs, N.J., 1980.
[23] A. Ruhe, The Two Sided Arnoldi Algorithm for Nonsymmetric Eigenvalue Problems, Matrix Pencils, Lecture Notes in Mathematics 973, Eds. B.K. Kågström and A. Ruhe, Springer–Verlag, Berlin, 1983, 104–120.

[24] A. Ruhe, Rational Krylov algorithms for nonsymmetric eigenvalue problems. II. Matrix pairs, *Linear Algebra Appl.,* **197/198** (1994), 283–295.
[25] Y. Saad, Variations on Arnoldi's method for computing eigenelements of large nonsymmetric matrices, *Linear Algebra Appl.,* **34** (1980), 269–295.
[26] Y. Saad, Projection Methods for Solving Large Eigenvalues Problems, Matrix Pencils, Lecture Notes in Mathematics 973, Eds. B.K. Kågström and A. Ruhe, Springer–Verlag, Berlin, 1983, 121–144.
[27] Y. Saad, Chebyshev acceleration techniques for solving nonsymmetric eigenvalue problems, *Math. Comput.,* **42** (1984), 567–588.
[28] D.C. Sorensen, Implicit application of polynomial filters in a k-step Arnoldi method, *SIAM J. Matrix Anal. Appl.,* **13** (1992), 357–385.
[29] G.W. Stewart, J.-G. Sun, Matrix Perturbation Theory, Academic Press, Boston, 1990.
[30] J.H. Wilkinson, The Algebraic Eigenvalue Problem, Clarendon Press, Oxford, 1965.