Numer. Math. 67: 21–40 (1994)

Numerische Mathematik

An adaptive Chebyshev iterative method for nonsymmetric linear systems based on modified moments

D. Calvetti[1,⋆], G.H. Golub[2,⋆⋆], and L. Reichel[3,⋆⋆⋆]

[1] Department of Pure and Applied Mathematics, Stevens Institute of Technology, Hoboken, NJ 07030, USA
[2] Department of Computer Science, Stanford University, Stanford, CA 94305, USA
[3] Department of Mathematics and Computer Science, Kent State University, Kent, OH 44242, USA

Received May 13, 1992/Revised version received May 13, 1993

Summary. Large, sparse nonsymmetric systems of linear equations with a matrix whose eigenvalues lie in the right half plane may be solved by an iterative method based on Chebyshev polynomials for an interval in the complex plane. Knowledge of the convex hull of the spectrum of the matrix is required in order to choose parameters upon which the iteration depends. Adaptive Chebyshev algorithms, in which these parameters are determined by using eigenvalue estimates computed by the power method or modifications thereof, have been described by Manteuffel [18]. This paper presents an adaptive Chebyshev iterative method, in which eigenvalue estimates are computed from modified moments determined during the iterations. The computation of eigenvalue estimates from modified moments requires less computer storage than when eigenvalue estimates are computed by a power method and yields faster convergence for many problems.

Mathematics Subject Classification (1991): 65F10

1. Introduction

The problem of solving a linear system of equations

$$A\mathbf{x} = \mathbf{b}\,, \quad A \in \mathbb{R}^{N\times N}\,, \quad \mathbf{x}, \mathbf{b} \in \mathbb{R}^N \tag{1.1}$$

with a large, sparse and nonsymmetric matrix A arises in many applications. A Chebyshev iterative method based on scaled Chebyshev polynomials p_n for an interval in the complex plane can be used to solve (1.1) when the spectrum of A lies in the right half plane. This includes matrices with a positive definite symmetric part. Manteuffel

Correspondence to: L. Reichel
⋆ Research supported in part by the Design and Manufacturing Institute at Stevens Institute of Technology
⋆⋆ Research supported in part by NSF grant CCR-8821078
⋆⋆⋆ Research supported in part by NSF grant DMS-9002884

[17, 18] discusses such Chebyshev iterative schemes and shows that the iterations depend on two parameters only, the center d and the focal length c of an ellipse in the complex plane $\mathbb{C}$ with foci at $d \pm c$. In these schemes, the p_n are Chebyshev polynomials for the interval between the foci, and are scaled so that $p_n(0) = 1$. The three-term recurrence relation for the p_n yields an inexpensive recurrence relation for computing a sequence of approximate solutions $\mathbf{x}_n$, $n = 1, 2, \ldots$, of (1.1). Let $\mathbf{x}_0$ denote a given initial approximate solution of (1.1), and introduce the residual vectors $\mathbf{r}_n := \mathbf{b} - A\mathbf{x}$, $n \geq 0$. The iterates x_n determined by the Chebyshev iterative method are such that

$$\mathbf{e}_n = p_n(A)\mathbf{e}_0 , \quad n \geq 0 , \tag{1.2}$$

where $\mathbf{e}_n$ denotes the error in $\mathbf{x}_n$, i.e.,

$$\mathbf{e}_n := \mathbf{x}^* - \mathbf{x}_n , \quad \mathbf{x}^* := A^{-1}\mathbf{b} .$$

Let the matrix A be diagonalizable and have spectral decomposition

$$A = W\Lambda W^{-1} , \quad \Lambda = \mathrm{diag}[\lambda_1\lambda_2, \ldots, \lambda_N] , \quad W = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_N] , \tag{1.3}$$

where the eigenvectors $\mathbf{w}_j$ are scaled so that $\|\mathbf{w}_j\| = 1$. Throughout this paper $\|\cdot\|$ denotes the Euclidean vector norm or the corresponding induced matrix norm. Let $\mathscr{S}(A)$ denote the spectrum of A. It follows from (1.3) that the error $\mathbf{e}_n$ can be bounded by

$$\|\mathbf{e}_n\| \leq \|W\| \, \|p_n(\Lambda)\| \, \|W^{-1}\| \, \|\mathbf{e}_0\| = \|W\| \, \|W^{-1}\| \, \|\mathbf{e}_0\| \max_{\lambda \in \mathscr{S}(A)} |p_n(\lambda)| . \tag{1.4}$$

We obtain from (1.2) that

$$\mathbf{r}_n = p_n(A)\mathbf{r}_0 , \quad n \geq 0 , \tag{1.5}$$

and, therefore, a bound similar to (1.4) holds for the residual vectors, also. Because of relation (1.5), the p_n are sometimes referred to as residual polynomials. If the parameters d and c are chosen so that the quantity

$$\max_{\lambda \in \mathscr{S}(A)} |p_n(\lambda)| \tag{1.6}$$

decreases rapidly with n, then, by (1.4), the norm $\|\mathbf{e}_n\|$ decreases rapidly as n increases; see, e.g., [17] for details, where the case when A cannot be diagonalized is treated, also. For pronouncedly nonnormal matrices A, i.e., when $\|W\| \, \|W^{-1}\|$ is "huge", it may be meaningful to consider pseudospectra of A instead of the spectrum; see [19] for a discussion. For simplicity, we will in the present paper only discuss convergence in terms of the spectrum $\mathscr{S}(A)$. For n sufficiently large, the scaled Chebyshev polynomials p_n for the interval between the foci at $d \pm c$ are of nearly constant magnitude on the boundary of any ellipse, which is not an interval, with foci at $d \pm c$. Chebyshev iteration is an attractive solution method if parameters d and c exist, such that there is an ellipse with foci at $d \pm c$ which contains $\mathscr{S}(A)$ and is not very close to the origin. In particular, this ellipse must not contain the origin. Assuming that such an ellipse exists, its center d and focal length c can be determined if $\mathscr{S}(A)$ is explicitly known. However, in general, $\mathscr{S}(A)$ is neither known nor easy to determine.

In [18] Manteuffel describes algorithms for dynamic estimation of the parameters d and c based on the power method applied to A, or modifications thereof. The parameters d and c are chosen so that $d \pm c$ are the foci of the the smallest ellipse containing available estimates of eigenvalues of A. As new estimates of the eigenvalues of A become available during the iterations, it may be necessary to refit the ellipse so that it encloses all available eigenvalue estimates of A. Manteuffel [18] proposes a combinatorial approach for fitting the ellipse. More recently other schemes have also been suggested; see [4, 16]. We review Manteuffel's adaptive Chebyshev algorithms in Sect. 2.

A modification of Manteuffel's adaptive schemes is proposed by Elman et al. [7], who replace the power method and its modifications by the Arnoldi process and the GMRES algorithm. The Arnoldi process is applied to compute eigenvalue estimates of A, and these estimates are used to determine new parameters d and c. We remark that these eigenvalue estimates obtained are the same as the eigenvalue estimates that Manteuffel computes by the power method [17]. Having computed eigenvalue estimates by the Arnoldi process, the best available approximate solution of (1.1), say $\mathbf{x}_n$, can be improved quite inexpensively by the GMRES algorithm before restarting Chebyshev iteration with the new parameters d and c. The scheme proposed by Elman et al. [7] is a hybrid iterative method because it combines Chebyshev iteration with the GMRES algorithm. A recent survey of hybrid iterative schemes can be found in [19].

This paper presents an adaptive Chebyshev algorithm that uses modified moments to compute approximations of eigenvalues of A. The computed modified moments and the recursion coefficients of the p_n are input to the modified Chebyshev algorithm, which determines a nonsymmetric tridiagonal matrix. We compute the eigenvalues of this tridiagonal matrix and consider them as estimates of eigenvalues of A. These estimates are used to compute parameters d and c by determining the smallest ellipse that contains the estimates. From the location of the foci at $d \pm c$ of this ellipse, the parameters d and c can easily be computed. We remark that the tridiagonal matrix determined by our algorithm also can be computed by the nonsymmetric Lanczos process. However, the latter would require the computation of matrix-vector products with both the matrices A and A^T; our scheme only requires the evaluation of matrix-vector products with the matrix A.

The computation of each modified moment requires the evaluation of an inner product of two N-vectors. The adaptive procedure that we describe in this paper requires 2κ modified moments to estimate κ eigenvalues of A. The simultaneous calculation of iterates $\mathbf{x}_n$ and modified moments makes it possible to compute new eigenvalue estimates from modified moments and refit the ellipse in order to determine new values for the parameters d and c as soon as the rate of convergence of the computed residual vectors $\mathbf{r}_n$ falls below a certain tolerance.

Our numerical experiments indicate that modified moments only have to be computed during the first couple of iterations in order to determine parameters d and c that yield a high rate of convergence. When such parameters have been found, the iterations can proceed without computing further modified moments, and therefore without computing further inner products, until an accurate approximate solution of (1.1) has been found. Typically, the vast majority of the iterations can be carried out without computing modified moments and inner products. The simplicity of Chebyshev iteration with fixed parameters d and c allows efficient implementations on parallel and vector computers; see Dongarra et al. [5, Chap. 7.1.6] for a recent discussion.

Our scheme for computing modified moments and eigenvalue estimates for a nonsymmetric matrix A is an extension of an algorithm described by Golub and Kent [12] for the computation of modified moments and eigenvalue estimates for a symmetric matrix. The computation of modified moments requires the residual vector $\mathbf{r}_0$ be available. This is the only N-vector that our adaptive Chebyshev algorithm requires stored, in addition to the N-vectors required by nonadaptive Chebyshev iteration. We note that the adaptive Chebyshev algorithms proposed by Manteuffel and Ashby [18, 2] and by Elman et al. [7] require more N-vectors to be stored than our schemes. Details of the storage requirements are discussed in Sect. 2.

This paper is organized in the following way. In Sect. 2 we outline nonadaptive Chebyshev iteration and schemes used by Manteuffel [18] and Elman et al. [7] for determining eigenvalue estimates of A. The problem of determining the ellipse that encloses the eigenvalue estimates and yields the smallest convergence factor is treated in Sect. 3. This section follows the presentation by Manteuffel [18]. In Sect. 4 we discuss how inner products of residual vectors can be interpreted as modified moments, and we define a nonsymmetric tridiagonal matrix $H_\kappa \in \mathbb{R}^{\kappa\times\kappa}$ whose entries are the recurrence coefficients of monic polynomials that are orthogonal with respect to a bilinear form defined by the modified moments. This section extends results by Golub and Kent [12]. In Sect. 5 we study some properties of modified moments with respect to a complex measure with support in $\mathbb{C}^2$, and derive the modified Chebyshev algorithm. Our presentation follows Golub and Gutknecht [12]. We use the modified Chebyshev algorithm to compute the elements of the matrix H_κ from the modified moments and the recurrence coefficients of the residual polynomials. The eigenvalues of H_κ approximate eigenvalues of A and are used to determine suitable parameters d and c for Chebyshev iteration. In Sect. 6 we show that the tridiagonal matrices H_κ that are generated by our algorithm are similar to tridiagonal matrices that can be determined by the nonsymmetric Lanczos process. The results of numerical experiments comparing our adaptive scheme with a scheme due to Manteuffel as implemented by Ashby and Manteuffel [2] are presented in Sect. 7.

2. Adaptive Chebyshev algorithms

In this section we outline the adaptive Chebyshev algorithms by Manteuffel [17, 18] and Elman et al. [7] and introduce notation that will be used in the remainder of the paper. A more detailed discussion of the material presented can be found in [2, 7, 17, 18]. Given the two parameters d and c, Chebyshev iteration for (1.1) can be defined as follows. Let $\mathbf{x}_0$ be the initial approximate solution, let $\mathbf{r}_0 := \mathbf{b} - A\mathbf{x}_0$ and $\boldsymbol{\Delta}_0 := \frac{1}{d}\mathbf{r}_0$. The iterates $\mathbf{x}_n$ for $n = 1, 2, \ldots$ are defined by

$$\begin{aligned} \mathbf{x}_n &:= \mathbf{x}_{n-1} + \boldsymbol{\Delta}_{n-1} , \\ \mathbf{r}_n &:= \mathbf{b} - A\mathbf{x}_n , \\ \boldsymbol{\Delta}_n &:= \alpha_n \mathbf{r}_n + \beta_n \boldsymbol{\Delta}_{n-1} , \end{aligned} \tag{2.1}$$

where

$$\alpha_n := \frac{2}{c}\,\frac{T_n\left(\frac{d}{c}\right)}{T_{n+1}\left(\frac{d}{c}\right)} , \quad \beta_n := \frac{T_{n-1}\left(\frac{d}{c}\right)}{T_{n+1}\left(\frac{d}{c}\right)} , \tag{2.2}$$

and $T_n(\lambda)$ is the Chebyshev polynomial

$$T_n(\lambda) := \cosh\left(n \cosh^{-1}(\lambda)\right) \ .$$

The residual polynomials p_n in (1.2) and (1.5) are given by

$$p_n(\lambda) = \frac{T_n\left(\frac{d-\lambda}{c}\right)}{T_n\left(\frac{d}{c}\right)} \ . \tag{2.3}$$

Let d and c be the center and focal length, respectively, of the smallest ellipse containing $\mathscr{S}(A)$. The assumption that $\mathscr{S}(A)$ lies in the right half plane and is symmetric with respect to the real axis implies that $d > 0$ and $d^2 > c^2$. It therefore suffices to consider pairs of real numbers (d, c^2) that lie in

$$\mathscr{R} := \left\{(d, c^2) : d > 0,\ d^2 > c^2\right\} \ . \tag{2.4}$$

For each $\lambda \in \mathbb{C}$ define the *asymptotic convergence factor*

$$r(\lambda, d, c^2) := \left| \frac{d - \lambda + ((d-\lambda)^2 - c^2)^{1/2}}{d + (d^2 - c^2)^{1/2}} \right| \ . \tag{2.5}$$

It can be shown that for large n the component of the error $\mathbf{e}_n$ in the direction of the eigenvectors $\mathbf{w}_j$, cf. (1.3), in each iteration is multiplied by a factor of magnitude roughly equal to $r(\lambda_j, d, c^2)$. Therefore, for sufficiently large n, the dominating eigenvector components of $\mathbf{e}_n$ are in the directions of eigenvectors $\mathbf{w}_j$ associated with eigenvalues λ_j with largest convergence factor (2.5). From the relation $\mathbf{r}_n = A\mathbf{e}_n$, it follows that for n sufficiently large the residual error $\mathbf{r}_n$ also is dominated by the same eigenvector components. It is desirable to choose d and c so that the asymptotic convergence factor (2.5) associated with each eigenvalue of A is small. This suggests to let the parameters d and c be the solution of the mini-max problem

$$\min_{(d,c^2)\in\mathscr{R}} \max_{\lambda\in\mathscr{S}(A)} r(\lambda, d, c^2) \ . \tag{2.6}$$

The eivenvalues of A with largest convergence factors are vertices of the convex hull of $\mathscr{S}(A)$. Let $\mathscr{H}(A)$ denote the set of vertices of the convex hull of $\mathscr{S}(A)$. The solution of (2.6) is a function only of the eigenvalues of A in $\mathscr{H}(A)$. Therefore, Manteuffel's adaptive Chebyshev schemes, as well as our algorithm, seek to determine estimates of the elements of $\mathscr{H}(A)$.

In one of the adaptive schemes described in [18], the power method is applied to A in order to approximate eigenvalues of the matrix. When the power method is applied to A, eigenvalues of large magnitude are typically determined most accurately. Therefore, it may be difficult to determine the vertices of $\mathscr{H}(A)$ that are closest to the origin with high accuracy in this manner. Manteuffel [18] suggests the following approach to circumvent this problem. Let

$$B(z) := \frac{d - z + ((d-z)^2 - c^2)^{1/2}}{d + (d^2 - c^2)^{1/2}} \tag{2.7}$$

and apply the power method to $B(A)$ in order to estimate the eigenvalues of largest magnitude of this operator. If b_j is an eigenvalue of $B(A)$ and

$$g := d + (d^2 - c^2)^{1/2} \ ,$$

then

$$\lambda_j := d - \frac{1}{2}\left(gb_j + \frac{c^2}{gb_j}\right) \tag{2.8}$$

is an eigenvalue of A and $r(\lambda_j, d, c^2) = |B(b_j)|$. Therefore, the power method applied to $B(A)$ typically yields highest accuracy for eigenvalues of A with largest convergence factor. In order to avoid the nonlinearity of the relation (2.8) between λ_j and b_j, Manteuffel [18] also proposed to apply the power method to the matrix

$$\hat{A} := 2g(dI - A) = g^2 B(A) + c^2[B(A)]^{-1} . \tag{2.9}$$

The linearity of the relation between A and $\hat{A}$ makes it simple to compute the eigenvalue of A from those of $\hat{A}$, and, moreover, the power method applied to $\hat{A}$ typically yields highest accuracy for eigenvalues with largest convergence factor.

In the implementation of adaptive Chebyshev iteration by Ashby and Manteuffel [2], the power method is applied to A, $B(A)$ or $\hat{A}$ in order to determine four estimates of eigenvalues. The implementation [2] based on the power method applied to A or $B(A)$ requires the storage of four N-vectors in addition to the vectors used for nonadaptive Chebyshev iteration (2.1). The implementation [2] based the power method applied to $\hat{A}$ requires the storage of five additional N-vectors. The adaptive schemes based on the different power methods also differ in their operation count. The power method applied to A is, generally, the most expensive scheme; it requires four matrix-vector products with the matrix A for the computation of each set of four eigenvalue estimates, in addition to the matrix-vector products needed to compute the iterates $\mathbf{x}_n$ by Chebyshev iteration.

An alternative to the power methods for computing estimates of eigenvalues of A is provided by the Arnoldi process [1]. This method is a Galerkin scheme for approximating m eigenvalues of A, in which the test and trial spaces are a Krylov subspace

$$K_m(A, \mathbf{v}) := \text{span}\{\mathbf{v}, A\mathbf{v}, \ldots, A^{m-1}\mathbf{v}\} , \tag{2.10}$$

where $\mathbf{v}$ is a vector in $\mathbb{R}^N$. The scheme requires the storage of an orthonormal basis of $K_m(A, v)$, and the operation count is $O(m^2N)$. In the application of the Arnoldi process to the computation of eigenvalue estimates for Chebyshev iteration discussed in [7] m is chosen to be four; see [7] for further details.

3. Fitting of the ellipse

In this section we discuss how to compute the parameters d and c for Chebyshev iteration. Let $\mathscr{F}(d, c, a)$ denote the ellipse with center d, focal length c and semi-major axis $a \geq 0$, i.e.,

$$\mathscr{F}(d, c, a) := \{z \in \mathbb{C} : |z - d + c| + |z - d - c| \leq 2a\} .$$

Let Π_n be the set of polynomials of degree at most n, and also define the subset $\tilde{\Pi}_n := \{q : q \in \Pi_n, q(0) = 1\}$. The following theorem shows that the residual polynomials p_n for Chebyshev iteration with parameters d and c minimize the limit, as $n \to \infty$, of the n^{th} root of the quantity (1.6) when $\mathscr{S}(A) := \mathscr{F}(d, c, a)$.

Theorem 1. ([17]) *Assume that* $0 \notin \mathscr{F}(d,c,a)$. *Let* $t_n \in \tilde{\Pi}_n$ *satisfy*

$$\max_{\lambda \in \mathscr{F}(d,c,a)} |t_n(\lambda)| = \min_{q \in \tilde{\Pi}_n} \max_{\lambda \in \mathscr{F}(d,c,a)} |q(\lambda)| ,$$

and let p_n *be given by (2.3). Then*

$$\lim_{n\to\infty} \left[\max_{\lambda \in \mathscr{F}(d,c,a)} |t_n(\lambda)| \right]^{1/n} = \lim_{n\to\infty} \left[\max_{\lambda \in \mathscr{F}(d,c,a)} |p_n(\lambda)| \right]^{1/n} . \tag{3.1}$$

Moreover,

$$\lim_{n\to\infty} |p_n(\lambda)|^{1/n} = r(\lambda, d, c^2)$$

for any $\lambda \in \mathbb{C}$ *and, in particular,*

$$\lim_{n\to\infty} |p_n(\lambda)|^{1/n} = r\left(d + \frac{c^2}{|c|^2}\, a, d, c^2 \right) < 1 \tag{3.2}$$

for any λ *on the boundary of* $\mathscr{F}(d,c,a)$. □

In the terminology used, e.g., in [17, 22], equality (3.1) shows that the residual polynomials p_n given by (2.3) yield an asymptotically optimal rate of convergence with respect to $\mathscr{F}(d,c,a)$. We note that this result can be improved in several ways. Since the polynomials p_n are scaled Faber polynomials for $\mathscr{F}(d,c,a)$, results by Eiermann [6] on Faber polynomials imply that

$$\max_{\lambda \in \mathscr{F}\, d,c,a)} |p_n(\lambda)| / \max_{\lambda \in \mathscr{F}(d,c,a)} |t_n(\lambda)| \le \gamma \log n , \quad n \ge 1 , \tag{3.3}$$

for some constant γ independent of n. Moreover, Fischer and Freund [8] have recently shown that for many ellipses the right hand side of (3.3) can be replaced by one.

Formula (3.2) shows that, if d and c are the center and focal length of a small ellipse enclosing the spectrum of A, and if this ellipse is not very close to the origin, then $\max_{\lambda \in \mathscr{S}(A)} |p_n(\lambda)|$ converges rapidly to zero as n increases. Moreover, if the matrix A is not very far from normal, then (1.4) shows that the norm $\|\mathbf{e}_n\|$ also converges rapidly to zero with increasing n.

We now outline the scheme of Manteuffel [17] for computing the best ellipse with respect to the spectrum $\mathscr{S}(A)$, i.e., we compute the best parameters d and c for Chebyshev iteration, when the spectrum $\mathscr{S}(A)$ is given. However, we remark that when carrying out Chebyshev iteration, $\mathscr{S}(A)$ is generally not known. The adaptive Chebyshev algorithm therefore computes the best ellipse with respect to a set of eigenvalue estimates computed during the iterations. This set is typically updated a few times during the iterations.

Since A is real, the set $\mathscr{H}(A)$ is symmetric with respect to the real axis, and the foci of the smallest ellipse enclosing $\mathscr{H}(A)$ are either real or complex conjugate. The center d and focal length c of the smallest ellipse containing $\mathscr{H}(A)$ are such that $(d, c^2) \in \mathscr{R}$, where $\mathscr{R}$ is defined by (2.4). Thus, the mini-max problem (2.6) can be replaced by the simpler mini-max problem (3.4)

$$\min_{(d,c^2) \in \mathscr{R}} \max_{\lambda \in \mathscr{H}^+(A)} r(\lambda, d, c^2) , \tag{3.4}$$

where $\mathscr{H}^+(A) := \{\lambda \in \mathscr{H}(A) : \mathrm{Im}(\lambda) \ge 0\}$. The theorem below is helpful for the solution of (3.4).

Theorem 2. ([3, 17]) *Let the set $\mathscr{M} \subset \mathscr{R}$ be closed and bounded, and let $\mathscr{S}(A) = \{\lambda_i\}_{i=1}^N$. Then $\{r(\lambda_i, d, c^2)\}_{i=1}^N$ is a finite set of real-valued functions of two variables (d, c^2), continuous on $\mathscr{M}$. Let $m(d, c^2) := \max_i(\lambda_i, d, c^2)$. Then $m(d, c^2)$ has a minimum at some point (d_0, c_0^2). If (d_0, c_0^2) is in the interior of $\mathscr{M}$ then one of the following statements holds:*

1. *The point (d_0, c_0^2) is a local minimum of $r(\lambda_i, d, c^2)$ for some i such that $r(\lambda_i, d_0, c_0^2) = m(d_0, c_0^2)$.*
2. *The point (d_0, c_0^2) is a local minimum among the loci $\{(d, c^2) \in \mathscr{M} : r(\lambda_i, d, c^2) = r = (\lambda_j, d, c^2)\}$ for some i and j such that $m(d_0, c_0^2) = r(\lambda_j, d_0, c_0^2) = r(\lambda_i, d_0, c_0^2)$.*
3. *The point (d_0, c_0^2) is such that for some i, j and k, $m(d_0, c_0^2) = r(\lambda_i, d_0, c_0^2) = r(\lambda_j, d_0, c_0^2) = r(\lambda_k, d_0, c_0^2)$.*

Manteuffel [17] presents an algorithm for the solution of (3.4) based on the following observations that are a consequence of Theorem 2.

1. If $\mathscr{H}^+(A) = \{\lambda_1\}$, then $d = x_1$ and $c^2 = -y_1^2$, where $\lambda_1 = x_1 + \mathrm{i}y_1$, $\mathrm{i} := \sqrt{-1}$.

2. If $\mathscr{H}^+(A) = \{\lambda_1, \lambda_2\}$, then the optimal parameters (d, c^2) correspond to a point lying on the intersection of the two surfaces

$$r(\lambda_1, d, c^2) = r(\lambda_2, d, c^2) \ .$$

The point corresponding to the optimal parameters when the positive convex hull contains only two points is called the *pairwise best point*, and the associated ellipse passing through λ_1 and λ_2 is called the *pairwise best ellipse*.

3. If $\mathscr{H}^+(A)$ contains three or more eigenvalues, then the solution to (3.4) must be either a pairwise best point or it is the intersection of three surfaces. Given the pairwise best point for two eigenvalues λ_1 and λ_2, this is the best point if the associated pairwise best ellipse contains all eigenvalues in the closure of its interior. If no pairwise best point is the solution to (3.4), then determine the *three-way point* on the intersection of the three surfaces:

$$r(\lambda_1, d, c^2) = r(\lambda_2, d, c^2) = r(\lambda_3, d, c^2)$$

and the associated *three-way ellipse*. If the associated three-way ellipse contains all eigenvalues of A in the closure of its interior then the three-way point is a feasible point. The three-way feasible point with smallest convergence factor is the solution to the mini-max problem (3.4).

A detailed description, and some further simplifications, of the scheme for fitting the ellipse outlined above are presented in [17].

4. Modified moments

In this section we define moments and modified moments and discuss how they can be used to gain spectral information about the matrix A while computing approximate solutions $\mathbf{x}_n$ of (1.1) by Chebyshev iteration (2.1). Let A have spectral resolution (1.3), let p_n be residual polynomials (2.3) and express $\mathbf{r}_0$ in the basis of eigenvectors $\{\mathbf{w}_j\}_{j=1}^N$, i.e.,

$$\mathbf{r}_0 = \sum_{i=1}^{N} \alpha_i \mathbf{w}_i \ .$$

Then it follows from $p_n(A) = W p_n(\Lambda) W^{-1}$ that

$$\mathbf{r}_n = \sum_{i=1}^{N} \alpha_i p_n(\lambda_i) \mathbf{w}_i \ .$$

Introduce the inner product

$$(\mathbf{r}_k, \mathbf{r}_l) = \mathbf{r}_k^{\mathrm{T}} \mathbf{r}_l = \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j \mathbf{w}_i^{\mathrm{T}} \mathbf{w}_j p_k(\lambda_i) p_l(\lambda_j) \ ,$$

and let $\gamma(\lambda, \eta)$ be the complex symmetric measure with support in $\mathbb{C}^2$ and with "jumps" of height $\gamma_{ij} := \alpha_i \alpha_j \mathbf{w}_i^{\mathrm{T}} \mathbf{w}_j$ at the points $(\lambda_i, \lambda_j) \in \mathbb{C}^2$, for $\lambda_i, \lambda_j \in \mathscr{S}(A)$. Then the inner product can be written as

$$(\mathbf{r}_k, \mathbf{r}_l) = \sum_{i=1}^{N} \sum_{j=1}^{N} p_k(\lambda_i) p_l(\lambda_j) \gamma_{ij} = \int_{\mathbb{C}} \int_{\mathbb{C}} p_k(\lambda) p_l(\eta) d\gamma(\lambda, \eta) \ . \tag{4.1}$$

We remark that if the matrix A is symmetric, then its eigenvalues are all real and we can choose the eigenvector matrix W to be orthogonal. In this case (4.1) simplifies to

$$(\mathbf{r}_k, \mathbf{r}_l) = \sum_{i=1}^{N} \alpha_i^2 p_k(\lambda_i) p_l(\lambda_i) = \int_{\mathbb{R}} p_k(\lambda) p_l(\lambda) d\alpha(\lambda) \ , \tag{4.2}$$

where α is a real measure with support on $\mathbb{R}$ and with jumps of height α_i^2 at $\lambda_i \in \mathscr{S}(A)$.

Let

$$\tilde{\gamma}_i := \sum_{j=1}^{N} \gamma_{ij} \ , \quad 1 \le i \le N \ , \tag{4.3}$$

and introduce the linear functional ϕ associated with the measure γ by

$$\phi(q) := \sum_{i=1}^{N} \tilde{\gamma}_i q(\lambda_i) \ , \tag{4.4}$$

where q is a polynomial. We are now in a position to define *moments* associated with the measure γ by

$$\mu_k := \phi(\lambda^k) \ , \quad k = 0, 1, 2, \ldots , \tag{4.5}$$

as well as *modified moments* with respect to the residual polynomials

$$\nu_k := \phi(p_k) \ , \quad k = 0, 1, 2, \ldots \quad . \tag{4.6}$$

For future references we note that

$$\nu_k = (\mathbf{r}_k, \mathbf{r}_0) \ , \quad k = 0, 1, 2, \ldots \quad . \tag{4.7}$$

Let $\langle\langle \cdot , \cdot \rangle\rangle$ dnote the bilinear form generated by ϕ,

$$\langle\langle f, g\rangle\rangle := \phi(fg) \,, \tag{4.8}$$

where f and g are polynomials, and assume that there is a family of monic formal orthogonal polynomials $\{\pi_i\}_{i=0}^{N}$ associated with the measure γ, i.e.,

$$\langle\langle \pi_k, \pi_l\rangle\rangle \begin{cases} = 0 & \text{if } k \neq l \,, \quad 0 \leq k \,, \quad l \leq N \,, \\ \neq 0 & \text{if } k = l \,, \quad 0 \leq k < N \,. \end{cases}$$

The bilinear form (4.8) has the property that

$$\langle\langle zf, g\rangle\rangle = \langle\langle f, zg\rangle\rangle \,,$$

and this implies that the π_k satisfy a three-term recurrence relation

$$\begin{aligned} &\pi_{k+1}(z) = (z - \hat{\alpha}_k)\pi_k(z) - \hat{\beta}_k \pi_{k-1}(z) \,, \quad 0 \leq k < N \,, \\ &\pi_0(z) = 1 \,, \quad \pi_{-1}(z) = 0 \,, \end{aligned} \tag{4.9}$$

see, e.g., [9] for a proof. Introduce the matrices

$$H_\kappa := \begin{bmatrix} \hat{\alpha}_0 & \hat{\beta}_1 & & & 0 \\ 1 & \hat{\alpha}_1 & \hat{\beta}_2 & & 0 \\ & 1 & \ddots & \ddots & \\ & & \ddots & & \hat{\beta}_{\kappa-1} \\ 0 & & & 1 & \hat{\alpha}_{\kappa-1} \end{bmatrix} \,, \quad 1 \leq \kappa \leq N \,. \tag{4.10}$$

In the following section we show how to construct the tridiagonal matrices (4.10) from the modified moments (4.6) and the parameters d and c of Chebyshev iteration. We show in Sect. 6 that the matrices H_κ are similar to the matrices determined by the nonsymmetric Lanczos process. This motivates the use of eigenvalues of H_κ to determine an ellipse for Chebyshev iteration. A discussion based on formula (4.2) on how modified moments can be applied to estimate eigenvalues of a symmetric positive definite matrix can be found in [12].

5. Computing eigenvalue estimates

This section derives a matrix identity that connects modified moments associated with a complex measure with the recurrence coefficients of a family or orthogonal polynomials associated with this measure. This identity is the basis of the modified Chebyshev algorithm for computing recursion coefficients for the orthogonal polynomials, and has been discussed in more detail, also including degenerate cases, by Golub and Gutknecht [11]. The recursion coefficients determine tridiagonal matrices H_κ, see (4.10), whose eigenvalues are estimates of eigenvalues of A. We discuss the computation of these estimates in the end of this section. Throughout this section we assume that $N = \infty$ in the formulas of Sect. 4, and, in particular, that the measure γ has infinitely many points of support, that the moments μ_k given by (4.5) are defined and finite for all $k \geq 0$, and that there is a complete family of monic formal orthogonal polynomials $\{\pi_k\}_{k=0}^{\infty}$ associated with γ.

Let $\{\tau_n\}_{n=0}^{\infty}$ be a family of polynomials that satisfy a three-term recurrence relation. We are especially interested in the case when the τ_n are the residual polynomials

p_n, given by (2.3), for Chebyshev iteration. It is easy to show that the p_n satisfy the three-term recurrence relation

$$\frac{1}{\alpha_n}\, p_{n+1}(\lambda) = \left(\frac{\beta_n}{\alpha_n} + \frac{1}{\alpha_n} - \lambda\right) p_n(\lambda) - \frac{\beta_n}{\alpha_n}\, p_{n-1}(\lambda)\,, \quad n \geq 1\,, \tag{5.1}$$
$$p_1(\lambda) = 1 - \frac{\lambda}{d}\,, \quad p_0(\lambda) = 1\,,$$

where the coefficients α_n and β_n are defined by (2.2). Introduce the quantities

$$\sigma_{mn} := \phi(\tau_m \pi_n)\,.$$

It follows from (4.6) that, if $\tau_k = p_k$, then

$$\sigma_{m0} = \nu_m\,, \quad m = 0, 1, 2, \ldots\,, \tag{5.2}$$

and the orthogonality of the π_n yields $\phi(\tau_m \pi_n) = 0$ for $m < n$, thus

$$\sigma_{mn} = 0 \quad \text{for} \quad m < n\,. \tag{5.3}$$

In order to derive matrix relations that will be used in the calculation of the recurrence coefficients $\hat{\alpha}_k$ and $\hat{\beta}_k$ for the polynomials π_k, we introduce the following semi-infinite vectors

$$\boldsymbol{\pi} = [\pi_0, \pi_1, \ldots]^{\mathrm{T}}\,, \quad \boldsymbol{\tau} = [\tau_0, \tau_1, \ldots]^{\mathrm{T}}\,,$$

and semi-infinite matrices

$$H := \begin{bmatrix} \hat{\alpha}_0 & \hat{\beta}_1 & & & 0 \\ 1 & \hat{\alpha}_1 & \hat{\beta}_2 & & \\ & 1 & \hat{\alpha}_2 & \hat{\beta}_3 & \\ 0 & & \ddots & \ddots & \ddots \end{bmatrix}\,, \quad T := \begin{bmatrix} \tau_{00} & \tau_{01} & & 0 \\ \tau_{10} & \tau_{11} & \tau_{12} & \\ & \tau_{21} & \tau_{22} & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix}\,,$$

$$S := \begin{bmatrix} \sigma_{00} & & & 0 \\ \sigma_{10} & \sigma_{11} & & \\ \sigma_{20} & \sigma_{21} & \sigma_{22} & \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}\,,$$

where H and T are tridiagonal and S is lower triangular. The matrix H_κ defined by (4.10) is the leading $\kappa \times \kappa$ principal submatrix of H. The nonvanishing entries of H, i.e., the recurrence coefficients for the polynomials π_k given by (4.9), are to be computed. The nontrivial entries τ_{jk} of the matrix T are recurrence coefficients of the polynomials τ_j and are assumed to be explicitly known. In particular, if $\tau_j = p_j$, then we obtain from (5.1) that

$$\tau_{00} = d\,, \quad \tau_{10} = -d\,,$$
$$\tau_{n-1,n} = -\frac{\beta_n}{\alpha_n}\,, \quad \tau_{nn} = \frac{\beta_n + 1}{\alpha_n}\,, \quad \tau_{n+1,n} = -\frac{1}{\alpha_n}\,, \quad n \geq 1\,.$$

We write the three-term recurrence relations for the π_k and τ_k in the form

$$z\boldsymbol{\pi}^{\mathrm{T}}(z) = \boldsymbol{\pi}(z)^{\mathrm{T}} H\,, \quad z\boldsymbol{\tau}^{\mathrm{T}}(z) = \boldsymbol{\tau}^{\mathrm{T}}(z) T\,. \tag{5.4}$$

Define the functional $\hat{\phi}$ on the set of vectors of polynomials by

$$\hat{\phi}\left([q_0, q_1, \ldots]^{\mathrm{T}}\right) := [\phi(q_0), \phi(q_1), \ldots]^{\mathrm{T}} \ , \quad q_n \in \Pi_n \ .$$

Applying $\hat{\phi}$ to the rank-one matrix $\boldsymbol{\tau}\boldsymbol{\pi}^{\mathrm{T}}$ yields $\hat{\phi}(\boldsymbol{\tau}\boldsymbol{\pi}^{\mathrm{T}}) = S$, and it follows from (5.4) that

$$\begin{aligned} SH &= \hat{\phi}(\boldsymbol{\tau}\boldsymbol{\pi}^{\mathrm{T}})H = \hat{\phi}(\boldsymbol{\tau}\boldsymbol{\pi}^{\mathrm{T}}H) = \hat{\phi}(\boldsymbol{\tau}z\boldsymbol{\pi}^{\mathrm{T}}) \\ &= \hat{\phi}\left((z\boldsymbol{\tau}^{\mathrm{T}})^{\mathrm{T}}\boldsymbol{\pi}^{\mathrm{T}}\right) = \hat{\phi}(T^{\mathrm{T}}\boldsymbol{\tau}\boldsymbol{\pi}^{\mathrm{T}}) = T^{\mathrm{T}}\hat{\phi}(\boldsymbol{\tau}\boldsymbol{\pi}^{\mathrm{T}}) = T^{\mathrm{T}}S \ . \end{aligned} \tag{5.5}$$

This matrix identity is the basis of the *modified Chebyshev algorithm*, described in [10, 11, 20, 23], for computing the recurrence coefficients for the polynomials π_k from the recurrence coefficients for the τ_k and the modified moments ν_k. We now derive the modified Chebyshev algorithm for computing the entries of H_κ. Equating elements in the left and right hand sides of equation (5.5) yields

$$\sigma_{i,j+1} + \hat{\alpha}_j\sigma_{ij} + \hat{\beta}_j\sigma_{i,j-1} = \tau_{i-1,i}\sigma_{i-1,j} + \tau_{ii}\sigma_{ij} + \tau_{i+1,i}\sigma_{i+1,j} \ . \tag{5.6}$$

If $i < j-1$, then both the right hand side and left hand side of (5.6) vanish, because S is lower triangular. When $i = j-1$, formula (5.6) yields

$$\hat{\beta}_j\sigma_{j-1,j-1} = \tau_{j,j-1}\sigma_{jj} \ , \tag{5.7}$$

and for $i = j$ we obtain

$$\hat{\alpha}_j\sigma_{jj} + \hat{\beta}_j\sigma_{j,j-1} = \tau_{jj}\sigma_{jj} + \tau_{j+1,j}\sigma_{j+1,j} \ . \tag{5.8}$$

The coefficients $\hat{\alpha}_j$ and $\hat{\beta}_j$ are computed by (5.7) and (5.8), and we note that this requires only the diagonal and subdiagonal elements of the matrix S. These entries of S can be generated recursively starting from $\{\sigma_{j0}\}_{j=0}^{2\kappa-1}$ defined by (5.2). The computations proceed as follows. Initialize

$$\begin{aligned} &\hat{\alpha}_0 := \tau_{00} + \tau_{10}\,\frac{\nu_1}{\nu_0} \ , \\ &\sigma_{j0} := \nu_j \ , \qquad 0 \le j \le 2\kappa - 1 \ , \\ &\sigma_{j1} := \tau_{j+1,j}\sigma_{j+1,0} + (\tau_{jj} - \hat{\alpha}_0)\sigma_{j0} + \tau_{j-1,j}\sigma_{j-1,0} \ , \quad 1 \le j \le 2\kappa - 2 \ . \end{aligned}$$

Then compute for $j = 1, 2, \ldots, \kappa - 1$:

$$\begin{aligned} &\hat{\beta}_j := \tau_{j,j-1}\frac{\sigma_{jj}}{\sigma_{j-1,j-1}} \ , \\ &\hat{\alpha}_j := \tau_{jj} + \tau_{j+1,j}\frac{\sigma_{j+1,j}}{\sigma_{jj}} - \hat{\beta}_j\frac{\sigma_{j,j-1}}{\sigma_{jj}} \ , \\ &\sigma_{i,j+1} := \tau_{i+1,i}\sigma_{i+1,j} + (\tau_{ii} - \hat{\alpha}_j)\sigma_{ij} + \tau_{i-1,i}\sigma_{i-1,j} - \hat{\beta}_j\sigma_{i,j-1} \ , \quad j \le i < 2\kappa - j \ , \end{aligned}$$

where we use the property (5.3). Thus, the computation of the entries of H_κ requires 2κ modified moments $\{\nu_j\}_{j=0}^{2\kappa-1}$. We compute these ν_j from the residual vectors $\{\mathbf{r}_j\}_{j=0}^{2\kappa-1}$ by (4.7). The eigenvalues of H_κ are estimates of eigenvalues of A, and in the numerical examples of Sect. 7 we computed them by the EISPACK [21] subroutine HQR. Our computational experience indicates that for reason of numerical stability κ should not be chosen too large, e.g., $\kappa \le 15$. A discussion on the stability of the modified Chebyshev algorithm can be found in [10]. Another restriction on the size of κ stems from that the matrices H_κ have to exist. We will discuss the existence

of the matrices H_κ in Sect. 6. In the present section we assume that the matrices H_κ exist for all $\kappa \geq 1$.

After each $2\kappa - 1$ iterations by the Chebyshev method (2.1), we can determine a new matrix H_κ and compute its spectrum. Let $\mathscr{S}(H)$ denote the union of sets of eigenvalues of all the computed matrices H_κ. We determine the parameters d and c by fitting an ellipse to the available set $S(H)$ using the scheme outlined in Sect. 3 with $\mathscr{S}(A)$ replaced by $\mathscr{S}(H)$. The spectrum of each computed matrix H_κ increases the set $\mathscr{S}(H)$. If during the iteration with the adaptive Chebyshev method one finds that the parameters d and c change insignificantly when eigenvalues of new matrices H_κ are included in the set $\mathscr{S}(H)$, then Chebyshev iteration can typically proceed until convergence with fixed values of d and c and without computing further modified moments.

6. Relation to the nonsymmetric Lanczos process

In this section we show that the tridiagonal matrix H_κ computed from the modified moments as described in Sect. 5 is similar to a complex symmetric tridiagonal matrix that can be generated by the nonsymmetric Lanczos process. Let the π_k be the monic orthogonal polynomials defined by (4.9) and let the eigenvalues of A be denoted by λ_j. In view of that π_N is uniquely determined by the recurrence relation (4.9), and since the polynomial

$$p(z) := \prod_{j=1}^{N} (z - \lambda_j)$$

is monic and satisfies $\langle\langle p, \pi_j \rangle\rangle = 0$ for $0 \leq j < N$, it follows that $\pi_N = p$. Therefore

$$\pi_N(A) = 0 \ . \tag{6.1}$$

It is convenient to write (4.9) in matrix form. Introduce the matrix

$$P := \begin{bmatrix} \pi_0(\lambda_1) & \pi_0(\lambda_2) & \dots & \pi_0(\lambda_N) \\ \pi_1(\lambda_1) & \pi_1(\lambda_2) & \dots & \pi_1(\lambda_N) \\ \vdots & \vdots & & \vdots \\ \pi_{N-1}(\lambda_1) & \pi_{N-1}(\lambda_2) & \dots & \pi_{N-1}(\lambda_N) \end{bmatrix} ,$$

and the row vector

$$\boldsymbol{\pi}_N(z) := [\pi_0(z), \pi_1(z), \dots, \pi_{N-1}(z)] \ .$$

Then (4.9) can be written as

$$\boldsymbol{\pi}_N(z) H_N = z\boldsymbol{\pi}_N(z) - [0, \dots, 0, \pi_N(z)] \ . \tag{6.2}$$

Setting $z = \lambda_k$ in (6.2) for $1 \leq k \leq N$ yields

$$P^{\mathrm{T}} H_N = \Lambda P^{\mathrm{T}} \ , \tag{6.3}$$

where Λ is given by (1.3). Introduce the matrix

$$\Gamma := \operatorname{diag}\left[\tilde{\gamma}_1^{1/2}, \tilde{\gamma}_2^{1/2}, \dots, \tilde{\gamma}_N^{1/2}\right] \ ,$$

where the $\tilde{\gamma}_i$ are given by (4.3). It follows from the orthogonality of the π_k with respect to the bilinear form (4.8) that, for certain constants $d_k \neq 0$,

$$(P\Gamma)(P\Gamma)^{\mathrm{T}} = D = \mathrm{diag}[d_2, d_2, \ldots, d_N] \ ,$$

and, therefore,

$$(D^{-1/2}P\Gamma)(D^{-1/2}P\Gamma)^{\mathrm{T}} = I \ . \tag{6.4}$$

From (6.3) and (6.4) we obtain the identity

$$D^{-1/2}H_N D^{1/2} = (D^{-1/2}P\Gamma)\Lambda(D^{-1/2}P\Gamma)^{\mathrm{T}} \ , \tag{6.5}$$

which shows that the matrix

$$\tilde{H}_N := \begin{bmatrix} \tilde{\alpha}_0 & \tilde{\beta}_1 & & & 0 \\ \tilde{\beta}_1 & \tilde{\alpha}_1 & \tilde{\beta}_2 & & \\ & \tilde{\beta}_2 & \ddots & \ddots & \\ & & \ddots & & \tilde{\beta}_{N-1} \\ 0 & & & \tilde{\beta}_{N-1} & \tilde{\alpha}_{N-1} \end{bmatrix} := D^{-1/2}H_N D^{1/2} \tag{6.6}$$

is complex symmetric. For future reference, we define $\tilde{H}_\kappa$ as the leading $\kappa \times \kappa$ principal submatrix of $\tilde{H}_N$. It follows from (6.4) that the polynomials

$$\tilde{\pi}_k := d_k^{-1/2}/\pi_k \ , \quad 0 \leq k < N \ ,$$

are orthonormal with respect to the inner product (4.8), and the entries of the matrix (6.6) are the recurrence coefficients for the $\tilde{\pi}_k$.

Introduce the $N \times l$ matrices

$$V_l := \left[\tilde{\pi}_0(A)\mathbf{r}_0, \tilde{\pi}_1(A)\mathbf{r}_0, \ldots, \tilde{\pi}_{l-1}(A)\mathbf{r}_0\right] \ ,$$
$$U_l := \left[\tilde{\pi}_0(A^{\mathrm{T}})\mathbf{r}_0, \tilde{\pi}_1(A^{\mathrm{T}})\mathbf{r}_0, \ldots, \tilde{\pi}_{l-1}(A^{\mathrm{T}})\mathbf{r}_0\right] \ .$$

The columns of V_l form a basis of the Krylov subspace $K_l(A, \mathbf{r}_0)$, cf. (2.10), and the columns of U_l form a basis of $K_l(A^{\mathrm{T}}, \mathbf{r}_0)$. It follows from (4.9) and (6.1) that

$$\begin{aligned} AV_l &= V_l\tilde{H}_l + \mathbf{e}_l^{\mathrm{T}}\tilde{\pi}_l(A)\mathbf{r}_0 \ , \\ A^{\mathrm{T}}U_l &= U_l\tilde{H}_l + \mathbf{e}_l^{\mathrm{T}}\tilde{\pi}_l(A^{\mathrm{T}})\mathbf{r}_0 \ , \end{aligned} \qquad 1 \leq l \leq N \ , \tag{6.7}$$

where we define $\tilde{\pi}_N := 0$. In view of (1.3) and

$$\alpha_l\mathbf{r}_0^{\mathrm{T}}\mathbf{w}_l = \alpha_l \sum_{j=1}^{N} \alpha_j \mathbf{w}_j^{\mathrm{T}}\mathbf{w}_l = \tilde{\gamma}_l \ , \quad 1 \leq l \leq N \ ,$$

we have that

$$\begin{aligned} \mathbf{r}_0^{\mathrm{T}}\tilde{\pi}_j(A)\tilde{\pi}_k(A)\mathbf{r}_0 &= \mathbf{r}_0^{\mathrm{T}}W\tilde{\pi}_j(\Lambda)\tilde{\pi}_k(\Lambda)W^{-1}\mathbf{r}_0 \\ &= \sum_{l=1}^{N} \tilde{\pi}_j(\lambda_l)\tilde{\pi}_k(\lambda_l)\alpha_l\mathbf{r}_0^{\mathrm{T}}\mathbf{w}_l \\ &= \sum_{l=1}^{N} \tilde{\pi}_j(\lambda_l)\tilde{\pi}_k(\lambda_l)\tilde{\gamma}_l \ . \end{aligned} \tag{6.8}$$

Equation (6.8) shows that

$$U_l^{\mathrm{T}} V_l = I \ , \quad 1 \le l \le N \ . \tag{6.9}$$

Thus, the columns of U_N and V_N are biorthogonal.

Now let $V_n = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N]$ and $U_N = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N]$. Then (6.7) can be written as

$$\begin{aligned} A\mathbf{v}_{j+1} &= \tilde{\beta}_j \mathbf{v}_j + \tilde{\alpha}_j \mathbf{v}_{j+1} + \tilde{\beta}_{j+1} \mathbf{v}_{j+2} \ , \\ A^{\mathrm{T}} \mathbf{u}_{j+1} &= \tilde{\beta}_j \mathbf{u}_j + \tilde{\alpha}_j \mathbf{u}_{j+1} + \tilde{\beta}_{j+1} \mathbf{u}_{j+2} \ , \end{aligned} \quad 1 \le j \le N-2 \ , \tag{6.10}$$

where $\mathbf{u}_0 := \mathbf{v}_0 := 0$. The columns of V_N and U_N, and the entries of $\tilde{H}_N$, can be generated recursively by formulas (6.10) and (6.9). This is the nonsymmetric Lanczos process. Thus, the matrix H_κ determined by the modified Chebyshev algorithm as described in Sect. 5, is similar to a complex symmetric matrix $\tilde{H}_\kappa$ that can be computed by the Lanczos process with initial vectors $\mathbf{u}_1 := \mathbf{v}_1 := s\mathbf{r}_0$ for some suitable scaling factor s.

We are now in a position to discuss the existence of the matrices H_κ obtained by the modified Chebyshev algorithm. The above connection with the nonsymmetric Lanczos process shows that a matrix H_κ fails to exist precisely when the nonsymmetric Lanczos process breaks down. Gutknecht [14, 15] has presented a look-ahead procedure that can be used in conjunction with formula (6.10) to avoid break-down of the Lanczos process. Golub and Gutknecht [11] describe how break-downs can be circumvented in the modified Chebyshev algorithm. However, a modified Chebyshev algorithm that does not break down is more complicated than the algorithm described in Sect. 5. In our implementation we have therefore used the modified Chebyshev algorithm as described in Sect. 5, and we choose κ small enough in an adaptive manner during the iterations, so that the computed matrices H_κ exist. As already pointed out above, κ also has to be chosen small enough to avoid numerical instability.

7. Numerical examples

This section presents numerical experiments, in which the performance of our new adaptive Chebyshev algorithm for nonsymmetric linear systems based on modified moments is compared with the adaptive Chebyshev method based on the power method applied to the matrix $\hat{A}$ by Manteuffel [18] as implemented by CHEBYCODE [2], where $\hat{A}$ is defined by (2.9). In our experiments we found this power method to generally work better than the power method applied to the matrix A. The hybrid algorithm by Elman et al. sometimes converges faster than adaptive Chebyshev iteration; see [7]. However, the lack of software for the hybrid scheme makes a comparison difficult and is therefore omitted in the present paper.

All programs used are written in FORTRAN 77. Our new adaptive scheme has been implemented by using parts of CHEBYCODE, e.g., the subroutines for computing the Chebyshev iterates and for determining and updating the smallest ellipse containing all computed eigenvalue estimates of the matrix A. We carried out the numerical experiments on an IBM RISC 6000/550 workstation using double precision arithmetic, i.e., with approximately 15 significant digits.

The test problems are derived by discretizing the elliptic partial differential equation

$$-\Delta u + 2p_1 u_x + 2p_2 u_y - p_3 u = f \tag{7.1}$$

with constant coefficients p_1, p_2 and p_3 on the unit square $\Omega : 0\{(x, y,) : 0 \leq x, y \leq 1\}$, and with boundary condition $u(x, y) = 0$ on $\partial\Omega$. The function f is chosen so that $u(x, y) = xe^{xy} \sin(\pi x) \sin(\pi y)$ solves (7.1). We discretize (7.1) by symmetric finite differences on a uniform $(n+2) \times (n+2)$ grid, including boundary points, and use the standard five-point stencil to approximate the Laplacian Δ. This yields a linear system of $N := n^2$ equations for n^2 unknowns u_{ij}, $1 \leq i, j \leq n$, where u_{ij} approximates the solution u of (7.1) at the grid point $(ih, jh), h := \frac{1}{n+1}$. We scale the linear system obtained in this manner by h^2 and write it as $\tilde{A}\mathbf{x} = \mathbf{b}$. A typical equation of this system reads

$$(4-p_3h^2)u_{ij} - (1+p_1h)u_{i-1,j} - (1-p_1h)u_{i+1,j} - (1+p_2h)u_{i,j-1} - (1-p_2h)u_{i,j+1} = h^2 f_{ij} \ ,$$

where $f_{ij} = f(ih, jh)$. In order to keep the issues of interest clear, no preconditioner is used. In practical applications, however, the use of a preconditioner is often desirable. To obtain systems of equations with different properties, we modify the matrix $\tilde{A}$ by adding a multiple of the identity, i.e., we solve $A\mathbf{x} = \mathbf{b}$, where $A := \tilde{A} + \delta I$ and $\delta \geq 0$ is a constant. As the value of δ increases, the spectrum of the matrix A is shifted away from the origin.

In the following tables "pm($\hat{A}$)" denotes the adaptive Chebyshev algorithm based on the power method applied to the matrix $\hat{A}$ as implemented by the code [2]. We denote our adaptive scheme based on modified moments associated with the matrix A by "mm(A)". The number of iterations required by the methods is listed in the column labeled "steps". The number of eigenvalue estimates computed in each adaptive step is shown in the column labeled "eigenvalues". The column in the tables labeled "maxadapt" displays the maximum number of times the ellipse is fitted in the scheme mm(A). The column labeled "frequency" shows how often the ellipse is fitted. For instance, if maxadapt = 10 and frequency = 20, then the ellipse is fitted after every 20 iterations until the ellipse has been fitted 10 times. In the examples the number of eigenvalues κ estimated by the adaptive procedure varies from 4 to 14. We observed that, in general, a larger number of eigenvalues can be estimated for problems of smaller size. When the dimension of the matrix A becomes large, e.g., $N = 10,000$ or $N = 25,600$, we found that the number of eigenvalue estimates computed in each adaptive step has to be reduced in order to obtain useful estimates. This depends on that round-off errors introduced during the computation of the modified moments can give rise to significant errors in the modified moments when the vector length is large. Furthermore, we noticed that when the norm of the residual vector grows rapidly because a poor ellipse is used, then the ellipse should be refitted frequently. This is the case during the initial iterations in Example 7.2 with a pronouncedly nonsymmetric matrix. After each fitting of the ellipse only 2κ modified moments (4.7) have to be computed, and therefore only 2κ inner products are computed, independently of the frequency $\geq 2\kappa$ chosen. Moreover, after the ellipse has been fitted maxadapt times no more modified moments, and therefore no more inner products, are computed. We remark that our code for the adaptive Chebyshev algorithm based on modified moments is a research code and lacks the sophistication of a production code. We believe that its performance can be improved by careful coding and by implementation of strategies for choosing the frequency and maxadapt parameters.

In the derivation of the modified Chebyshev algorithm of Sect. 5 we assumed that the tridiagonal matrix H exists. If during the computations it would turn out that the

$\kappa \times \kappa$ leading principal submatrix H_κ of H does not exist, then the modified Chebyshev algorithm is curtailed and a submatrix of H_κ is determined, whose spectrum yields eigenvalue estimates of A. In all examples new iterates are determined until the norm of the residual vector has decreased by at least a factor $1 \cdot 10^{-8}$. Numerous computed examples show that the total number of iterations required is quite insensitive to the choice of initial ellipse. In the examples reported below, we chose the initial values of the center d and square of the focal length c of the initial ellipse to be $d = 4 - p_3 h^2$ and $c^2 = 0.1$.

Example 7.1. We select $p_1 = 40$, $p_2 = 40$, $p_3 = 30$ and the dimension of A, $N = 6,400$. Table 1 shows that the adaptive schemes mm(A) and pm($\hat{A}$) achieve convergence in approximately the same number of Chebyshev iteration steps when $\delta = 0.03$. In this example only two adaptive steps are taken by the schemes mm(A), with 14 eigenvalue estimates computed in each adaptive step. Table 2 differs from Table 1 only in the choice of δ. In Table 2, $\delta = 0.01$ and the spectrum of the matrix A is closer to the origin than in Table 1. On comparing Tables 1 and 2, we see that the number of iterations required for pm($\hat{A}$) to achieve convergence is nearly twice as large in Table 2 than in Table 1, while the entries for mm(A) are roughly the same in both tables. Thus, for the present example the adaptive scheme pm($\hat{A}$) is much more sensitive to the choice of δ, i.e., to the distance of the spectrum to the origin, than our schemes based on modified moments. We remark that the foci of all ellipses determined by the three adaptive schemes lie on the real axis.

Table 1. $p_1 = 40$, $p_2 = 40$, $p_3 = 30$; $\delta = 0.03$; $N = 6,400$

Adaptive method	Eigenvalues	Maxadapt	Frequency	Steps	$\|\mathbf{r}_{\text{last}}\|/\|\mathbf{r}_0\|$
pm($\hat{A}$)				273	.57D-10
mm(A)	14	2	28	266	.88D-10

Table 2. $p_1 = 40$, $p_2 = 40$, $p_3 = 30$; $\delta = 0.01$; $N = 6,400$

Adaptive method	Eigenvalues	Maxadapt	Frequency	Steps	$\|\mathbf{r}_{\text{last}}\|/\|\mathbf{r}_0\|$
pm($\hat{A}$)				580	.71D-10
mm(A)	14	2	28	276	.65D-10

Example 7.2. We select $p_1 = 90$, $p_2 = 90$, $p_3 = 200$, $\delta = 0.01$ and $N = 2,500$. The matrix A in this example is highly nonsymmetric and many of its eigenvalues have large imaginary part. The large growth of the norm of the residual vectors in the first few iterations limits the number of eigenvalue estimates which can be computed in each adaptive step to 6. Furthermore, in order to obtain reasonable estimates of the ellipse containing the spectrum of A in its interior, the adaptive procedure is applied 6 times. We point out that the total number of eigenvalues of A estimated in this example is actually smaller than in Example 7.1. Table 3 shows that for this example pm($\hat{A}$) requires more iterations than mm(A) to achieve convergence.

Example 7.3. We select $p_1 = 2$, $p_2 = 4$, $p_3 = 30$, $\delta = 0.01$ and the dimension of A equal to 4,900. All eigenvalue estimates for this example are real for both adaptive

Table 3. $p_1 = 90$, $p_2 = 90$, $p_3 = 200$; $\delta = 0.01$; $N = 2,500$

Adaptive method	Eigenvalues	Maxadapt	Frequency	Steps	$\|\mathbf{r}_{\text{last}}\|/\|\mathbf{r}_0\|$
pm($\hat{A}$)				365	.31D-9
mm(A)	6	6	12	346	.18D-9

Table 4. $p_1 = 2$, $p_2 = 4$, $p_3 = 30$; $\delta = 0.01$; $N = 4,900$

Adaptive method	Eigenvalues	Maxadapt	Frequency	Steps	$\|\mathbf{r}_{\text{last}}\|/\|\mathbf{r}_0\|$
pm($\hat{A}$)				935	.17D-10
mm(A)	6	6	12	413	.92D-11

Table 5. $p_1 = 200$, $p_2 = 200$, $p_3 = 300$; $\delta = 0.2$; $N = 10,000$

Adaptive method	Eigenvalues	Maxadapt	Frequency	Steps	$\|\mathbf{r}_{\text{last}}\|/\|\mathbf{r}_0\|$
pm($\hat{A}$)				398	.51D-9
mm(A)	4	6	8	279	.93D-9

Table 6. $p_1 = 100$, $p_2 = 100$, $p_3 = 300$; $\delta = 0.2$; $N = 25,600$

Adaptive method	Eigenvalues	Maxadapt	Frequency	Steps	$\|\mathbf{r}_{\text{last}}\|/\|\mathbf{r}_0\|$
pm($\hat{A}$)				355	.29D-9
mm(A)	4	6	16	295	.28D-9

schemes used. The number of Chebyshev iteration steps required by each adaptive Chebyshev iteration scheme in order to achieve convergence is reported in Table 4. Notice that in this example the number of Chebyshev iteration steps required for pm($\hat{A}$) to achieve converges in more than twice the number of iteration steps needed by mm(A).

Example 7.4. In this example we choose $p_1 = 200$, $p_2 = 200$, $p_3 = 300$, $\delta = 0.2$ and $N = 10,000$. Table 5 shows that pm($\hat{A}$) requires 40% more iteration steps than mm(A) in order to achieve convergence.

Example 7.5. We select $p_1 = 100$, $p_2 = 100$, $p_3 = 300$, $\delta = 0.2$ and $N = 25,600$. In this example the axis of the final ellipse determined is parallel with the imaginary axis. Table 6 shows that mm(A) achieves convergence in approximately 20% fewer iterations than pm($\hat{A}$).

8. Conclusions

This paper presents an adaptive Chebyshev algorithm for solving large, sparse non-symmetric linear systems based on modified moments. A major advantage of this scheme is that it requires fewer N-vectors be stored in computer memory than adaptive schemes based on power methods or the Arnoldi process. Our numerical examples illustrate that our scheme often yields significantly faster convergence than the power method applied to $\hat{A}$, when the matrix A has eigenvalues close to the origin. When

the eigenvalues are not close to the origin, adaptive schemes based on the power method can yield as rapid convergence as our scheme. The choices of how often the ellipse is to be fitted is important for the performance of our new method. Computed examples indicate that the ellipse does not have to be fitted many times, and, therefore, the iterations can be carried out by evaluating only fairly few inner products. Our scheme therefore is attractive for implementation on parallel MIMD and SIMD computers. Such implementations should adaptively determine how frequently the ellipse ought to be fitted, and when an ellipse has been determined that can be used until convergence.

Acknowledgements. We would like to thank Steve Ashby for providing the code described in [2].

References

1. Arnoldi, W.E. (1951): The principle of minimized iterations in the solution of the matrix eigenvalue problem. Q. Appl. Math. **9**, 17–29
2. Ashby, S.F. (1985): CHEBYCODE: a FORTRAN implementation of Manteuffel's adaptive Chebyshev algorithm. Report UIUCDCS-R-85-1203, Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL
3. Bartle, R.G. (1964): Elements of Real Analysis. Wiley, New York
4. Chatelin, F., Godet-Thobie, S. (1991): Stability analysis in aeronautical industries, in: M. Durand, F. El Dabaghi, eds., High Performance Computing II. pp. 415–422. Elsevier Science Publishers
5. Dongarra, J.J., Duff, I.S., Sorensen, D.C., van der Vorst, H.A. (1991): Solving Linear Systems on Vector and Shared Memory Computers. SIAM, Philadelphia
6. Eiermann, M. (1989): On semiiterative methods generated by Faber polynomials. Numer. Math. **56**, 139–156
7. Elman, H.C., Saad, Y., Saylor, P.E. (1986): A hybrid Chebyshev Krylov subspace algorithm for solving nonsymmetric systems of linear equations. SIAM J. Sci. Stat. Comput. **7**, 840–855
8. Fischer, B., Freund, R. (1991): Chebyshev polynomials are not always optimal. J. Approx. Theory **65**, 261–272
9. Gautschi, W. (1968): Construction of Gauss-Christoffel quadrature formulas. Math. Comput. **22**, 251–270
10. Gautschi, W. (1982): On generating orthogonal polynomials. SIAM J. Sci. Stat. Comput. **3**, 289–317
11. Golub, G.H., Gutknecht, M.H. (1990): Modified moments for indefinite weight functions. Numer. Math. **57**, 607–624
12. Golub, G.H., Kent, M. (1989): Estimates of eigenvalues for iterative methods. Math. Comput. **53**, 619–626
13. Golub, G.H., Varga, R.S. (1961): Chebyshev semi-iterative methods, successive over-relaxation methods, and second order Richardson iterative methods I+II. Numer. Math. **3**, 147–168
14. Gutknecht, M.H. (1992): A completed theory of the unsymmetric Lanczos process and related algorithms. Part I, SIAM J. Matrix Anal. Appl. **13**, 594–639
15. Gutknecht, M.H.: A completed theory of the unsymmetric Lanczos process and related algorithms. Part II, SIAM J. Matrix Anal. Appl. (to appear)
16. Ho, D. (1990): Tchebyshev acceleration technique for large scale nonsymmetric matrices. Numer. Math. **56**, 721–734
17. Manteuffel, T.A. (1977): The Chebyshev iteration for nonsymmetric linear systems. Numer. Math. **28**, 307–327
18. Manteuffel, T.A. (1978): Adaptive procedure for estimation of parameters for the nonsymmetric Chebyshev iteration. Numer. Math. **31**, 187–208
19. Nachtigal, N.M., Reichel, L., Trefethen, L.N. (1992): A hybrid GMRES algorithm for nonsymmetric linear systems. SIAM J. Matrix Anal. Appl. **13**, 796–825
20. Sack, R.A., Donovan, A.F. (1972): An algorithm for Gaussian quadrature given modified moments. Numer. Math. **18**, 465–478

21. Smith, B.T., Boyle, J.M., Ikebe, Y., Klema, V.C., Moler, C.B. (1970): Matrix Eigensystem Routines: EISPACK Guide, 2nd ed., Springer, Berlin Heidelberg New York
22. Varga, R.S. (1962): Matrix Iterative Analysis. Prentice-Hall, Englewood Cliffs, NJ
23. Wheeler, J.C. (1974): Modified moments and Gaussian quadratures. Rocky Mt. J. Math. **4**, 287–295