Arnoldi Type Algorithms for Large Unsymmetric Multiple Eigenvalue Problems*

Zhongxiao Jia†

Januray 27, 1995

Abstract

Two Arnoldi type algorithms are proposed for large unsymmetric *multiple* eigenvalue problems when the matrix A involved is *diagonalizable*. The theoretical background is established. The algorithms can adaptively determine the multiplicity of a eigenvalue and a basis of the associated *eigenspace*. Numerical experiments show reliability of the algorithms.

Keywords. multiple eigenvalue, diagonalizable, Arnoldi's algorithm, eigenprojector, convergence, error bounds

AMS(MOS) subject classifications. 65F15, 15A18

1 Introduction

The Lanczos algorithm is a very powerful tool for extracting a few extreme eigenvalues and associated eigenvectors of large symmetric matrices [9, 11]. 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, 14]. It can be used to compute outer part of the spectrum and corresponding eigenvectors [6, 7, 8, 14, 15, 17]. In order to improve overall performance, Saad [16] suggests to use it in conjunction with the Chebyshev iteration. There are other variants available; see, e.g. [13, 17].

*This work was supported by the Graduiertenkolleg Mathematik at the University of Bielefeld. Presented at the Workshop on Reliability of Computations, CERFACS, March 29–April 1, 1993 and the Lanczos International Centenary Conference, Raleigh, Dec. 12–17, 1993.

†Fakultät für Mathematik, Universität Bielefeld, email: `jia@mathematik.uni-bielefeld.de`, Postfach 10 01 31, 33501 Bielefeld, Germany

To apply Arnoldi's algorithm and its variants to practical problems, one must account for some difficulties which can arise in applications [2, 3], e.g.

Difficulty* *Multiple eigenvalues are a common occurrence.*

In the symmetric case, Parlett *et al.* [10] use the Lanczos algorithm with selective orthogonalization to solve Difficulty*. It maintains the semi-orthogonality among the Lanczos vectors so as to avoid the occurrence of spurious eigenvalues and determines the multiplicities of the required eigenvalues and the associated eigenspaces by restarting. Its 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 [10] cannot solve Difficulty* here, 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 [10]. Theoretically speaking, a simple simulation of the idea used in [10] 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 [12] 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 [7, 8]. They can be used to compute outer part of the spectrum and corresponding eigenvectors, up to a multiplicity equal to the block size when A is diagonalizable. However, if the multiplicities of the required eigenvalues are bigger than the block size, the block algorithms themselves are not able to determine the multiplicity of a multiple eigenvalue and the associated eigenspace.

In this paper, we attempt to design Arnoldi type algorithms for solving such a Difficulty* when A is *diagonalizable*. As is seen from [7, 8], although the idea proposed seems not complicated, it is important not only in its own right but also indispensable for the block Arnoldi methods when the block size is less 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 proposed in Section 4. Some of the results, i.e. theoretical error bounds for eigenvectors, are also new for both Arnoldi's method and the perturbation analysis of a general eigenvalue problem; in Section 4 we present two Arnoldi type algorithms to solve Difficulty* for the diagonalizable A, the second of which is the first with the Chebyshev iteration [5, 16]; in Section 5, we discuss implementations of the algorithms; in Section 6, we report on several 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 labelled in the decreasing order of their real parts

$$Re(\lambda_1) \geq Re(\lambda_2) \geq \cdots \geq Re(\lambda_M), \tag{1}$$

where the multiplicities of λ_i are $d_i, i = 1, 2, \ldots, M$. Under the above assumption let $\mathcal{P}_i$ be the d_i-dimensional eigenspace associated with λ_i and the columns of $\Phi_{id_i} = (\varphi_{i1}, \varphi_{i2}, \ldots, \varphi_{id_i})$ form a basis of $\mathcal{P}_i$, where $\|\varphi_{ij}\| = 1$ and $\|\cdot\|$ denotes the 2-norm. Let $\Psi_{id_i} = (\psi_{i1}, \psi_{i2}, \ldots, \psi_{id_i})$, where ψ_{ij} are the left eigenvectors associated with λ_i such that $\psi_{ik}^H \varphi_{ij} = \delta_{kj}$, $k, j = 1, 2, \ldots, d_i$. Here "H" denotes the conjugate transpose and δ_{kj} the Kronecker delta. We want to compute a few eigenvalues λ_i with largest real parts and determine the corresponding multiplicities d_i and bases of $\mathcal{P}_i$, $i = 1, 2, \ldots, r$.

We denote by $\mathcal{K}_m(v, A)$ the Krylov subspace spanned by $v, Av, \ldots, A^{m-1}v$, 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, $i = 1, 2, \ldots, M$.

2.2 The underlying Arnoldi algorithm

The 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 $l = 1, 2, \ldots, m$ do
   2.1. $w = Av_l$.
   2.2. For $j = 1, 2, \ldots, l$ do
   $h_{jl} = v_j^H A v_l$,
   $w = w - h_{jl} v_j$.
   2.3. $h_{l+1l} = \|w\|$.
   2.5. $v_{l+1} = w / h_{l+1l}$.

For practical considerations, refer to [14].

This algorithm produces an orthonormal basis $\{v_l\}_1^m$ of $\mathcal{K}_m(v_1, A)$. Define the matrix $V_m = (v_1, v_2, \ldots, v_m)$. 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 $H_m = V_m^H A V_m$, whose entries are the h_{jl} generated 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. The corresponding approximate eigenvectors, called the Ritz vectors of A in $\mathcal{K}_m(v_1, A)$, are

$$\varphi_i^{(m)} = V_m y_i^{(m)}, \tag{2}$$

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 a-posteriori bound

$$\|r_i^{(m)}\| = \|(A - \lambda_i^{(m)} I)\varphi_i^{(m)}\| = h_{m+1m} \mid e_m^H y_i^{(m)} \mid, \tag{3}$$

in which $e_m = (0, 0, \ldots, 0, 1)^H$. (3) can be used for a stopping criterion which checks cheaply the size of the residual without computing $\varphi_i^{(m)}$ by (2).

Since Algorithm 1 has to save all the vectors v_l generated previously and its amount of computation increases quadratically with steps, Arnoldi's algorithm usually has to be used iteratively in practice; see, e.g. [14, 17]. In order to improve efficiency, an iterative Arnoldi algorithm can be accelerated by the Chebyshev iteration [5, 16]. The Chebyshev iteration is for the purpose of amplifying 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 $v_1^{(1)}, v_1^{(2)}, \ldots, v_1^{(k)}$. Then each $v_1^{(j)}, 1 \le j \le k$, can be expanded in the eigenspaces $\mathcal{P}_i, i = 1, 2, \ldots, M$ as

$$\begin{aligned} v_1^{(j)} &= b_{j1}\varphi_{i1} + b_{j2}\varphi_{i2} + \cdots + b_{jd_i}\varphi_{id_i} + u_i^{(j)}, & (4) \\ u_i^{(j)} &\in \mathcal{P}_1 \oplus \cdots \oplus \mathcal{P}_{i-1} \oplus \mathcal{P}_{i+1} \oplus \cdots \oplus \mathcal{P}_M, \ 1 \le j \le k, & (5) \end{aligned}$$

where "$\oplus$" denotes the direct sum.

Let

$$B_k = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1d_i} \\ b_{21} & b_{22} & \cdots & b_{2d_i} \\ \vdots & \vdots & \cdots & \vdots \\ b_{k1} & b_{k2} & \cdots & b_{kd_i} \end{pmatrix}. \tag{6}$$

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

$$v_1^{(j)} = \beta_j \tilde{\varphi}_{ij} + u_i^{(j)}, \ u_i^{(j)} \in \mathcal{P}_1 \oplus \cdots \oplus \mathcal{P}_{i-1} \oplus \mathcal{P}_{i+1} \oplus \cdots \oplus \mathcal{P}_M, \tag{7}$$

where $\tilde{\varphi}_{ij}, j = 1, 2, \ldots, k$ are also normalized 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 $k > d_i$, $\tilde{\varphi}_{ij}$, $j = d_i + 1, \ldots, k$ *belong* to the span of $\{\tilde{\varphi}_{ij}\}_{j=1}^{d_i}$.

Define $\tilde{\Phi}_{ik} = (\tilde{\varphi}_{i1}, \tilde{\varphi}_{k2}, \ldots, \tilde{\varphi}_{ik})$ and $\tilde{\Psi}_{ik} = (\tilde{\psi}_{i1}, \tilde{\psi}_{i2}, \ldots, \tilde{\psi}_{ik})$, 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 [14], we have the following a-priori residual bound.

Theorem 1 *Let $\gamma_m(v_1^{(j)}) = \|\pi_m(v_1^{(j)})A(I - \pi_m(v_1^{(j)}))\|$. Then for the eigenpairs λ_i, $\tilde{\varphi}_{ij}$, $1 \le i \le M$, $1 \le j \le k$, we have*

$$\|(A_m(v_1^{(j)}) - \lambda_i I)\tilde{\varphi}_{ij}\| \le \sqrt{\gamma_m^2(v_1^{(j)}) + |\lambda_i|^2}\|(I - \pi_m(v_1^{(j)}))\tilde{\varphi}_{ij}\|, \tag{8}$$

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

Therefore, whether some of the eigenpairs $\lambda_{ij}^{(m)}, \tilde{\varphi}_{ij}^{(m)}$, $i = 1, 2, \ldots, m, j = 1, 2, \ldots, k$ 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 has been shown [7, 8, 14, 15] that the right hand sides of (8) approach zero as m increases, usually first for outer part of the spectrum. Furthermore, it has been shown [6, 8] that for $v_1^{(j)}, 1 \le j \le k$, the Ritz pairs $\lambda_{ij}^{(m)}, \tilde{\varphi}_{ij}^{(m)}$ converge to $\lambda_i, \tilde{\varphi}_{ij}, 1 \le j \le k$ provided that the eigenproblem of $A_m(v_1^{(j)})$ is not too ill-conditioned and the right hand sides of (8) tend to zero. According to the previous statement, we can find an approximate basis of $\mathcal{P}_i$ in this way provided that d_i initial vectors $v_1^{(j)}$ are chosen such that the assumption on B_k of (6) is satisfied. For $k > d_i$, $\tilde{\varphi}_{ij}^{(m)}, j = d_i + 1, \ldots, k$ 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 if the matrix $\tilde{\Phi}_{i2}^{(m)} = (\tilde{\varphi}_{i1}^{(m)}, \tilde{\varphi}_{i2}^{(m)})$ is column full rank. If yes, this shows that $d_i \ge 2$, and we then take the columns of $\tilde{\Phi}_{i2}^{(m)}$ as elements of an approximate basis of $\mathcal{P}_i$; otherwise, $d_i = 1$ and $\tilde{\varphi}_{i1}^{(m)}$ is an approximate basis of $\mathcal{P}_i$. We then stop, else continue restarting with a new initial vector $v_1^{(3)}$. Proceeding in this way until at some $(k+1)$th restarting, we have that the matrix $\tilde{\Phi}_{ik}^{(m)} = (\tilde{\varphi}_{i1}^{(m)}, \tilde{\varphi}_{i2}^{(m)}, \ldots, \tilde{\varphi}_{ik}^{(m)})$ is column full rank, but $\tilde{\Phi}_{ik+1}^{(m)} = (\tilde{\varphi}_{i1}^{(m)}, \tilde{\varphi}_{i2}^{(m)}, \ldots, \tilde{\varphi}_{ik+1}^{(m)})$ 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 $\lambda_{ij}^{(m)}, \tilde{\varphi}_{ij}^{(m)}, i = 1, 2, \ldots, r, j = 1, 2, \ldots, k$, 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 $\lambda_{ij}^{(m)}, \tilde{\varphi}_{ij}^{(m)}, i = 1, 2, \ldots, r, j = 1, 2, \ldots, k$ 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 [19, p. 69].

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

$$| \lambda_i - \lambda_{ij}^{(m)} | \leq \|\tilde{\psi}_{ij}\| \cdot \|\tilde{r}_{ij}^{(m)}\| + O(\|\tilde{r}_{ij}^{(m)}\|^2). \tag{9}$$

From Theorem 2, we see that if λ_i is not very ill-conditioned, then $\lambda_{ij}^{(m)}$ is a good approximation to λ_i provided that $\|\tilde{r}_{ij}^{(m)}\|$ is small.

Before deriving error bounds for approximate eigenvectors $\tilde{\varphi}_{ij}^{(m)}$, we need the following lemma [6, 8].

Lemma 1 *Let* $x_1, x_2, \ldots, x_s$ *be* s *vectors and* $\alpha_1, \alpha_2, \ldots, \alpha_s$ s *scalars, and define the matrix* $X = (x_1, x_2, \ldots, x_s)$, *where* s *is any integer number. Then*

$$\| \sum_{j=1}^{s} \alpha_j x_j \| \geq \frac{1}{\inf_{D\ diag.} \kappa(XD)} \min_{1 \leq j \leq s} | \alpha_j | \| \sum_{j=1}^{s} x_j \|, \tag{10}$$

where D*'s are* $s \times s$ *nonsingular diagonal matrices and* $\kappa(XD)$ *denotes the condition number 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* $g_{i,m}^{(j)} = \min_{l \neq i} | \lambda_l - \lambda_{ij}^{(m)} |$. *Define the matrix*

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

Then

$$\sin \theta(\tilde{\varphi}_{ij}, \tilde{\varphi}_{ij}^{(m)}) \leq \|(I - P_i) \tilde{\varphi}_{ij}^{(m)}\| \leq \frac{\inf_{D\ diag.} \kappa(X_{ij}^{(m)} D)(1 + \|P_i\|)}{g_{i,m}^{(j)}} \|\tilde{r}_{ij}^{(m)}\|. \tag{11}$$

If A *is symmetric, then*

$$\sin \theta(\tilde{\varphi}_{ij}, \tilde{\varphi}_{ij}^{(m)}) = \|(I - P_i) \tilde{\varphi}_{ij}^{(m)}\| \leq \frac{\|\tilde{r}_{ij}^{(m)}\|}{g_{i,m}^{(j)}}. \tag{12}$$

Proof. It is well known that

$$\begin{aligned} P_l P_i &= \delta_{il} P_l, & (13) \\ \sum_{l=1}^{M} P_l &= I. & (14) \end{aligned}$$

Hence we have

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

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

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

On the other hand,

$$\begin{aligned}\|(I-P_i)(A-\lambda_{ij}^{(m)}I)\tilde{\varphi}_{ij}^{(m)}\| &\leq \|I-P_i\|\|(A-\lambda_{ij}^{(m)}I)\tilde{\varphi}_{ij}^{(m)}\| \\ &\leq (1+\|P_i\|)\|\tilde{r}_{ij}^{(m)}\|.\end{aligned}$$

By (14), we obtain

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

Therefore, combining the above three relations gives

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

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

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

Thus, (11) holds.

If A is symmetric, its eigenvectors are mutually orthogonal. In this case, $\|I-P_i\|=1$, $\inf_{D \text{ diag.}}\kappa(X_{ij}^{(m)}D)=1$ and $\sin\theta(\tilde{\varphi}_{ij},\tilde{\varphi}_{ij}^{(m)}) = \|(I-P_i)\tilde{\varphi}_{ij}^{(m)}\|$. Thus, (12) holds. $\square$

Thus, the sensitivity of $\tilde{\varphi}_{ij}$ depends on the sensitivity of the eigenvalue problem of A, the gap $g_{i,m}^{(j)}$ of $\lambda_{ij}^{(m)}$ and other exact eigenvalues of A as well as $\|\tilde{r}_{ij}^{(m)}\|$. If $\|\tilde{r}_{ij}^{(m)}\|$ is small, $\tilde{\varphi}_{ij}^{(m)}$ is a good approximation to $\tilde{\varphi}_{ij}$ provided that the eigenvalue problem of A is not too ill-conditioned and $g_{i,m}^{(j)}$ is not too small. However, if A is symmetric, the sensitivity of $\tilde{\varphi}_{ij}$ depends mainly on $g_{i,m}^{(j)}$ and $\|\tilde{r}_{ij}^{(m)}\|$.

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

Assume that A is diagonalizable and has eigenpairs $\lambda_i, \varphi_i, i = 1, 2, \ldots, N$, where $\|\varphi_i\| = 1$ and $\lambda_i \neq \lambda_j$ for $i \neq j$. Let $\psi_i, i = 1, 2, \ldots, N$ be the left eigenvectors associated with λ_i such that $\psi_i^H \varphi_j = \delta_{ij}$. Assume that $A + E$ is also diagonalizable and has eigenpairs $\tilde{\lambda}_i, \tilde{\varphi}_i, i = 1, 2, \ldots, N$, where E is a small perturbation matrix, $\tilde{\lambda}_i$ are ordered to approximate λ_i in one to one correspondence. Then how do $\tilde{\varphi}_i$ approximate φ_i, $i = 1, 2, \ldots, N$?

From the proof of Theorem 3, it is easy to answer this question.

Theorem 4 *Under the above assumptions and notation, let* $g_i = \min_{j \neq i} \mid \tilde{\lambda}_i - \lambda_j \mid$, $i = 1, 2, \ldots, N$, *and define the matrix*

$$X_i = (P_1 \tilde{\varphi}_i, \ldots, P_{i-1} \tilde{\varphi}_i, P_{i+1} \tilde{\varphi}_i, \ldots, P_N \tilde{\varphi}_i).$$

Then for $i = 1, 2, \ldots, N$, *the following relations hold:*

$$\sin \theta(\varphi_i, \tilde{\varphi}_i) \leq \|(I - P_i) \tilde{\varphi}_i\| \leq \frac{\inf_{D \ diag.} \kappa(X_i D)(1 + \|P_i\|)}{g_i} \|E\|. \tag{15}$$

This theorem establishes some new error bounds for eigenvectors for the general eigenvalue problem posed above. However, it seems difficult to compare this theorem with those existing results, e.g. [18].

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*

$$\sigma_{min}(\tilde{\Phi}_{ik}^{(m)}) \leq \sigma_{min}(\tilde{\Phi}_{ik}) + \sqrt{k} \max_{1 \leq j \leq k} \|\tilde{\varphi}_{ij} - \tilde{\varphi}_{ij}^{(m)}\|. \tag{16}$$

In particular, if $k > d_i$, *then*

$$\begin{aligned} \sigma_{min}(\tilde{\Phi}_{ik}^{(m)}) &\leq \sqrt{k} \max_{1 \leq j \leq k} \|\tilde{\varphi}_{ij} - \tilde{\varphi}_{ij}^{(m)}\| & (17) \\ &\approx \frac{\sqrt{k} \cdot C_{ik}}{g_{i,m}} \max_{1 \leq j \leq k} \|\tilde{r}_{ij}^{(m)}\| \text{ for small } \|\tilde{r}_{ij}^{(m)}\| & (18) \end{aligned}$$

where

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

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 [19, p. 101–102], we have

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

If $k > d_i$, then $\sigma_{min}(\tilde{\Phi}_{ik}) = 0$. 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 (17) and (18). $\square$

Based on this theorem, we can decide if $\tilde{\Phi}_{ik}^{(m)}, i = 1, 2, \ldots, r$, are approximately column rank deficient in the sense of (18) and thus determine the multiplicities d_i and approximate bases of $\mathcal{P}_i$.

4 Two Arnoldi type algorithms

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

Algorithm 2

1. Set $k = 1$, define the set $\mathcal{S} = \{1, 2, \ldots, r\}$ and give a tolerance tol.

2. Start: Choose a real initial vector $v_1^{(k)}$ of norm one and the number $m > r$, 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. Then select the r eigenvalues $\lambda_{1k}^{(m)}, \ldots, \lambda_{rk}^{(m)}$ with largest real parts as approximations to $\lambda_1, \ldots, \lambda_r$.

4. Test convergence of r approximating eigenpairs $\lambda_{ik}^{(m)}$, $\tilde{\varphi}_{ik}^{(m)}$ using (3). If they are all below tol, then go to Step 5, else go to Step 6.

5. For all $i \in \mathcal{S}$, set $\tilde{\Phi}_{ik}^{(m)} = (\tilde{\varphi}_{i1}^{(m)}, \ldots, \tilde{\varphi}_{ik}^{(m)})$. For $k > 1$, decide if $\tilde{\Phi}_{ik}^{(m)}$ for all $i \in \mathcal{S}$ are column rank deficient in the sense of (18). 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 $\tilde{\varphi}_{ik}^{(m)}, i = 1, 2, \ldots, r$ obtained from Step 4 by (2), 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.

Algorithm 3

1. Set $k = 1$, and define the set $\mathcal{S} = \{1, 2, \ldots, r\}$ and give a tolerance tol.

2. Start: Choose a real initial vector $v_1^{(k)}$ of norm one, and the steps m of Arnoldi's process and the steps n of the Chebyshev iteration, where $m > r$.

3. Perform m steps of Arnoldi's process starting with $v_1^{(k)}$ and compute the m eigenvalues of the resulting Hessenberg matrix. Then select the r eigenvalues $\lambda_{1k}^{(m)}, \ldots, \lambda_{rk}^{(m)}$ with largest real parts as approximations to $\lambda_1, \ldots, \lambda_r$, and set $R_{r,k}^{(m)}$ $= \{\lambda_{r+1k}^{(m)}, \ldots, \lambda_{mk}^{(m)}\}$.

4. Test convergence of r approximating eigenpairs $\lambda_{ik}^{(m)}, \tilde{\varphi}_{ik}^{(m)}$ using (3). If they are all below tol, then go to Step 8, else go to Step 5.

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

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

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 $\tilde{\Phi}_{ik}^{(m)} = (\tilde{\varphi}_{i1}^{(m)}, \ldots, \tilde{\varphi}_{ik}^{(m)})$. For $k > 1$, decide if $\tilde{\Phi}_{ik}^{(m)}$ for all $i \in \mathcal{S}$ are column rank deficient in the sense of (18). 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.

For details on Steps 5–7, we refer to [5, 16].

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 computing effort by the exploitation of symmetry.

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 $\{v_1, v_2, \ldots, v_m\}$ might be far from orthonormal. To maintain the mutually orthogonality of $\{v_l\}_1^m$, Saad [14] suggests to use the modified Gram-Schmidt method developed in [4] 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, Algorithms 2–3 are 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 [16] that in this case $\mathcal{K}_l(v_1, A)$ will be invariant, which implies, in particular, that the l Ritz pairs $\lambda_i^{(l)}, \varphi_i^{(l)}, i = 1, 2, \ldots, l$ are exact. Therefore, such breakdowns are "lucky" things. In finite precision, exact breakdowns are rare, but near breakdowns are possible. At this time, the right hand sides of (3) are very small. In terms of the theoretical analysis in Section 3, the l Ritz pairs are usually good approximations to some eigenpairs of A. 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 $v_1^{(k)}, k = 1, 2, \ldots$ 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 (6) is satisfied in practice.

Now we consider how to construct a new initial $v_1^{(k)}$ for Step 6 of Algorithms 2–3. First, we give the detail which performs Step 6 of Algorithm 2. Assuming v_1' at Step 6 to be a new starting vector, Saad [14] suggests the following combination:

$$\begin{aligned}
\alpha \cdot v_1' &= \sum_{i=1}^{r} \|(A - \lambda_{ik}^{(m)} I)\tilde{\varphi}_{ik}^{(m)}\| Re(\tilde{\varphi}_{ik}^{(m)}) && (19)\\
&= V_m \sum_{i=1}^{r} \|(A - \lambda_{ik}^{(m)} I)\tilde{\varphi}_{ik}^{(m)}\| V_m Re(y_{ik}^{(m)}) && (20)\\
&= V_m y, && (21)
\end{aligned}$$

where α is a normalizing factor. Thus, we only need to compute the eigenvectors $y_{ik}^{(m)}$, $i = 1, 2, \ldots, r$ of the Hessenberg matrix $H_m^{(k)}$ derived by Algorithm 1 starting with $v_1^{(k)}$.

Now let us look at the detail which performs Step 6 of Algorithm 3. Saad [16] suggests the following combination:

$$\alpha \cdot v_1' = \sum_{i=1}^{r} \xi_i \tilde{\varphi}_{ik}^{(m)} = \sum_{i=1}^{r} \xi_i V_m y_{ik}^{(m)} = V_m \sum_{i=1}^{r} \xi_i y_{ik}^{(m)} = V_m y, \tag{22}$$

where α is a normalizing factor. Thus, we only need to compute the eigenvectors $y_{ik}^{(m)}$, $i = 1, 2, \ldots, r$ of $H_m^{(k)}$ and select appropriate coefficients ξ_i at the kth restarting. Saad takes $\xi_i = \kappa_i^{-n}, i = 1, 2, \ldots, r$, where

$$\kappa_i = \left| \frac{T_n((\lambda_i^{(m)} - d)/c)}{T_n((\lambda_r^{(m)} - d)/c)} \right|^{1/n},$$

are the convergence ratios of $\lambda_i^{(m)}$ such that the new initial $v_1^{(k)}$ of Step 7 has the same components in the directions of $\tilde{\varphi}_{ik}^{(m)}, i = 1, 2, \ldots, r$. Here $T_n(z)$ is the first kind

Chebyshev polynomial of degree n, and d, c are, respectively, the center and focal distance of an ellipse which contains the set $R_{r,k}^{(m)}$ but $\{\lambda_{1k}^{(m)}, \ldots, \lambda_{rk}^{(m)}\}$. Refer to [5, 16] for more details.

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 a-priori unknown in practice. But according to Theorem 5, we can adopt the following criterion:

If restarting proceeds until the kth, we for the first time have the inequality

$$\sigma_{min}(\tilde{\Phi}_{ik}^{(m)}) \le \frac{\sqrt{k} \cdot C'_{ik}}{g'_{i,m}} \max_{1 \le j \le k} \|\tilde{r}_{ij}^{(m)}\|, \tag{23}$$

where C'_{ik} is a moderate factor, say, no greater than 1000, which means that the eigenvalue problem of A is not too ill-conditioned, and $g'_{i,m} = \min_{l \ne i} \mid \lambda_{l1}^{(m)} - \lambda_{i1}^{(m)} \mid$, then λ_i is $(k-1)$ multiple. 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. In this case, it is only necessary to determine the rank of $\tilde{\Phi}_{ik}^{(m)}$.

6 Numerical experiments

We report on several numerical examples to show reliability of the algorithms. They are performed using MATLAB4.2a on a DECstation with the 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. We construct the matrix $A = X\Lambda X^{-1}$, where

$$\Lambda = \mathrm{diag}(\begin{bmatrix} 1.9 & 0.5 \\ -2 & 1.9 \end{bmatrix}, \begin{bmatrix} 1.9 & 0.5 \\ -2 & 1.9 \end{bmatrix}, \begin{bmatrix} 1.9 & 0.5 \\ -2 & 1.9 \end{bmatrix}, 1.8, 1.6, 1.4, 1-(j-1)/90),$$

$$j = 10, 11, \ldots, 90,$$

and X is generated randomly, $\kappa(X) = \|X^{-1}\|\|X\| \approx 10628$. The eigenvalue problem of A is thus mildly ill-conditioned, A has two three multiple eigenvalues $\lambda_1 = 1.9 + i$ and $\lambda_2 = 1.9 - i$ and the rest eigenvalues are simple.

Algorithms 2–3 were run on such an 90×90 A. We want to find the five right-most eigenvalues of A and determine their corresponding multiplicities. The algorithms stop as soon as all actual residual norms drop below $tol = 10^{-9}$ and k satisfies condition (23), where we took all $C_{ik} = 100$. Tables 1–2 show the results and processes of determining λ_i, $d_i, i = 1, \ldots, 5$ obtained with Algorithms 2–3, respectively, where it

m	it	$m.v$	Residual norms				
			1	2	3	4	5
10	14	140	$9.8D-11$	$9.8D-11$	$1.3D-10$	$4.D-10$	$5.3D-10$
15	3	45	$1.5D-14$	$1.5D-14$	$1.4D-12$	$3.3D-11$	$2.2D-11$
12	7	84	$2.3D-13$	$2.3D-13$	$4.4D-12$	$7.5D-11$	$3.7D-11$
13	7	91	$3.8D-12$	$3.8D-13$	$1.7D-12$	$2.4D-10$	$2.2D-11$

m	k	$svd(\tilde{\Phi}_{1k}^{(m)})$	$svd(\tilde{\Phi}_{2k}^{(m)})$	$svd(\tilde{\Phi}_{3k}^{(m)})$	$svd(\tilde{\Phi}_{4k}^{(m)})$	$svd(\tilde{\Phi}_{5k}^{(m)})$
10	1	1	1	1	1	1
15	2	1.4142128 0.0014664	1.4142128 0.0014664	1.4142136 $4.2D-10$	1.4142136 $3.2D-9$	1.4142136 $5.1D-9$
12	3	1.7320414 0.0055557 0.0013792	1.7320414 0.0055557 0.0013792			
13	4	1.9999999 0.0057390 0.0026689 $6.2D-11$	1.9999999 0.0057390 0.0026689 $6.2D-11$			

	λ_1	λ_2	λ_3	λ_4	λ_5
Multiplicity	3	3	1	1	1

Table 1: *The process of determining the multiplicities of the five right-most eigenvalues of Example 1 using Algorithm 2.*

denotes the number of iterations and $svd(X)$ the set of all singular values of a matrix X. For $m = 12$ and $m = 8, n = 20$, the computed eigenvalues are

$$\lambda_{1,2} \approx 1.90000000000052 \pm 0.99999999999894i, \lambda_3 \approx 1.79999999998354,$$

$$\lambda_4 \approx 1.60000000004152, \lambda_5 \approx 1.39999999991501$$

and

$$\lambda_{1,2} \approx 1.9000000000007 \pm 1.00000000000004i, \lambda_3 \approx 1.79999999998809,$$

$$\lambda_4 \approx 1.60000000001141, \lambda_5 \approx 1.39999999997681.$$

It can be seen from Tables 1–2 that the algorithms solved Example 1 reliably.

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

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

m	n	it	$m.v$	Residual norms				
				1	2	3	4	5
8	20	4	92	$8.1D-15$	$8.1D-15$	$3.1D-12$	$4.5D-12$	$1.9D-11$
10	15	3	60	$4.6D-14$	$4.6D-14$	$5.8D-12$	$2.6D-11$	$3.7D-11$
11	10	3	53	$1.3D-15$	$1.3D-15$	$2.6D-13$	$1.5D-12$	$3.3D-12$
9	15	4	81	$4.5D-15$	$4.5D-15$	$1.3D-13$	$2.4D-13$	$1.D-13$

m	n	k	$svd(\tilde{\Phi}_{1k}^{(m)})$	$svd(\tilde{\Phi}_{2k}^{(m)})$	$svd(\tilde{\Phi}_{3k}^{(m)})$	$svd(\tilde{\Phi}_{4k}^{(m)})$	$svd(\tilde{\Phi}_{5k}^{(m)})$
8	20	1	1	1	1	1	1
10	15	2	1.4128785 0.6143862	1.4128785 0.6143862	1.4142136 $7.7D-11$	1.4142136 $6.6D-10$	1.4142136 $3.3D-10$
11	10	3	1.7305780 0.0713464 0.0030518	1.7305789 0.0713464 0.0030518			
9	15	4	1.9986014 0.0074719 0.0031036 $2.D-13$	1.9986014 0.0074719 0.0031036 $2.D-13$			

	λ_1	λ_2	λ_3	λ_4	λ_5
Multiplicity	3	3	1	1	1

Table 2: *The process of determining the multiplicities of the five right-most eigenvalues of Example 1 using Algorithm 3.*

$$j = 7, 8, \ldots, 90,$$

and X is generated randomly, $\kappa(X) \approx 1368$. Therefore, the eigenvalue problem of A is mildly ill-conditioned. A has two double multiple eigenvalues $\lambda_1 = 1.66, \lambda_2 = 1.62$ which are quite clustered, and the rest eigenvalues are simple.

Both algorithms were run on such an 90×90 A. We want to find the four right-most eigenvalues and determine their multiplicities. The stopping criterion was as in Example 1. Tables 3–4 show the results and processes of determining $\lambda_i, d_i, i = 1, 2, 3, 4$ obtained with Algorithms 2–3, respectively. For $m = 32$ and $m = 10, n = 30$, the computed eigenvalues are

$$\lambda_1 \approx 1.65999999998200, \lambda_2 \approx 1.61999999994564,$$

$$\lambda_3 \approx 1.30000000000078, \lambda_4 \approx 0.99999999999259$$

and

$$\lambda_1 \approx 1.66000000037090, \lambda_2 \approx 1.61999999971908,$$

$$\lambda_3 \approx 1.30000000006560, \lambda_4 \approx 1.00000000003075.$$

m	it	$m.v$	Residual norms			
			1	2	3	4
30	19	570	$7.5D-11$	$3.2D-11$	$1.3D-10$	$7.3D-11$
32	10	320	$8.9D-12$	$9.2D-11$	$3.4D-11$	$1.4D-11$
29	87	2523	$5.4D-10$	$8.6D-10$	$1.9D-10$	$1.1D-10$

m	k	$svd(\tilde{\Phi}_{1k}^{(m)})$	$svd(\tilde{\Phi}_{2k}^{(m)})$	$svd(\tilde{\Phi}_{3k}^{(m)})$	$svd(\tilde{\Phi}_{4k}^{(m)})$
30	1	1	1	1	1
32	2	1.4141428 0.0141505	1.4141441 0.0140142	1.4142136 $6.4D-10$	1.4142136 $2.7D-10$
29	3	1.7319786 0.0015821 $9.1D-9$	1.7315264 0.0042619 $2.5D-9$		

	λ_1	λ_2	λ_3	λ_4
Multiplicity	2	2	1	1

Table 3: *The process of determining the multiplicities of the four right-most eigenvalues of Example 2 using Algorithm 2.*

We see from Table 3 that Algorithm 2 found $\lambda_i, d_i, i = 1, 2, 3, 4$ reliably though the convergence is not satisfactory. Note that a minor change of m has a strong effect on the number of iterations. This is not surprising since Arnoldi's method is inefficient for clustered eigenvalues based on the theoretical analysis [7, 14, 15]. In fact, in the experiments, we found that no convergence occurred at cost of $it = 915$ if taking $m = 20$. On the contrary, it is obvious that Algorithm 3 solved Example 2 rapidly.

Example 3. We verify a symmetric matrix. Due to the unitary invariance of its eigenvalue problem, it is enough to consider diagonal matrices. Let us select a diagonal matrix A of order $N = 85$, with the following eigenvalue distribution:

$$\lambda_1 = 1.8, d_1 = 2;\ \ \lambda_2 = 1.1, d_2 = 2;$$

$$\lambda_{i-2} = \cos[(2i-5)\pi/2(N-2)], d_{i-2} = 1, \text{ for } i = 5, 6, \ldots, N.$$

Note that $\lambda_i, i = 3, 4, \ldots, N-2$, are clustered.

Algorithms 2–3, which reduce to the symmetric Lanczos algorithm and Lanczos-Chebyshev algorithm, were run on such a matrix A. We want to find the three largest eigenvalues and determine the corresponding multiplicities. We require that two algorithms stop as soon as all actual residual norms decrease below $tol = 10^{-7}$ and k

m	n	it	$m.v$	Residual norms			
				1	2	3	4
10	30	15	570	$5.7D-10$	$6.9D-10$	$1.6D-10$	$4.8D-11$
15	20	6	190	$1.5D-10$	$1.9D-10$	$4.8D-11$	$1.3D-12$
20	15	7	230	$3.9D-12$	$4.8D-12$	$1.3D-12$	$4.2D-13$

m	n	k	$svd(\tilde{\Phi}_{1k}^{(m)})$	$svd(\tilde{\Phi}_{2k}^{(m)})$	$svd(\tilde{\Phi}_{3k}^{(m)})$	$svd(\tilde{\Phi}_{4k}^{(m)})$
10	30	1	1	1	1	1
15	20	2	1.4047301 0.1635030	1.2134230 0.7263640	1.4142136 $6.1D-10$	1.4142136 $1.D-10$
20	15	3	1.7211508 0.1940105 $4.D-10$	1.5248784 0.8214292 $1.3D-9$		

	λ_1	λ_2	λ_3	λ_4
Multiplicity	2	2	1	1

Table 4: *The process of determining the multiplicities of the four right-most eigenvalues of Example 2 using Algorithm 3.*

satisfies condition (23), where all $C_{ik} = 1$. Tables 5–6 list the results computed by Algorithms 2–3. For $m = 15$ and $m = 15, n = 15$, the computed eigenvalues are

$$\lambda_1 \approx 1.80000000000000, \lambda_2 \approx 1.10000000000000, \lambda_3 \approx 0.99493483506257$$

and

$$\lambda_1 \approx 1.80000000000000, \lambda_2 \approx 1.10000000000000, \lambda_3 \approx 0.99493483506264.$$

As in Examples 1–2, the algorithms solved this symmetric multiple eigenproblem reliably. Note that in Table 6 the smallest singular value of $\tilde{\Phi}_{32}^{(m)}$ is of order 10^{-6}, which is about 25 times greater than $4.1D-8$. This is not surprising since

$$1/g_{3,m} \approx 1/g'_{3,m} \approx 220.$$

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

m	it	$m.v$	Residual norms		
			1	2	3
15	31	465	$4.8D-12$	$2.7D-9$	$7.5D-8$
20	13	260	$4.7D-15$	$1.4D-10$	$3.8D-8$
25	11	275	$1.8D-13$	$1.D-10$	$1.9D-8$

m	k	$svd(\tilde{\Phi}_{1k}^{(m)})$	$svd(\tilde{\Phi}_{2k}^{(m)})$	$svd(\tilde{\Phi}_{3k}^{(m)})$
15	1	1	1	1
20	2	1.0798203	1.3685855	1.4142136
		0.9132294	0.3563363	$7.D-7$
25	3	1.3987115	1.6946216	
		1.0215703	0.3581308	
		$2.2D-12$	$2.9D-9$	

	λ_1	λ_2	λ_3
Multiplicity	2	2	1

Table 5: *The process of determining the multiplicities of the three largest eigenvalues of Example 3 using Algorithm 2.*

7 Concluding remarks

We have proposed two 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 also 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 [9], other variants of Arnoldi's method, e.g. [13, 17] and the power method [19]. At the same time, we point out that this idea is essential for the block Arnoldi methods when the block size is less than or equal to the multiplicities of the required eigenvalues [7, 8]. Also, we have only described the algorithms for computing right-most eigenvalues and determine their multiplicities as well as corresponding eigenspaces. But they can be easily modified to work for other set of the spectrum, e.g. the eigenvalues with smallest real parts or with largest moduli.

Acknowledgements. This paper is based on part of [8]. I thank my superviser Professor L. Elsner very much 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.

m	n	it	$m.v$	Residual norms		
				1	2	3
15	15	9	255	$1.1D-14$	$2.9D-10$	$4.1D-8$
10	25	7	220	$1.4D-14$	$1.D-9$	$2.1D-8$
20	10	10	290	$3.9D-16$	$3.1D-10$	$2.8D-8$

m	n	k	$svd(\tilde{\Phi}_{1k}^{(m)})$	$svd(\tilde{\Phi}_{2k}^{(m)})$	$svd(\tilde{\Phi}_{3k}^{(m)})$
15	15	1	1	1	1
10	25	2	1.1478018 0.8261064	1.4064420 0.1480572	1.8290003 $1.D-6$
20	10	3	1.3794671 1.0474184 $1.D-14$	1.5054727 0.8564764 $2.6D-9$	

	λ_1	λ_2	λ_3
Multiplicity	2	2	1

Table 6: *The process of determining the multiplicities of the three largest eigenvalues of Example 3 using Algorithm 3.*

References

[1] W. E. Arnoldi, The principle of minimized iteration in the solution of the matrix eigenvalue problem, *Quart. Appl. Math.*, 9:11–29 (1951).

[2] F. Chatelin and S. Godet-Thobie, Stability analysis in aeronautical industries, *in Proceedings of the 2nd Symposium on High Performance Computing*, France, M. Durand and F. El. Dabaghi eds., Elsevier/North-Holland, 1991, pp. 415–422.

[3] J. Cullum and R. A. Willoughby, A practical procedure for computing eigenvalues of large sparse nonsymmetric matrices, *in Large Scale Eigenvalue Problem*, J. Cullum and R. A. Willoughby, eds., Elsevier/North-Holland, 1986, pp. 199–240.

[4] G. W. Daniel, W. B. Gragg, L. Kaufmann, and G. W. Stewart, Reorthogonalization and stable algorithms for updating the Gram-Schmidt QR factorization, *Math. Comput.*, 30:772–795 (1976).

[5] D. Ho, F. Chatelin and M. Bennani, Arnoldi-Tchebychev procedure for large nonsymmetric matrices, *Math. Mod. Numer. Anal.*, 24:53–65 (1990).

[6] Z. Jia, The convergence of generalized Lanczos methods for large unsymmetric eigenproblems, *SIAM J. Matrix Anal. Appl.*, 16 (1995), to appear.

[7] Z. Jia, Generalized block Lanczos methods for large unsymmetric eigenproblems, *Submitted for publication*, 1994.

[8] Z. Jia, *Some Numerical Methods for Large Unsymmetric Eigenproblems*, Ph.D. Thesis, University of Bielefeld, Bielefeld, 1994.

[9] C. Lanczos, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, *J. Res. Nat. Bur. Stand.*, 45:225–280 (1950).

[10] B. N. Parlett and D. S. Scott, The Lanczos algorithm with selective orthogonalization, *Math. Comput.*, 33:217–238 (1979).

[11] B. N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs, NJ., 1980.

[12] A. Ruhe, The two sided Arnoldi algorithm for nonsymmetric eigenvalue problems, *Matrix Pencils*, B. K. Kågström and A. Ruhe, eds., Lecture Notes in Mathematics, Springer–Verlag, Berlin, 973:104–120 (1983).

[13] A. Ruhe, Rational Krylov algorithms for nonsymmetric eigenvalue problems. II. Matrix pairs, *Linear Algebra Appl.*, 197/198:283–295 (1994).

[14] Y. Saad, Variations on Arnoldi's method for computing eigenelements of large nonsymmetric matrices, *Linear Algebra Appl.*, 34:269–295 (1980).

[15] Y. Saad, Projection methods for solving large eigenvalues problems, *Matrix Pencils*, Lecture Notes in Mathematics, B. K. Kågström and A. Ruhe, eds., Springer–Verlag, Berlin, 973:121–144 (1983).

[16] Y. Saad, Chebyshev acceleration techniques for solving nonsymmetric eigenvalue problems, *Math. Comput.*, 42:567–588 (1984).

[17] D. C. Sorensen, Implicit application of polynomial filters in a k-step Arnoldi method, *SIAM J. Matrix Anal. Appl.*, 13:357–385 (1992).

[18] G. W. Stewart and J-G. Sun, *Matrix Perturbation Theory*, Academic Press, Boston, 1990.

[19] J. H. Wilkinson, *The Algebraic Eigenvalue Problem,* Clarendon Press, Oxford, 1965.