A Polynomial Acceleration of the Projection Method for Large Nonsymmetric Eigenvalue Problems

A. Nishida and Y. Oyanagi

Department of Information Science
University of Tokyo
Tokyo, Japan 113

Abstract

This study proposes a method for the acceleration of the projection method to compute a few eigenvalues with the largest real parts of a large nonsymmetric matrix.

In the field of the solution of the linear system, an acceleration using the least squares polynomial which minimizes its norm on the boundary of the convex hull formed with the unwanted eigenvalues are proposed. We simplify this method for the eigenvalue problem using the similar property of the orthogonal polynomial. This study applies the Tchebychev polynomial to the iterative Arnoldi method and proves that it computes necessary eigenvalues with far less complexity than the QR method. Its high accuracy enables us to compute the close eigenvalues that can not be obtained by the simple Arnoldi method.

1 Introduction

In the fluid dynamics and the structural analysis, there are a number of cases where a few eigenvalues with the largest real parts of a nonsymmetric matrix are required. In economic modeling, the stability of a model is interpreted in terms of the dominant eigenvalues of a large nonsymmetric matrix [1, 7].

Several methods have been proposed for this problem. The method proposed by Arnoldi in 1951 and the subspace iteration method due to Rutishauser , which are variants of the projection method, have been the most effective for this purpose. The Arnoldi method, however, has a drawback of the expense of too much memory space. This problem can be solved by using the method iteratively [6]. Although the iterative Arnoldi method is quite effective and may excel the subspace iteration method in performance, the dimension of the subspace is inevitably large, in particular when the wanted eigenvalues are clustered. Moreover it favors the convergence on the envelope of the spectrum.

To overcome this difficulty Saad proposed a Tchebychev acceleration technique for the Arnoldi method [7], which is an expansion of the similar technique for symmetric matrices. In the nonsymmetric case, we have to consider the distribution of the eigenvalues in the complex plane. The normalized Tchebychev function $P_n(\lambda) = T_n(\frac{d-\lambda}{c})/T_n(\frac{d}{c})$ has the property

$$\lim_{n\to\infty} P_n(\lambda) = \begin{cases} 0 & \lambda \text{ is inside of } \hat{F}(d,c) \\ \infty & \lambda \text{ is outside of } \hat{F}(d,c) \end{cases},$$

where d and $d \pm c$ are the center and the focal points of the ellipse $\hat{F}(d,c)$, which passes through the origin. Considering the optimal ellipse which encloses the unnecessary eigenvalues obtained adaptively by the previous step of the Arnoldi method and applying this polynomial to the matrix of the problem, we can make the new matrix whose necessary eigenvalues are made dominant [5]. We continue the subsequent Arnoldi iterations with the new matrix. This algorithm was refined and expanded by Ho [4] to the case where the reference eigenvalues do not have the largest or the smallest real parts. However, the adaptive methods based on the optimal ellipse have a defect of making the excessively large ellipse compared with the distribution of the unwanted eigenvalues.

In this paper, we use the convex hull proposed for the solution of the nonsymmetric linear system [8] instead of Manteuffel's optimal ellipse. The least squares polynomials minimize the L_2 norm defined on the boundary of the convex hull which encloses the unnecessary eigenvalues. From the maximum modulus principle, the absolute value of the polynomial is guaranteed to take on a maximum on the boundary of the convex hull. The polynomials can be generated without any numerical integration, using the orthogonality of the Tchebychev functions. In the eigenvalue problem, we can directly use the ortho-normal polynomial generated by the Tchebychev functions as the mini-max polynomial, since we have no need to normalize the polynomial at the origin.

The numerical experiments show that the method is effective for this purpose. The iteration of the Arnoldi method proposed by Saad is used in our algorithm and contributes to the economization of the memory space, which is consumed mainly by the coefficients of the polynomials.

2 Background

This section gives an outline of the methods referred to in this paper. The Arnoldi method, which is a variant of the projection method, plays the main role in our problem. The principle of the acceleration technique using the optimal ellipse [5] is explained briefly, since the properties used in this method are also important in our algorithm. We then describe the Tchebychev-Arnoldi method using the optimal ellipse and the least-squares based method, which were developed for solving the linear system by Saad [7, 8].

2.1 The Arnoldi method

If $u \neq 0$, let $K_l = \text{lin}(u, Au, \cdots, A^{l-1}u)$ be the Krylov subspace generated by u. Arnoldi's method computes an orthogonal basis $\{v_i\}_1^l$ of K_l in which the map is represented by an upper Hessenberg matrix i.e., an upper triangular matrix with sub-diagonal elements:

1. $v_1 = u / \parallel u \parallel_2, \quad h_{1,1} = (Av_1, v_1);$
2. for $j = 1, \cdots, l-1$, put

$$x_{j+1} = Av_j - \sum_{i=1}^{j} h_{ij} v_i, \quad h_{j+1,j} = \parallel x_{j+1} \parallel_2,$$

$$v_{j+1} = h_{j+1}^{-1} x_{j+1}, \quad h_{i,j+1} = (Av_{j+1}, v_i), \quad (i \leq j+1).$$

The algorithm terminates when $x_j = 0$, which is impossible if the minimal polynomial of A with respect to u is of degree $\geq l$. If this condition is satisfied, $H_l = (h_{ij})$ is an irreducible Hessenberg matrix.

In the iterative variant [7], we start with an initial vector u and fix a moderate value m, then compute the eigenvectors of H_m. We begin again, using as a starting vector a linear combination of the eigenvectors. No proof exists for the convergence of this method.

2.2 The adaptive Tchebychev-Arnoldi method

The original idea of using the Tchebychev polynomial for filtering the desired eigenvalues was proposed by Manteuffel in 1977 [5]. It was applied to the solution of nonsymmetric linear systems.

If x_0 is the initial guess at the solution x, an iteration is defined with the general step $x_n = x_{n-1} + \sum_{i=1}^{n-1} \gamma_{ni} r_i$ where the γ_{ij}'s are constants and $r_i = b - Ax_i$ is the residual at step i. Let $e_i = x - x_i$ be the error at the ith step then an inductive argument yields $e_n = [I - As_n(A)]e_0 \equiv P_n(A)e_0$ where $s_n(z)$ and $P_n(z)$ are polynomials of degree n such that $P_n(0) = 1$. To make $\parallel e_n \parallel \leq \parallel P_n(A) \parallel \parallel e_0 \parallel$ small, the Tchebychev polynomial is used as the sequence of polynomials.

The Tchebychev polynomials are given by $T_n(z) = \cosh(n\cosh^{-1}(z))$. Let $F(d,c)$ be the member of the family of ellipses in the complex plane centered at d with the focal points at $d+c$ and $d-c$, where d and c are complex numbers. Suppose $z_i \in F_i(0,1)$, $z_j \in F_j(0,1)$; then

$$\Re\mathrm{e}\,(\cosh^{-1}(z_i)) < \Re\mathrm{e}\,(\cosh^{-1}(z_j)) \Leftrightarrow F_i(0,1) \subset F_j(0,1),$$

$$\Re\mathrm{e}\,(\cosh^{-1}(z_i)) = \Re\mathrm{e}\,(\cosh^{-1}(z_j)) \Leftrightarrow F_i(0,1) = F_j(0,1).$$

Consider the scaled and translated Tchebychev polynomials $P_n(\lambda) = T_n(\frac{d-\lambda}{c})/T_n(\frac{d}{c})$. Using the definition of the cosh, we can see that

$$P_n(\lambda) = \frac{e^{n\cosh^{-1}(\frac{d-\lambda}{c})} + e^{-n\cosh^{-1}(\frac{d-\lambda}{c})}}{e^{n\cosh^{-1}(\frac{d}{c})} + e^{-n\cosh^{-1}(\frac{d}{c})}} \doteq e^{n\cosh^{-1}(\frac{d-\lambda}{c}) - n\cosh^{-1}(\frac{d}{c})}$$

for large n.

Let $r(\lambda) = \lim_{n\to\infty} |P_n(\lambda)^{\frac{1}{n}}|$, then we have $r(\lambda) = e^{\Re\mathrm{e}\,(\cosh^{-1}(\frac{d-\lambda}{c}) - \cosh^{-1}(\frac{d}{c}))}$. From the above lemma and the definition of $r(\lambda)$, we have that if $\lambda_i \in F_i(d,c), \lambda_j \in F_j(d,c)$, then

$$r(\lambda_i) < r(\lambda_j) \Leftrightarrow F_i(d,c) \subset F_j(d,c), \quad r(\lambda_i) = r(\lambda_j) \Leftrightarrow F_i(d,c) = F_j(d,c), \quad r(\lambda) = 1 \Leftrightarrow \lambda \in \hat{F}(d,c),$$

where $\hat{F}(d,c)$ is the member of the family passing through the origin. Thus we have

$$\lim_{n\to\infty} P_n(\lambda) = \begin{cases} 0 & \lambda \text{ is inside of } \hat{F}(d,c) \\ \infty & \lambda \text{ is outside of } \hat{F}(d,c) \end{cases}.$$

Suppose that we can find the optimal ellipse that contains all the eigenvalues of A except for the r wanted ones [5]. Then the algorithm runs a certain number of steps of the Tchebychev iteration and take the resulting vector z_n as the initial vector in the Arnoldi process. From the Arnoldi purification process one obtains a set of m eigenvalues, r of which are the approximation to the r wanted ones, while the remaining ones will be used for adaptively constructing the best ellipse.

- *Start:* Choose an initial vector v_1, a number of Arnoldi steps m and a number of Tchebychev steps n.
- *Iterate:*
  1. Perform the m steps of the Arnoldi algorithm starting with v_1. Compute the m eigenvalues of the resulting Hessenberg matrix. Select the r eigenvalues of the largest real parts $\tilde{\lambda}_1, \cdots, \tilde{\lambda}_r$ and take $\tilde{R} = \{\tilde{\lambda}_{r+1}, \cdots, \tilde{\lambda}_m\}$. If satisfied stop, otherwise continue.
  2. Using $\tilde{R}$, obtain the new estimates of the parameters d and c of the best ellipse. Then compute the initial vector z_0 for the Tchebychev iteration as a linear combination of the approximate eigenvectors $\tilde{u}_i, \quad i = 1, \cdots, r$.
  3. Perform n steps of the Tchebychev iteration to obtain z_n. Take $v_1 = z_n / \parallel z_n \parallel$ and back to 1.

2.3 The least-squares based method

It has been shown that the least-squares based method for solving linear systems is competitive with the ellipse based methods and are more reliable [8].

By the maximum principle, the maximum modulus of $|1 - \lambda s_n(\lambda)|$ is found on the boundary of some region H of the complex plane that includes the spectrum of A and it is sufficient to regard the problem as being defined on the boundary. We use the least squares residual polynomial minimizing the L_2 norm $\parallel 1 - \lambda s_n(\lambda) \parallel_w$ with respect to some weight $w(\lambda)$ on the boundary of H [8]. Suppose that the $\mu+1$ points $h_0, h_1, \cdots, h_\mu$ constitute the vertices of H. On each edge E_ν, $\nu = 1, \cdots, \mu$, of the convex hull, we choose a weight function $w_\nu(\lambda)$. Denoting by c_ν the center of the νth edge and by d_ν the half width, i.e., $c_\nu = (h_\nu + h_{\nu-1})/2$, $d_\nu = (h_\nu - h_{\nu-1})/2$, the weight function on each edge is defined by $w_\nu(\lambda) = 2|d_\nu^2 - (\lambda - c_\nu)^2|^{-\frac{1}{2}}/\pi$. The inner product on the space of complex polynomials is defined by $\langle p,q\rangle = \sum_{\nu=1}^{\mu}\int_{E_\nu} p(\lambda)\overline{q(\lambda)}w_\nu(\lambda)|d\lambda|$. An algorithm using explicitly the modified moments $\langle t_i(\lambda), t_j(\lambda)\rangle$, where $\{t_j\}$ is some suitable basis of polynomials, is developed for the problem of computing the least squares polynomials in the complex plane.

We express the polynomial $t_j(\lambda)$ in terms of the Tchebychev polynomials $t_j(\lambda) = \sum_{i=0}^{j}\gamma_{i,j}^{(\nu)}T_i(\xi)$ where $\xi = (\lambda - c_\nu)/d_\nu$ is real. The expansion coefficients $\gamma_{i,j}^{(\nu)}$ can be computed easily from the three term recurrence of the polynomials $\beta_{k+1}t_{k+1}(\lambda) = (\lambda-\alpha_k)t_k(\lambda) - \delta_k t_{k-1}(\lambda)$. The problem $\min_{s\in P_{n-1}} \| 1-\lambda s_n(\lambda)\|_w$ is to find $\eta = (\eta_0, \eta_1, \cdots, \eta_{n-1})^T$ of $s_n(\lambda) = \sum_{i=0}^{n-1}\eta_i t_i(\lambda)$ so that $J(\eta) = \| 1-\lambda s_n(\lambda)\|_w$ is minimum.

2.4 Approach

In the previous section we described the outline of the least-squares based method on any arbitrary area. It has a difficulty on the application to other purposes due to the constraint $P_n(0) = 1$.

We use the fact that the eigenvalue problem does not require any such condition to the polynomial and propose a new simple algorithm to get the mini-max polynomial to accelerate the convergence of the projection method. The minimum property of the Tchebychev functions described below is important to prove the optimality of this polynomial.

Let a non-negative weight function $w(\lambda)$ be given in the interval $a \geq \lambda \geq b$. The orthogonal polynomials $p_0(\lambda), p_1(\lambda), \cdots$, when multiplied by suitable factors C, possess a minimum property:

the integral $\int(\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_0)^2 w(\lambda)d\lambda$ takes on its least value when the polynomial in the integrand is $Cp_n(\lambda)$. The polynomial in the integrand may be written as a linear combination of the $p_i(\lambda)$, in the form $(Cp_n(\lambda) + c_{n-1}p_{n-1}(\lambda) + \cdots c_0)$. Since the functions $p_n(\lambda)\sqrt{w(\lambda)}$ are orthogonal, and in fact, orthogonal if the $p_i(\lambda)$ are appropriately defined, the integral is equal to $C^2 + \sum_{\nu=0}^{n-1}c_\nu^2$, which assumes its minimum at $c_0 = c_1 = \cdots = c_{n-1} = 0$.

Using the above property, we describe the new method to generate the coefficients of the ortho-normal polynomials in terms of the Tchebychev weight below.

We use the three term recurrence $\beta_{n+1}p_{n+1}(\lambda) = (\lambda - \alpha_n)p_n(\lambda) - \beta_n p_{n-1}(\lambda)$, where $p_i(\lambda)$ satisfies the orthonormality. Because of the condition of the use of the Tchebychev polynomial $p_n(\lambda) = \sum_{i=0}^{n}\gamma_{i,n}^{(\nu)}T_i[(\lambda - c_\nu)/d_\nu]$, the constraints $\langle p_0, p_0\rangle = 2\sum_{\nu=1}^{\mu}|\gamma_{0,0}^{(\nu)}|^2 = 1$, $\langle p_1, p_1\rangle = \sum_{\nu=1}^{\mu}[2|\gamma_{0,1}^{(\nu)}|^2 + |\gamma_{1,1}^{(\nu)}|^2] = 1$, and $\langle p_0, p_1\rangle = 2\sum_{\nu=1}^{\mu}\gamma_{0,0}^{(\nu)}\overline{\gamma}_{1,1}^{(\nu)} = 0$ must hold.

Moreover each expansion of $p_i(\lambda)$ at any edge must be consistent. The condition $2\sum_{\nu=1}^{\mu}|\gamma_{0,0}^{(\nu)}|^2 = 1$ derives $|\gamma_{0,0}^{(\nu)}| = 1/2\mu,\ \nu = 1, \cdots \mu$, and we can choose $1/\sqrt{2\mu}$ as $\gamma_{0,0}^{(\nu)}$. The consistency of $p_1(\lambda) = \gamma_{0,1}^{(\nu)} + \gamma_{1,1}^{(\nu)}(\lambda - c_\nu)/d_\nu = (\gamma_{1,1}^{(\nu)}/d_\nu)\lambda + \gamma_{0,1}^{(\nu)} - \gamma_{1,1}^{(\nu)}c_\nu/d_\nu$ derives $\gamma_{1,1}^{(\nu)}/d_\nu = \gamma_{1,1}^{(\nu')}/d_{\nu'}$, $\gamma_{0,1}^{(\nu)} - \gamma_{1,1}^{(\nu)}c_\nu/d_\nu = \gamma_{0,1}^{(\nu')} - \gamma_{1,1}^{(\nu')}c_{\nu'}/d_{\nu'}$ $(\nu \neq \nu')$. It can be rewritten as $\gamma_{1,1}^{(\nu)} = d_\nu t$, $\gamma_{0,1}^{(\nu)} - c_\nu t = \gamma_{0,1}^{(\nu')} - c_{\nu'}t$ where t is a real number.

Using the condition $\langle p_0, p_1\rangle = \sum_{\nu=1}^{\nu}\gamma_{0,0}^{(\mu)}\overline{\gamma_{0,1}^{(\nu)}} = \sum_{\nu=1}^{\mu}\overline{\gamma_{0,1}^{(\nu)}}/\sqrt{2\mu} = 0$, the sum $\sum_{\nu=1}^{\mu}(\gamma_{0,1}^{(\nu)} - c_\nu t) = -\sum_{\nu=1}^{\mu}c_\nu t = \mu(\gamma_{0,1}^{(\nu')} - c_{\nu'}t)$ $\ (1 \leq \nu' \leq \mu)$ derives the relation $\gamma_{0,1}^{(\nu)} = c_\nu t - (\sum_{\nu'=1}^{\mu}c_{\nu'})t/\mu$. Putting it into the equation $\langle p_1, p_1\rangle = \sum_{\nu=1}^{\mu}[2|\gamma_{0,1}^{(\nu)}|^2 + |\gamma_{1,1}^{(\nu)}|^2] = 1$, we have $2\sum_{\nu=1}^{\mu}|c_\nu - (\sum_{\nu'=1}^{\mu}c_{\nu'})/\mu|^2t^2 + \sum_{\nu=1}^{\mu}|d_\nu|^2t^2 = 1$. It derives $t = 1/\sqrt{S}$ where $S = \sum_{\nu=1}^{\mu}[2|c_\nu - (\sum_{\nu'=1}^{\mu}c_{\nu'})/\mu|^2 + |d_\nu|^2]$. From the above constraints, we can determine the values of all the coefficients of the polynomial using the values of d_ν, c_ν,and μ.

3 Algorithm

We describe the details of our algorithm below. Following the definition of some notations, a mini-max polynomial, which is derived from the definition of the new norm on the boundary, is combined with the Arnoldi method as an accelerator. We also mention the block version of the Arnoldi method [9] in order to increase parallelism and to detect the multiplicity of the computed eigenvalues.

3.1 The definition of the L_2 norm

This section defines the L_2 norm on the boundary of the convex hull and other notations. We described in the previous section the outline of the method by the least squares polynomials, which is based on the convex hull generated from the unwanted eigenvalues. In this section we begin with the computation of the convex hull.

Suppose a nonsymmetric matrix A is given. We must implement some adaptive scheme in order to estimate the approximate eigenvalues. The combination of the eigensolver and the accelerating technique is described subsequently.

Nonsymmetric matrices have complex eigenvalues which distribute symmetrically in terms of the real axis. Hence we can consider only the upper half of the complex plane. Using the sorted eigenvalues, we start from the rightmost point and make it the vertex h_0 of the convex hull H. We compute the gradient between h_i and the other eigenvalues with smaller real parts, and choose the point with the smallest gradient as h_{i+1}.

Since the eigenvalues obtained by the Arnoldi method is roughly ordered in terms of the absolute value, the adoption of the bubble sort is appropriate. This algorithm uses the property that only the points in the upper half plane are concerned. It requires $O(n^2)$ complexity in the worst case and $O(n)$ in the best case.

We then define the L_2 norm on the boundary of the convex hull H constituted by the points $h_0, \cdots, h_\mu$. On each edge $E_\nu \quad (\nu = 1, 2, \cdots \mu)$, we denote the center and the half width by $c_\nu = (h_\nu + h_{\nu-1})/2$ and $d_\nu = (h_\nu - h_{\nu-1})/2$, respectively, and define the weight function by $w_\nu(\lambda) = 2[d_\nu^2 - (\lambda - c_\nu)^2]^{-\frac{1}{2}}/\pi \ \ (\lambda \in E_\nu)$. Using the above definition, the inner product on ∂H is defined by

$$\langle p, q\rangle = \int_{\partial H} p(\lambda)\overline{q(\lambda)}w(\lambda)|d\lambda| = \sum_{\nu=1}^{\mu} \int_{E_\nu} p(\lambda)\overline{q(\lambda)}w_\nu(\lambda)|d\lambda| = \sum_{\nu=1}^{\mu} \langle p, q\rangle_\nu .$$

$\| p(\lambda) \|_w = \langle p, p\rangle^{\frac{1}{2}}$ satisfies the following theorem.

Theorem 3.1 *In an inner product space, the norm $\| u \|$ of an element u of the complex linear space has the following properties:*

1. $\| ku \| = |k| \, \| u \|$.
2. $\| u \| > 0$ *unless* $u = 0$; $\quad \| u \| = 0$ *implies* $u = 0$.
3. $\| u + v \| \le \| u \| + \| v \|$.

3.2 The computation of the coefficient γ

We have mentioned the expression of the polynomial i.e., $p_n(\lambda) = \sum_{i=0}^{n} \gamma_{i,n}^{(\nu)} T_i[\lambda - c_\nu/d_\nu]$. Using the three term recurrence of the Tchebychev polynomials, a similar recurrence $\beta_{k+1}p_{k+1}(\lambda) = (\lambda - \alpha_k)p_k(\lambda) - \delta_k p_{k-1}(\lambda)$ on the $p_i(\lambda)$ holds. Denoting ξ_ν by $\xi_\nu = (\lambda - c_\nu)/d_\nu$, the equation can be rewritten as

$$\beta_{k+1}p_{k+1}(\lambda) = (d_\nu \xi + c_\nu - \alpha_k) \sum_{i=0}^{k} \gamma_{i,k}^{(\nu)} T_i(\xi) - \delta_k \sum_{i=0}^{k-1} \gamma_{i,k-1}^{(\nu)} T_i(\xi).$$

From the relations $\xi T_i(\xi) = [T_{i+1}(\xi) + T_{i-1}(\xi)]/2, \ i > 0$ and $\xi T_0(\xi) = T_1(\xi)$, it is expressed by

$$\sum \gamma_i \xi T_i(\xi) = \frac{1}{2}\gamma_1 T_0(\xi) + (\gamma_0 + \frac{1}{2}\gamma_2)T_1(\xi) + \cdots + \frac{1}{2}(\gamma_{i-1} + \gamma_{i+1})T_i(\xi) + \cdots + \frac{1}{2}(\gamma_{n-1} + \gamma_{n+1})T_n(\xi) \quad (\gamma_{n+1} = 0)$$

and arranged into

$$\beta_{n+1}p_{n+1}(\lambda) = d_\nu[\frac{\gamma_{1,n}^{(\nu)}}{2}T_0(\xi) + (\gamma_{0,n}^{(\nu)} + \frac{\gamma_{2,n}^{(\nu)}}{2})T_1(\xi) + \cdots + \sum_{i=2}^{n} \frac{1}{2}(\gamma_{i-1,n}^{(\nu)} + \gamma_{i+1,n}^{(\nu)})T_i(\xi)]$$

$$+(c_\nu-\alpha_n)\sum_{i=0}^{n}\gamma_{i,n}^{(\nu)}T_i(\xi)-\delta_n\sum_{i=0}^{n-1}\gamma_{i,n-1}^{(\nu)}T_i(\xi)\qquad(T_{-1}=T_1).$$

Comparing the equation with $p_{n+1}(\lambda)=\sum_{i=0}^{n+1}\gamma_{i,n+1}^{(\nu)}T_i(\xi)$, we find the following relations

$$\beta_{n+1}\gamma_{0,n+1}^{(\nu)}=\frac{1}{2}d_\nu\gamma_{1,n}^{(\nu)}+(c_\nu-\alpha_n)\gamma_{0,n}^{(\nu)}-\delta_n\gamma_{0,n-1}^{(\nu)},$$

$$\beta_{n+1}\gamma_{1,n+1}^{(\nu)}=d_\nu(\gamma_{0,n}^{(\nu)}+\frac{1}{2}\gamma_{2,n}^{(\nu)})+(c_\nu-\alpha_n)\gamma_{1,n}^{(\nu)}-\delta_n\gamma_{1,n-1}^{(\nu)},$$

and

$$\beta_{n+1}\gamma_{i,n+1}^{(\nu)}=\frac{d_\nu}{2}[\gamma_{i+1,n}^{(\nu)}+\gamma_{i-1,n}^{(\nu)}]+(c_\nu-\alpha_n)\gamma_{i,n}^{(\nu)}-\delta_n\gamma_{i,n-1}^{(\nu)}$$

$$i=2,...,n+1\qquad(\gamma_{-1,n}^{(\nu)}=\gamma_{1,n}^{(\nu)},\quad\gamma_{i,n}^{(\nu)}=0\quad i>n).$$

The choice of the initial values $\gamma_{0,0}^{(\nu)}$, $\gamma_{0,1}^{(\nu)}$ and $\gamma_{1,1}^{(\nu)}$ is described in the previous section.

3.3 The computation of the coefficients in the three term recurrence

Using the relation $\beta_{k+1}p_{k+1}(\lambda)=(\lambda-\alpha_k)p_k(\lambda)-\delta_k p_{k-1}(\lambda)$ and the orthogonality of the Tchebychev polynomials, we derive

$$\beta_{k+1}=\langle p_{k+1},p_{k+1}\rangle^{1/2}=\sum_{\nu=1}^{\mu}\int_{E_\nu}p_{k+1}\overline{p_{k+1}}w_\nu(\lambda)|d\lambda|=\sum_{\nu=1}^{\mu}{\sum_{i=0}^{k+1}}'\gamma_{i,k+1}^{(\nu)}\overline{\gamma_{i,k+1}^{(\nu)}}$$

where we denote by ${\sum'}_{i=0}^{n}a_i=2a_0+\sum_{i=1}^{n}a_i$.

α and δ are computed similarly:

$$\alpha_k=\langle\lambda p_k,p_k\rangle=\sum_{\nu=1}^{\mu}(c_\nu{\sum_{i=0}^{k}}'\gamma_{i,k}^{(\nu)}\overline{\gamma_{i,k}^{(\nu)}}+d_\nu{\sum_{i=0}^{k}}'\gamma_{i,k}^{(\nu)}\overline{\gamma_{i+1,k}^{(\nu)}}),\quad\delta_k=\langle\lambda p_k,p_{k-1}\rangle=\sum_{\nu=1}^{\mu}d_\nu v_\nu$$

where $v_\nu=\gamma_{1,k}^{(\nu)}\overline{\gamma_{0,k-1}^{(\nu)}}+(\gamma_{0,k}^{(\nu)}+\frac{1}{2}\gamma_{2,k}^{(\nu)})\overline{\gamma_{1,k-1}^{(\nu)}}+\sum_{i=2}^{k-1}\frac{1}{2}(\gamma_{i-1,k}^{(\nu)}+\gamma_{i+1,k}^{(\nu)})\overline{\gamma_{i,k-1}^{(\nu)}}$.

3.4 The polynomial iteration

The polynomial obtained in the above procedure is applied to the matrix of the problem. We describe the algorithm combined with the Arnoldi method. The Arnoldi method is expressed as follows.

$$\hat{v}_{j+1}=Av_j-\sum_{i=1}^{j}h_{ij}v_i,\quad h_{ij}=(Av_j,v_i),\quad i=1,...,j,$$

$$h_{j+1,j}=\parallel\hat{v}_{j+1}\parallel,\qquad v_{j+1}=\hat{v}_{j+1}/h_{j+1,j}$$

where v_1 is an arbitrary non-zero initial vector. The eigenvectors corresponding to the eigenvalues which have the largest real parts are selected and combined linearly. The remaining eigenvalues constitute the convex hull. Suppose we have each coefficients of the polynomial $p_n(\lambda)$, where n is some appropriate integer. Put the combined vector into v_0 and we obtain the new vector v_n in which the components of the necessary eigenvectors are amplified by operating the following recursion:

$$p_0(A)v_0=\gamma_{0,0}^{(\nu)}Ev_0\qquad(1\le\nu\le\mu)$$

$$p_1(A)v_0=\gamma_{0,1}^{(\nu)}Ev_0+\gamma_{1,1}^{(\nu)}/d_\nu\cdot(A-c_\nu E)v_0$$

$$p_{i+1}(A)v_0 = [(A - \alpha_i E)p_i(A)v_0 - \delta_i p_{i-1}(A)v_0]/\beta_{i+1}.$$

Denoting $p_i(A)v_0$ by w_i, the above recurrence is transformed into

$$w_0 = \gamma_{0,0}^{(\nu)} v_0$$

$$w_1 = \gamma_{0,1}^{(\nu)} v_0 + \gamma_{1,1}^{(\nu)}/d_\nu \cdot (Av_0 - c_\nu v_0) = (\gamma_{0,1}^{(\nu)} - \gamma_{1,1}^{(\nu)} c_\nu / d_\nu) v_0 + \gamma_{1,1}^{(\nu)}/d_\nu \cdot Av_0$$

$$w_{i+1} = [Aw_i - \alpha_i w_i - \delta_i w_{i-1}]/\beta_{i+1} \qquad (i = 2, \cdots, n_T).$$

3.5 The block Arnoldi method

Suppose that we are interested in computing the r eigenvalues with largest real part. If $V_1 \in R^{n,r}$ is a rectangular matrix having r orthonormal columns and m is some fixed integer which limits the dimension of the computed basis, the iterative block-Arnoldi method can be described as follows:

for $k = 1, ..., m-1$ **do**

1. $W_k = AV_k$**;**

2. for $i = 1, ..., k$ **do**
$H_{i,k} = V_i^T W_k$; $\quad W_k = W_k - V_i H_{i,k}$**;**
end for

3. $Q_k R_k = W_k$

4. $V_{k+1} = Q_k$**;** $\quad H_{k+1,k} = R_k$**;**

end for

The restriction of the matrix A to the Krylov subspace has a block-Hessenberg form:

$$H_m = U_m^T A U_m = \begin{pmatrix} H_{1,1} & H_{1,2} & & \cdots & H_{1,m} \\ H_{2,1} & H_{2,2} & & & H_{2,m} \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & & & \\ 0 & \cdots & 0 & H_{m,m-1} & H_{m,m} \end{pmatrix}.$$

The above algorithm gives:

$$AV_k = \sum_{i=1}^{k} V_i H_{i,k} + V_{k+1} H_{k+1,k}, \quad k = 1, ..., m$$

which can be written in a form as $AU_m = U_m H_m + [0, ..., 0, V_{m+1} H_{m+1,m}]$, where $U_m = \{V_1, ..., V_m\}$. If $\Lambda_m = \mathrm{diag}(\lambda_1, ..., \lambda_m)$ denotes the diagonal matrix of eigenvalues of H_m corresponding to the eigenvectors $Y_m = [y_1, ..., y_{mr}]$ then the above relation gives $AU_m Y_m - U_m H_m Y_m = [0, ..., 0, V_{m+1} H_{m+1,m}] Y_m$. Denoting by $X_m = U_m Y_m$ the matrix of approximate eigenvectors of A and by $Y_{m,r}$ the last r block of Y_m, we have $\| AX_m - X_m \Lambda_m \|_2 = \| H_{m+1,m} Y_{m,r} \|_2$. It is used for the stopping criterion.

We compute the r eigenvalues of large real parts by the following algorithm:

Choose an orthogonal basis $V_1 = [v_1, ..., v_r]$, an integer m which limits the dimension of the computed basis, and an integer k which stands for the degree of the computed Tchebychev polynomials:

1. Starting with V_1, obtain the block-Hessenberg matrix H_m using the Arnoldi method. Compute the eigenvalues of H_m using the QR method and select $\lambda_1, ..., \lambda_r$ the eigenvalues of largest real part. Compute their Ritz vectors $x_1, ..., x_r$. Compute their residual norms for convergence test.

2. From $\mathrm{sp}(H_m) - \{\lambda_1, ..., \lambda_r\}$ get the optimal ellipse.

3. Starting with $x_1, ..., x_r$ and using the parameters of the ellipse just found, perform k steps of Tchebychev iteration to obtain an orthonormal set of vectors $v_1, ..., v_r$ and go to 1.

4 Evaluation

4.1 Complexity of the algorithms

The cost in terms of the number of floating-point operations are as follows: We denote by n, nz, m, r, k respectively the order of the matrix, its number of nonzero entries, the number of block Arnoldi steps, the number of required eigenvalues, and the degree of the Tchebychev polynomial. The block Arnoldi method costs $\sum_{j=1}^{m}\{2r\,nz + 4nr^2 j + 2r(r+1)n\} = 2rm\,nz + 2mr(mr+2r+1)n$ flops. $10r^3m^3$ flops are required for the computation of the eigenvalues of H_m of order mr by the QR method, $r^3O(m^2)$ for the corresponding eigenvectors by the inverse iteration, and $2kr\,nz + O(n)$ for the Tchebychev iteration [3, 9]. The computation of the coefficients costs approximately $O(\mu k^2)$ flops, where μ is the number of the vertices of the convex hull.

The total superfluous complexity of the least-squares based method by Saad is $O(k^3)$ flops, while Manteuffel's adaptive Tchebychev method requires the solutions of the equations of up to the fifth degree for $O(k^2)$ combinations of every two eigenvalues on the convex hull.

4.2 Numerical results

This section reports the results of the numerical experiments of our method and evaluates its performance. The experiments are performed on HP9000/720 using double precision. The time unit is 1/60 second.

We start with the decision of each element of the matrix given in the problem. In this section, the scaled sequences of random numbers are assigned respectively to the real and the imaginary parts of the eigenvalues except for those which are to be selected. The matrices are block diagonals with 2×2 or 1×1 diagonal blocks. Each block is of the form $\begin{bmatrix} a & b/2 \\ -2b & a \end{bmatrix}$ to prevent the matrix to be normal and has eigenvalues $a \pm bi$. It is transformed by an orthogonal matrix generated from a matrix with random elements by the Schmidt's orthogonalization method. m_A and n_T denote the order of the Arnoldi method and the maximum order of the Tchebychev polynomials respectively. We compare this algorithm with the double-shifted QR method. The error is computed by the L_2 norm.

In this section we test the some variations of the distribution of the eigenvalues using the matrices of order 50, the cases of $\lambda_{\max} = 2, 1.5$, and 1.1 while the distribution of the other eigenvalues is $\Re\mathrm{e}\,\lambda \in [0,1]$, and $\Im\mathrm{m}\,\lambda \in [-1,1]$. We denote the number of the iterations by i_A.

Case 1. $\lambda_{\max}$ is 2, while the distribution of the other eigenvalues is $\Re\mathrm{e}\,\lambda \in [0,1]$, $\Im\mathrm{m}\,\lambda \in [-1,1]$. The effect of the iteration is significant, especially for the orthogonality-based method. This tendency becomes sharper as the maximum eigenvalue gets closer to the second eigenvalue.

Case 2. The maximum eigenvalue is 1.5, while the distribution of the other eigenvalues is $\Re\mathrm{e}\,\lambda \in [0,1]$, $\Im\mathrm{m}\,\lambda \in [-1,1]$. Some variations of the combination of the parameters i_A and n_T are examined.

Case 3. The maximum eigenvalue is 1.1, while the distribution of the other eigenvalues is: $\Re\mathrm{e}\,\lambda \in [0,1]$, $\Im\mathrm{m}\,\lambda \in [-1,1]$. In this test we examine the relation between the parameter i_A and the order of the Arnoldi method m_A. The table shows that it is more effective to decrease the order of the Arnoldi method than to decrease the number of the Arnoldi iteration.

4.3 Comparison with the adaptive method

Some test problems from the Harwell-Boeing sparse matrix collection [2], the spectra of which are shown in the following eight figures, are solved using the block Arnoldi method. Ho's algorithm is used for reference.

The stopping criterion is based on the maximum of all computed residuals $\max_{1\le i\le r} \frac{\|Ax_i - \lambda_i x_i\|_2}{\|x_i\|_2} \equiv \max_{1\le i\le r} \frac{\|H_{m+1,m} Y_{m,r,i}\|_2}{\|Y_{m,i}\|_2} \le \epsilon$. $Y_{m,r,i}$ and $Y_{m,i}$ stand for the i-th column of the $Y_{m,r}$ and Y_m.

The following table indicate that Ho's algorithm shows better performance than the orthogonality-based method in most conditions except for the cases where the moduli of the necessary eigenvalues are much larger than those of the unnecessary eigenvalues. We may derive from the result the poor optimality of the convex hull despite its low computation cost.

orthogonality-based					Arnoldi				QR	
i_A	m_A	n_T	error	time	i_A	m_A	error	time	error	time
1	5	15	1.1E-10	11	1	15	2.3E-09	18	5.1E-15	112
2	5	15	3.6E-15	23	2	15	8.9E-16	34		

Table 1. $\lambda_{\max} = 2$, while the distribution of the other eigenvalues is $\Re\mathrm{e}\,\lambda \in [0,1]$, $\Im\mathrm{m}\,\lambda \in [-1,1]$.

orthogonality-based					Arnoldi				QR	
i_A	m_A	n_T	error	time	i_A	m_A	error	time	error	time
1	5	15	5.0E-05	15	1	15	1.2E-05	17	3.6E-15	111
2	5	15	1.1E-08	25	2	15	5.3E-11	33		
3	5	15	1.9E-11	32	3	15	3.7E-15	49		
4	5	15	5.6E-14	42	4	15	3.3E-15	63		
5	5	15	3.4E-15	54	5	15	2.8E-15	79		
3	5	20	3.0E-15	42	3	15	3.7E-15	49		
3	5	19	8.9E-16	38						
3	5	18	4.5E-13	38						
1	5	60	1.3E-14	47						

Table 2. $\lambda_{\max} = 1.5$.

orthogonality-based					Arnoldi				QR	
i_A	m_A	n_T	error	time	i_A	m_A	error	time	error	time
1	50	10	3.2E-15	240	1	50	7.5E-13	235	5.2E-15	113
1	45	20	6.9E-15	206	1	45	4.1E-10	191		
2	30	20	3.2E-15	161	2	50	7.7E+00	531		
3	20	15	6.3E-12	111	1	50	7.5E-13	236		
4	15	20	3.5E-13	112						
5	10	20	2.9E-14	96						

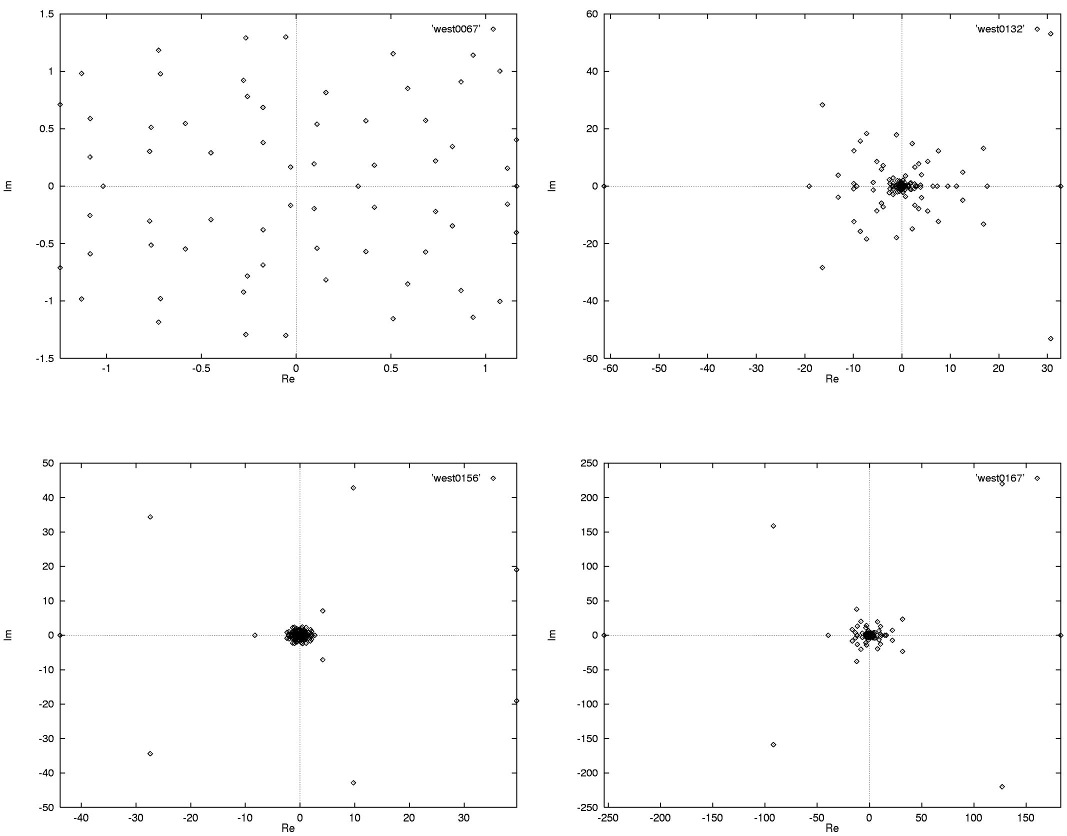
Table 3. $\lambda_{\max} = 1.1$.

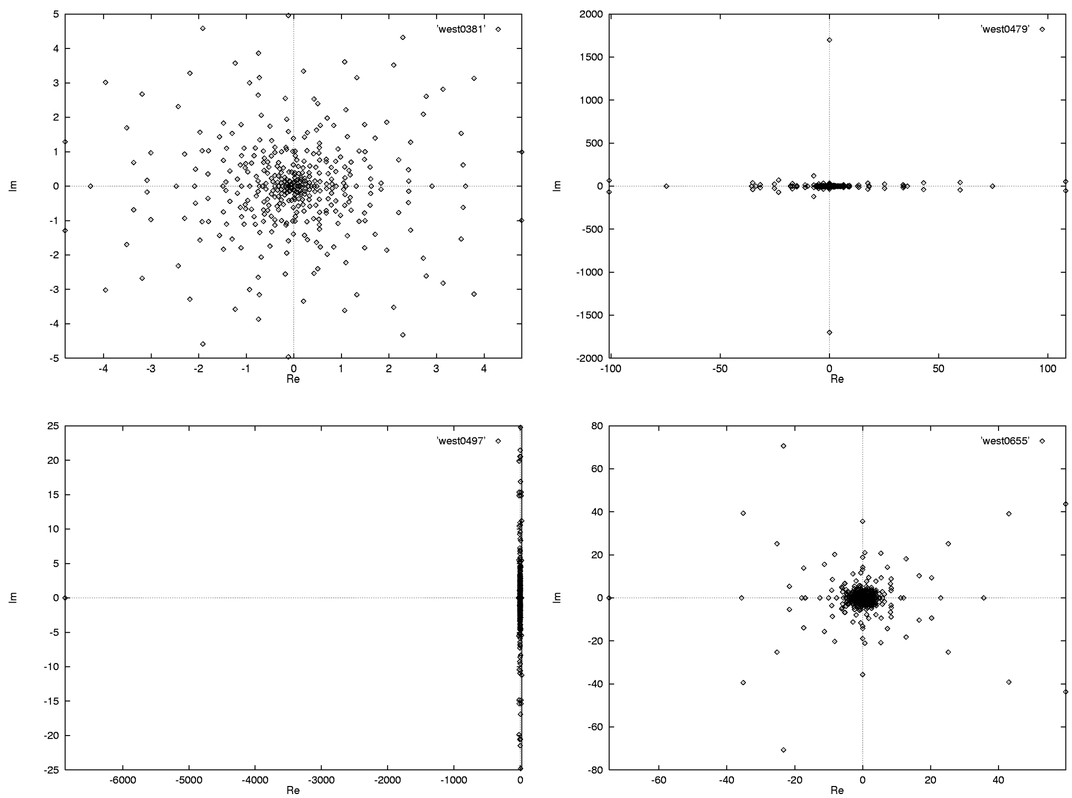
problem	`west0067`		`west0132`		`west0156`		`west0167`		`west0381`		`west0479`		`west0497`		`west0655`	
order of matrix	67		132		156		167		381		479		497		655	
size of basis	50		10		10		10		10		20		50		10	
degree of polynomial	10		10		10		10		10		10		10		10	
block size	1	1	1	1	1	2	1	1	2	2	2	2	2	2	1	2
number of iterations	3	2	6	33	2	10	2	4	25	5	8	6	3	2	5	4
computation time	312	245	37	201	19	95	19	36	723	735	376	320	857	652	172	316

Table 4. Test problems from `CHEMWEST`, a library in the Harwell-Boeing Sparse Matrix Collection. The results by Ho's algorithm (left) versus those by the orthogonality-based method (right) are listed.

5 Conclusion

This method requires the computation of $2rm\,nz + 2mr(mr+2r+1)n$ flops for the block Arnoldi method, $r^3[10m^3 + O(m^2)]$ for the computation of the eigenvalues of H_m, and $2kr\,nz + O(n)$ for the Tchebychev iteration.
It is less than those of the adaptive method, which requires the solutions of nonlinear equations for $O(k^2)$ combinations of the eigenvalues in most cases, and the least-squares based method, which costs $O(k^3)$ superfluous complexity.
We examined the other problems such as computational error by numerical experiments. The validity of our method was confirmed by the experiments using the Harwell-Boeing Sparse Matrix Collection, which is a set of standard test matrices for sparse matrix problems.
However, it can not show better performance than the adaptive algorithm in some cases where the moduli of the necessary eigenvalues are not larger than those of the unnecessary ones. A more detailed analysis of the precision of the methods is required.





References

[1] F. Chatelin, *Valeurs Propres de Matrices.* Paris: Masson, 1988.

[2] I. S. Duff, R. G. Grimes, and J. G. Lewis, "Sparse matrix test problems," *ACM Trans. Math. Softw.*, Vol. 15, 1989.

[3] G. H. Golub and C. F. V. Loan, *Matrix Computations.* Baltimore: The Johns Hopkins University Press, 1991.

[4] D. Ho, "Tchebychev acceleration technique for large scale nonsymmetric matrices," *Numer. Math.*, Vol. 56, 1990.

[5] T. A. Manteuffel, "The Tchebychev iteration for nonsymmetric linear systems," *Numer. Math.*, Vol. 28, 1977.

[6] Y. Saad, "Variations on Arnoldi's method for computing eigenelements of large unsymmetric matrices," *Lin. Alg. Appl.*, Vol. 34, 1980.

[7] Y. Saad, "Chebyshev acceleration techniques for solving nonsymmetric eigenvalue problems," *Math.Comp.*, Vol. 42, No. 166, 1984.

[8] Y. Saad, "Least squares polynomials in the complex plane and their use for solving nonsymmetric linear systems," *SIAM J. Numer. Anal.*, Vol. 24, No. 1, 1987.

[9] M. Sadkane, "A block Arnoldi-Chebyshev method for computing the leading eigenpairs of large sparse unsymmetric matrices," *Numer. Math.*, Vol. 64, 1993.