Int. J. Appl. Math. Comput. Sci., 2002, Vol.12, No.4, 533–538

amcs

INEQUALITY-BASED APPROXIMATION OF MATRIX EIGENVECTORS

ANDRÁS KOCSOR*, JÓZSEF DOMBI**, IMRE BÁLINT***

* Research Group on Artificial Intelligence, University of Szeged
Szeged, Hungary, H–6720, Aradi vrt. 1
e-mail: kocsor@inf.u-szeged.hu

** Department of Applied Informatics, University of Szeged
Szeged, Hungary, H–6720, Árpád tér 2
e-mail: dombi@inf.u-szeged.hu

*** Department of Theoretical Physics, University of Szeged
Szeged, Hungary, H–6720, Tisza L. krt. 84-86

Department of Pharmaceutical Analysis, University of Szeged
Szeged, Hungary, H–6720, Somogyi u. 4

Department of Natural Sciences, Dunaújváros Polytechnic
Dunaújváros, H–2400, Táncsics M. u. 1
e-mail: ibalint@physx.u-szeged.hu

A novel procedure is given here for constructing non-negative functions with zero-valued global minima coinciding with eigenvectors of a general real matrix A. Some of these functions are distinct because all their local minima are also global, offering a new way of determining eigenpairs by local optimization. Apart from describing the framework of the method, the error bounds given separately for the approximation of eigenvectors and eigenvalues provide a deeper insight into the fundamentally different nature of their approximations.

Keywords: eigenvectors, eigenvalues, inequalities, error bounds, iterative methods

1. Introduction

The determination of the eigenvectors and eigenvalues of large real matrices is of considerable importance to various fields of science and technology. The machinery for the solution of the problem has been worked out quite well, and we do not attempt to give a detailed overview of the relevant literature. However, the following classical monographs (Kato, 1966; Bellman, 1970; Golub *et al.*, 1996) are recommended. The methods have been devised either for determining all eigenvectors of the matrix simultaneously or successively. Various problems involve large matrices of sizes well beyond the range of the simultaneous determination of the eigenpairs, while quite a few problems require only a couple of them. The proposed family of novel algorithms are iterative and yield eigenpairs successively. Though the discussion here is restricted to real eigenvectors of real matrices, the method is also suitable for non-symmetrical matrices (with complex eigenpairs) and complex-valued matrices.

In Section 2 non-negative functions are introduced which play a key role in the algorithm. Section 3 contains error bounds for the differences between the true and approximate eigenvectors and eigenvalues, while Section 4 is devoted to the numerical results. Section 5 contains a short summary of the results obtained so far.

2. Eigenvector Functions and Their Properties

Definition 1. The inequality $\phi(\mathbf{x},\mathbf{y}) \leq \psi(\mathbf{x},\mathbf{y})$, $\mathbf{x},\mathbf{y} \in \mathbb{R}^n$, $\phi,\psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is called a proportionality inequality if

$$\phi(\mathbf{x},\mathbf{y}) \leq \psi(\mathbf{x},\mathbf{y}), \tag{1}$$

and the equality occurs if and only if $\mathbf{x}$ and $\mathbf{y}$ are linearly dependent.

Definition 2. For an arbitrary n-by-n matrix A, the function $\tau_A(\mathbf{x})$: $\mathbb{R}^n \to \mathbb{R}$ is called an 'eigenvector function' if and only if $\tau_A(\mathbf{x}) \geq 0$, the equality holding only when the vectors $\mathbf{x}$ and $A\mathbf{x}$ are linearly dependent.

Corollary 1. *For an arbitrary matrix A, using Definition 1, a special set of eigenvector functions can be con-*

structed which take the form of $\tau_A(\mathbf{x})$*:* $\mathbb{R}^n \to \mathbb{R}_+$ *(non-negative reals):*

$$\tau_A(\mathbf{x}) = \psi(\mathbf{x}, A\mathbf{x}) - \phi(\mathbf{x}, A\mathbf{x}). \tag{2}$$

Many inequality relations are suitable for defining eigenvector functions, such as Cauchy-Schwarz-Buniakowsky's, Hölder's, Minkowski's, Milne's inequality, the inequality of Aczél, Popoviciu and Bellman, the triangle inequality and the inequality based on Heron's formula (Hardy *et al.*, 1934; Dragomir, 1991; Mitrinović *et al.*, 1993). In this paper the constructed eigenvector functions make use of Cauchy-Schwarz-Buniakowsky's inequality. If $B \in \mathbb{R}^{n\times n}$ is a real, positive-definite matrix and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then

$$(\mathbf{x}^\top B\mathbf{y})^2 \le (\mathbf{x}^\top B\mathbf{x})(\mathbf{y}^\top B\mathbf{y}), \tag{3}$$

$$|\,\mathbf{x}^\top B\mathbf{y}\,| \le (\mathbf{x}^\top B\mathbf{x})^{1/2}(\mathbf{y}^\top B\mathbf{y})^{1/2} \tag{4}$$

are Cauchy-Schwarz-Buniakowsky-type inequalities, where the equality occurs iff $\mathbf{x}$ and $\mathbf{y}$ are proportional.[1]

Definition 3. Out of the multitude of possible Cauchy-Schwarz-Buniakowsky-type eigenvector functions, we decided to define the following four functions:

$$f_A(\mathbf{x}) := \|\mathbf{x}\|^2\|A\mathbf{x}\|^2 - (\mathbf{x}^\top A\mathbf{x})^2, \tag{5}$$

$$f_A^{(\omega)}(\mathbf{x}) := f_A(\mathbf{x})/\|\mathbf{x}\|^{4\omega}, \tag{6}$$

$$g_A(\mathbf{x}) := \|\mathbf{x}\|\|A\mathbf{x}\| - |\mathbf{x}^\top A\mathbf{x}|, \tag{7}$$

$$g_A^{(\omega)}(\mathbf{x}) := g_A(\mathbf{x})/\left(\|\mathbf{x}\|\|A\mathbf{x}\|\right)^{2\omega}, \tag{8}$$

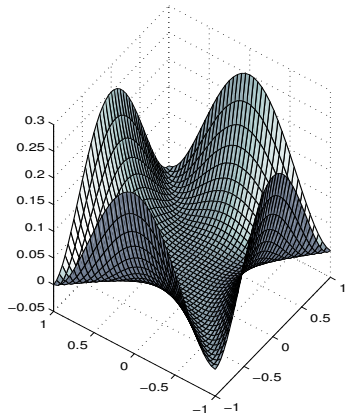
with $\|\cdot\|$ as the Euclidean norm. Equation (5) defines an n-variate, four-degree polynomial over $\mathbb{R}$, which was derived from (3) with $B = I$ (the identity matrix). Equation (6) defines the normalized version of (5). It is restricted to $\mathbf{x} \neq \mathbf{0}$ and its degree of homogeneity is $4-4\omega$, where $0 \le \omega < 1$. Functions (7) and (8) are based on relation (4). In (8), $\mathbf{x} \notin \{\mathbf{z}\colon\ A\mathbf{z} = \mathbf{0}\}$ and $0 \le \omega < 1$. If $\omega = 1/2$, (8) becomes $1 - |\cos(\xi)|$, where ξ is the angle between $\mathbf{x}$ and $A\mathbf{x}$.

Figures 1(a) and (b) show the graphs of $f_{A_0}^{(0)}(\mathbf{x})$ and $g_{A_1}^{(1/2)}(\mathbf{x})$ associated with the following 2×2 symmetric A_0 and non-symmetric A_1 matrices:

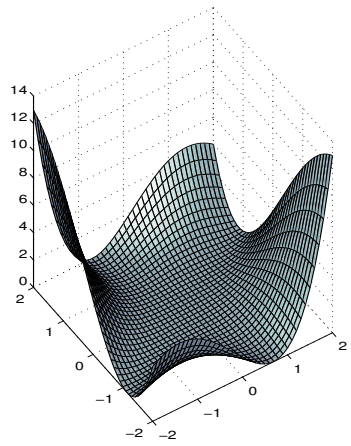
$$A_0 = \begin{pmatrix} 0.6 & -0.5 \\ -0.5 & 0.6 \end{pmatrix}, \quad A_1 = \begin{pmatrix} -0.5 & 0.6 \\ 0.3 & 0.6 \end{pmatrix}. \tag{9}$$

Lemma 1. *Let the eigenvector function* $\tau_A(\mathbf{x})$ *be* $(c > 0)$*-homogeneous.*

[1] The proof can be found in standard works, e.g. in (Hardy *et al.*, 1934).



(a)



(b)

Fig. 1. Graphs of $f_{A_0}^{(0)}(\mathbf{x})$ and $g_{A_1}^{(1/2)}(\mathbf{x})$.

(i) If $\tau_A(\mathbf{x})$ *is differentiable, then* $\nabla\tau_A(\mathbf{x}) = \mathbf{0} \Rightarrow \tau_A(\mathbf{x}) = 0$.

(ii) If $\tau_A(\mathbf{x})$ *is continuously differentiable, then* $\nabla\tau_A(\mathbf{x}) = \mathbf{0} \Longleftrightarrow \tau_A(\mathbf{x}) = 0$.

Proof. Since $\tau_A(\mathbf{x})$ is a differentiable, $(c > 0)$-degree, homogeneous function, Euler's homogeneity relation[2] (Eichhorn 1978) ensures that $\mathbf{x}^\top\nabla\tau_A(\mathbf{x}) = c\,\tau_A(\mathbf{x})$, so a zero-vector gradient implies a zero function value as well. The reverse case is trivial because the continuous differentiability of the non-negative function $\tau_A(\mathbf{x})$ implies that the derivative is a zero-vector when the function value is zero. ■

Lemma 2. *For an arbitrary, strictly monotonic, increasing function* $\kappa : \mathbb{R} \to \mathbb{R}$ *and a homogeneous eigenvector function* $\tau_A(\mathbf{x})$*, the composite function* $\tilde{\kappa}(\mathbf{x}) :=$

[2] Let $f\colon\ \mathbb{R}^n \to \mathbb{R}$ be r-homogeneous and differentiable on the open and connected set $D \subseteq \mathbb{R}^n$. Then

$$rf(\mathbf{w}) = \frac{\partial f(\mathbf{w})}{\partial w_1}w_1 + \frac{\partial f(\mathbf{w})}{\partial w_2}w_2 + \cdots + \frac{\partial f(\mathbf{w})}{\partial w_n}w_n, \ \mathbf{w} \in D.$$

$\kappa(\tau_A(\mathbf{x}))-\kappa(0)$ *is also an eigenvector function. Furthermore, if* κ *and* $\tau_A(\mathbf{x})$ *are differentiable, then* $\nabla\tilde{\kappa}(\mathbf{x}) = \mathbf{0} \Rightarrow \tilde{\kappa}(\mathbf{x}) = 0$.

Proof. Since $\tau_A(\mathbf{x})$ is an eigenvector function, the strictly increasing monotonous character of κ ensures that $\tilde{\kappa}(\mathbf{x})$ is also an eigenvector function. If κ and $\tau_A(\mathbf{x})$ are differentiable and $\tau_A(\mathbf{x})$ is homogeneous with positive degree, then $\nabla\tau_A(\mathbf{x}) = \mathbf{0} \;\Rightarrow\; \tau_A(\mathbf{x}) = \mathbf{0}$, and $\kappa' > 0$ implies $\nabla\tilde{\kappa}(\mathbf{x}) = \nabla\left(\kappa(\tau_A(\mathbf{x})) - \kappa(0)\right) = \kappa'(\tau_A(\mathbf{x}))\nabla\tau_A(\mathbf{x}) = \mathbf{0} \;\Rightarrow\; \nabla\tau_A(\mathbf{x}) = 0 \;\Rightarrow\; \tilde{\kappa}(\mathbf{x}) = 0$, which proves the assertion. ■

One example of a composite function is $\kappa(x) := \exp(x)$ with $\tau_A(\mathbf{x}) := f_A(\mathbf{x})$. The eigenvector functions defined above are suitable candidates for determining the eigenpairs of matrices via optimization.

3. Error Bounds for the Approximation of Eigenvalues and Eigenvectors

In this section the following mathematical assertions relate only to $f_A^{(0)}$ and $g_A^{(1/2)}$, but similar propositions can be made for general ω's and non-symmetric matrices, too. By approximating the true eigenpairs $(\mathbf{u}, \lambda)$ of a given matrix with $(\mathbf{x}, \nu)$, various measures can be used for the pairwise distances of the true and approximate eigenvalues, as well as those of the true and approximate eigenvectors. The following lemma shows that the widely used Wilkinson norm $\|A\mathbf{x} - \mathbf{x}\nu\|/\|\mathbf{x}\|$ of (Parlett 1980; Wilkinson 1965) is closely related to $f_A^{(0)}(\mathbf{x})$.

Lemma 3. *For any non-zero* $\mathbf{x} \in \mathbb{R}^n$ *and symmetric matrix* A*, if*

$$(f_A^{(0)}(\mathbf{x}))^{1/2}/\|\mathbf{x}\|^2 \le \epsilon, \tag{10}$$

then an eigenvalue λ *necessarily exists for which* $|\lambda - \sigma_A(\mathbf{x})| \le \epsilon$*, where* $\sigma_A(\mathbf{x}) = \mathbf{x}^\top A\mathbf{x}/\mathbf{x}^\top\mathbf{x}$ *is the Rayleigh quotient.*

Proof. For any non-zero $\mathbf{x} \in \mathbb{R}^n$ and for any symmetric matrix A,

$$\left(f_A^{(0)}(\mathbf{x})\right)^{1/2}/\|\mathbf{x}\|^2 = \|A\mathbf{x} - \mathbf{x}\sigma_A(\mathbf{x})\|/\|\mathbf{x}\|. \tag{11}$$

Direct computation yields the following sequence of equalities, which prove the above statement:

$$\begin{aligned}
&\|A\mathbf{x} - \mathbf{x}\sigma_A(\mathbf{x})\|^2\,(\mathbf{x}^\top\mathbf{x})^2 \\
&\quad = \left\|A\mathbf{x} - \mathbf{x}(\mathbf{x}^\top A\mathbf{x}/\mathbf{x}^\top\mathbf{x})\right\|^2(\mathbf{x}^\top\mathbf{x})^2 \qquad (12)\\
&\quad = \mathbf{x}^\top\mathbf{x}((\mathbf{x}^\top\mathbf{x})(A\mathbf{x})^\top(A\mathbf{x}) - (\mathbf{x}^\top A^\top\mathbf{x})(\mathbf{x}^\top A\mathbf{x})) \\
&\quad = \mathbf{x}^\top\mathbf{x} f_A^{(0)}(\mathbf{x}). \qquad (13)
\end{aligned}$$

Wilkinson's result (Parlett 1980; Wilkinson 1965) implies that for any non-zero $\mathbf{x} \in \mathbb{R}^{\mathbf{n}}$ and any scalar ν, an eigenvalue λ exists that satisfies the inequality

$$|\lambda - \nu| \le \|A\mathbf{x} - \mathbf{x}\nu\|/\|\mathbf{x}\|. \tag{14}$$

Taking into account (11) and replacing the Rayleigh-quotient with ν, we obtain the desired result. ■

The following theorems provide bounds for the approximation of the eigenvectors. Let the eigenvalues of the symmetric matrix A be $\lambda_1, \lambda_2, \ldots, \lambda_n$, and let the associated normalized eigenvectors be $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$, respectively. Let the angles between the eigenvectors and vector $\mathbf{x}$ be $\alpha_1, \alpha_2, \ldots, \alpha_n$. If $\sigma_A(\mathbf{x})$ approximates in the best manner λ_m, the angle between the associated eigenvector $\mathbf{u}_m$ and $\mathbf{x}$ provides a measure of the accuracy of approximation.

Theorem 1. *For an arbitrary non-zero* $\mathbf{x} \in \mathbb{R}^n$ *and a real, symmetric matrix* A*,*

$$\left(f_A^{(0)}(\mathbf{x})\right)^{1/2}/\|\mathbf{x}\|^2 \le \epsilon \implies \sin^2\alpha_m \le \epsilon/|\lambda_m - \bar{\lambda}_m|, \tag{15}$$

where $\lambda_m \neq 0$ *is the eigenvalue closest to* $\sigma_A(\mathbf{x})$ *and* $\bar{\lambda}_m$ *is the following convex linear combination of the complementary part of the spectrum:*

$$\bar{\lambda}_m = \frac{\sum\limits_{j,\, j\neq m} \lambda_j \cos^2\alpha_j}{\sum\limits_{j,\, j\neq m} \cos^2\alpha_j}. \tag{16}$$

Proof. The proof can be found in Appendix. ■

The essence of this theorem is that the accuracy of the approximation of an eigenvector depends not only on the bound ϵ but also on the structure of the spectrum. The value of $\bar{\lambda}_m$ depends on the distribution of eigenvalues in the complementary subset $\{\lambda_j\}_{j=1,\ldots,n} \setminus \lambda_m$, as well as on the actual position of the trial eigenvector $\mathbf{x}$ encoded in the parameters $\cos^2(\alpha_j)$. If the spectrum is almost degenerate, the bound, of course, loses its predictive power. However, when approximating the lowest eigenvalue, since $\bar{\lambda}_m$ is a convex sum of the remaining eigenvalues, it must be greater than or equal to the second one. If an estimation for $|\lambda_1 - \lambda_2|$ is available, Theorem 1 offers a good measure for judging the accuracy of the numerical calculation.

Theorem 2 is based on the relationship between $g_A^{(1/2)}(\mathbf{x})$ and $\tan^2(\alpha_m)$. The eigenvector $\mathbf{u}_m$ remains the best approximated eigenvector, α_m is again the angle between $\mathbf{x}$ and $\mathbf{u}_m$, and the new parameter

$$\bar{\bar{\lambda}}_m = \left(\frac{\sum\limits_{j,\, j\neq m} \lambda_j^2 \cos^2\alpha_j}{\sum\limits_{j,\, j\neq m} \cos^2\alpha_j}\right)^{1/2} \tag{17}$$

is similar in form to $\bar{\lambda}$ in Theorem 1.

Theorem 2. *For any real symmetric matrix A with eigenvalue $\lambda_m \neq 0$,*

$$\frac{1-|c_1|\tan^2(\alpha_m)}{1+\frac{c_2^2+1}{2}\tan^2(\alpha_m)} \leq 1-g_A^{(1/2)}(\mathbf{x}) \leq \frac{1+|c_1|\tan^2(\alpha_m)}{1+|c_2|\tan^2(\alpha_m)}, \quad (18)$$

where $c_1 = \bar{\lambda}_m/\lambda_m$ and $c_2 = \bar{\bar{\lambda}}_m/\lambda_m$.

Proof. The proof can be found in Appendix. ■

The gist of Theorem 2 is similar to that of Theorem 1, and expresses the intertwined nature of the accuracy of the approximation of an eigenvector and the structure of the spectrum. Similarly as in Theorem 1, we need acceptable estimates for the parameters c_1 and c_2 to keep the predictive power of the theorem, but unlike Theorem 1, Theorem 2 offers both lower and upper bounds on the accuracy of the approximation of an eigenvector. Beyond the bounds both theorems provide a clear insight into the nature of eigenvector approximations. While Lemma 3 ensures that an eigenvalue exists in a bounded vicinity of the Rayleigh-quotient, the bounds in Theorems 1 and 2 depend on the structure of the spectrum and on the actual position vector $\mathbf{x}$. In practice, their application requires therefore that we know something about the structure of the spectrum, either from an independent source or by making an educated guess.

4. Numerical Results

In order to illustrate how the eigenvector function-based search method works, we constructed two numerical algorithms. The first one is a general scheme, while the second is a version specifically tailored to sparse symmetric matrices. Definitions 2, 3 and Lemma 1 ensure that the minimum points of the eigenvector functions in question coincide with the eigenvectors of the underlying matrix. Thus the optimization of any of the constructed eigenvector functions (5)–(8) provides a single eigenpair of the symmetric matrix A.

Algorithm 1 describes the general scheme. Several minimization algorithms are available and suitable (Bazaraa *et al.*, 1993) for use in Algorithm 1, the choice depending mainly on the size of the matrix. The program function $bound_A(\mathbf{x})$ (any of the functions in Section 3) provides a measure for the actual accuracy of the approximation of an eigenvalue or an eigenvector. The input is a starting trial vector $\mathbf{x_0}$, and a positive bound ϵ is related to the underlying eigenvector function $\tau_A(\mathbf{x})$ and matrix A.

Algorithm 1. General optimization-based eigenvector search.

Require: $\tau_A(\mathbf{x})$ be an eigenvector function, $\mathbf{x_0}$ a starting trial vector, $bound_A(\mathbf{x})$ a measure of the accuracy of the actual approximation, and ϵ a suitable bound.

Ensure: an eigenpair approximation $(\mathbf{x_k}, \sigma_A(\mathbf{x_k}))$

$k \leftarrow 0$

while $bound_A(\mathbf{x_k}) \geq \epsilon$

$\mathbf{x_{k+1}} \leftarrow$ successor of $\mathbf{x_k}$, when minimizing $\tau_A(\mathbf{x})$ using an iterative minimization algorithm

$k \leftarrow k+1$

end while

$\sigma_A(\mathbf{x_k}) \leftarrow \mathbf{x_k}^\top A\mathbf{x_k}/\mathbf{x_k}^\top\mathbf{x_k}$.

The following numerical performance tests were accomplished on medium-sized symmetric and non-symmetric matrices with uniformly distributed random elements in the interval $[-1,1]$, making use of the functions (6) and (8). For the optimization algorithm the very efficient BFGS scheme (Bazaraa *et al.*, 1993) was chosen. It was started with a random trial vector and was terminated when the condition

$$\|A\mathbf{x} - \mathbf{x}\sigma_A(\mathbf{x})\|/\|\mathbf{x}\| < \epsilon \quad (19)$$

was met. Table 1 shows the average numbers of iteration steps necessary to attain an accuracy of 10^{-7} in the approximation of eigenvalues.

As the table clearly shows, the rate of convergence for both cases speeds up with an increase in matrix size. The degree of homogeneity also affects the algorithm's numerical efficiency. We found that $\omega = 0.5$ was the best value for both symmetric and non-symmetric matrices. In both cases the average number of iteration steps

Table 1. Convergence data obtained from determining the eigenpairs of symmetric and non-symmetric matrices using eigenvector functions $f_A^{(\omega)}(\mathbf{x})$ and $g_A(\mathbf{x})$.

	SYMMETRIC			
SIZE n	$f_A^{(\omega)}(\mathbf{x})$ $\omega=0.25$	$f_A^{(\omega)}(\mathbf{x})$ $\omega=0.5$	$f_A^{(\omega)}(\mathbf{x})$ $\omega=0.75$	$g_A(\mathbf{x})$
25	43	31	32	30
50	68	56	59	56
100	143	105	114	107
200	269	204	210	209
	NON-SYMMETRIC			
SIZE n	$f_A^{(\omega)}(\mathbf{x})$ $\omega=0.25$	$f_A^{(\omega)}(\mathbf{x})$ $\omega=0.5$	$f_A^{(\omega)}(\mathbf{x})$ $\omega=0.75$	$g_A(\mathbf{x})$
25	74	61	53	44
50	119	119	85	69
100	198	173	154	132
200	325	295	307	239

was quite low and close to n for symmetric matrices, and about $1.5n$ for non-symmetric matrices (n being the dimension), each step having an $O(n^2)$ operation-step dependence.

Algorithm 2 is a special variant of the general scheme, which is suitable for large, sparse matrices. Here the eigenvector function $\tau_A(\mathbf{x})$ is

$$f_A(\mathbf{x}) = \|\mathbf{x}\|^2\|A\mathbf{x}\|^2 - (\mathbf{x}^\top A\mathbf{x})^2, \tag{20}$$

where $\mathbf{A}$ is a sparse and symmetric matrix, while the corresponding gradient of $f_A(\mathbf{x})$ is

$$2(\mathbf{x}^\top\mathbf{x}A^2\mathbf{x} + \mathbf{x}(A\mathbf{x})^\top A\mathbf{x} - 2\mathbf{x}^\top A\mathbf{x}(A\mathbf{x})). \tag{21}$$

The bound $bound_A(\mathbf{x}) = (f_A(\mathbf{x}))^{1/2}/\|\mathbf{x}\|^2$ $(= \|A\mathbf{x} - \mathbf{x}\sigma_A(\mathbf{x})\|/\|\mathbf{x}\|)$ measures the accuracy of the actual eigenvalue approximation and ϵ is a suitable bound. Here α = linesearch($f_A(\mathbf{x})$,$\mathbf{x}$,$\mathbf{p}$) minimizes the function $f_A(\mathbf{x} + \alpha\mathbf{p})$ with respect to a scalar α. Here $\mathbf{p}$ is known as the search direction, while α is the so-called step length. In the WHILE-loop, $\mathbf{p_k}$ is refreshed by using a limited-memory quasi-Newton formula.

Algorithm 2. Optimization-based eigenpair search for sparse symmetric matrices.

Require: $f_A(\mathbf{x})$ be associated with a given sparse and symmetric matrix A, and ϵ a suitable bound.

Ensure: an eigenpair approximation $(\mathbf{x_k}, \sigma_A(\mathbf{x_k}))$

$\mathbf{g_0} \leftarrow \nabla f_A(\mathbf{x_0})$

$\mathbf{p_0} \leftarrow -\mathbf{g_0}$

$\alpha_0 \leftarrow$ linesearch($f_A(\mathbf{x})$,$\mathbf{x_0}$,$\mathbf{p_0}$)

$\mathbf{x_1} \leftarrow \mathbf{x_0} + \alpha_0\mathbf{g_0}$

$\mathbf{g_1} \leftarrow \nabla f_A(\mathbf{x_1})$

$\mathbf{y_0} \leftarrow \mathbf{g_1} - \mathbf{g_0}$

$\mathbf{s_0} \leftarrow \mathbf{x_1} - \mathbf{x_0}$

$k \leftarrow 1$

while $(f_A(\mathbf{x}))^{1/2}/\|\mathbf{x}\|^2 \geq \epsilon$

$$\mathbf{p_k} \leftarrow -\mathbf{g_k} + \frac{1}{\mathbf{y_{k-1}}^\top\mathbf{s_{k-1}}}(\mathbf{s_{k-1}}^\top\mathbf{g_k}\mathbf{y_{k-1}} + \mathbf{y_{k-1}}^\top\mathbf{g_k}\mathbf{s_{k-1}}) - \frac{\mathbf{s_{k-1}}^\top\mathbf{g_k}}{\mathbf{y_{k-1}}^\top\mathbf{s_{k-1}}} \times \left(1 + \frac{\mathbf{y_{k-1}}^\top\mathbf{y_{k-1}}}{\mathbf{y_{k-1}}^\top\mathbf{s_{k-1}}}\right)\mathbf{s_{k-1}}$$

$\alpha_k \leftarrow$ linesearch($f_A(\mathbf{x})$,$\mathbf{x_k}$,$\mathbf{p_k}$)

$\mathbf{x_{k+1}} \leftarrow \mathbf{x_k} + \alpha_k\mathbf{g_k}$

$\mathbf{g_{k+1}} \leftarrow \nabla f_A(\mathbf{x_{k+1}})$

$\mathbf{y_k} \leftarrow \mathbf{g_{k+1}} - \mathbf{g_k}$

$\mathbf{s_k} \leftarrow \mathbf{x_{k+1}} - \mathbf{x_k}$

$k \leftarrow k + 1$

end while

$\sigma_A(\mathbf{x_k}) \leftarrow \mathbf{x_k}^\top A\mathbf{x_k}/\mathbf{x_k}^\top\mathbf{x_k}$.

To test the algorithm, symmetric matrices of size 2000, 4000, 8000 and 16000 were employed with a 5% sparsity ratio. The required accuracy of the eigenvalues was the same as before. The test machine was a standard Pentium III PC with a 700 MHz processor and a 512 Mb core memory. The run-times were averaged over a set of 25 sample matrices, and the cpu times were found to be around 7, 29, 121 and 480 seconds using the MATLAB system.

5. Conclusions

Non-negative eigenvector functions were constructed, whose formulations were based on the so-called proportionality inequalities. The zero-valued minima of these functions coincide with the eigenvectors of an underlying matrix, yielding selected eigenpairs via an optimization procedure. The efficiency of the proposed method was illustrated by applying it to random medium-sized and large symmetric sparse matrices. The eigenvector functions were used to establish error bounds on the eigenvalue and eigenvector approximations. The first bound guarantees the accuracy of the approximation of the eigenvalues, while the other two bounds guarantee the accuracy of approximation if reasonable guesses can be made about the structure of the spectrum. Moreover, these bounds offer a clear insight into the nature of the approximation of eigenvalues and eigenvectors, which are quite different in character.

References

Bazaraa M.S. , Sherali H.D. and Shetty C.M. (1993): *Nonlinear Programming Theory and Algorithms*. — New York: Wiley.

Bellman R. (1970): *Introduction to Matrix Analysis (2nd Ed.)*. — New York: McGraw-Hill.

Dragomir S.S. and Arslangić Š.Z. (1991): *A refinement of the Cauchy-Buniakowski-Schwarz inequality for real numbers*. — Radovi Matematičui (Sarajevo) Vol. 7, No. 2, pp. 299–303.

Eichhorn W. (1978): *Functional Equations in Economics*. — Reading, MA: Addison-Wesley.

Golub G.H. and Van Loan C.F. (1996): *Matrix Computations (3rd Ed.)*. — Baltimore: The John Hopkins University Press.

Hardy G.H., Littlewood J.E. and Pólya G. (1934): *Inequalities*. — London: Cambridge University Press.

Kato T. (1966): *Perturbation Theory of Linear Operators*. — New York: Springer.

Mitrinović D.S., Pečarić and Fink J.E. (1993): *Classical and New Inequalies in Analysis*. — London: Kluwer.

Parlett B. (1980): *The Symmetric Eigenvalue Problem*. — Englewodd Cliffs: Prentice-Hall.

Wilkinson J.H. (1965): *The Algebraic Eigenvalue Problem*. — Oxford: Oxford University Press.

Appendix

Proof of Theorem 3. By Lemma 3, $(f_A^{(0)}(\mathbf{x}))^{1/2}/\|\mathbf{x}\|^2 \leq \epsilon \Rightarrow |\lambda_m - \sigma_A(\mathbf{x})| \leq \epsilon$. Taking into account $\cos\alpha_j = \mathbf{u}_j^\top \mathbf{x}/\|\mathbf{x}\|$ and $A = \sum_j \lambda_j \mathbf{u}_j \mathbf{u}_j^\top$, we have

$$\begin{aligned}|\lambda_m - \sigma_A(\mathbf{x})| &= \left|\lambda_m - \frac{\mathbf{x}^\top}{\|\mathbf{x}\|}\sum_j \lambda_j \mathbf{u}_j\mathbf{u}_j^\top \frac{\mathbf{x}}{\|\mathbf{x}\|}\right| \\ &= \left|\lambda_m - \lambda_m \cos^2\alpha_m - \bar{\lambda}_m \sum_{j,\ j\neq m} \cos^2\alpha_j\right| \\ &= \left|(\lambda_m - \bar{\lambda}_m)\sin^2\alpha_m\right| \leq \epsilon. \end{aligned} \tag{A1}$$

■

Proof of Theorem 4. We have $g_A^{(1/2)}(\mathbf{x}) = 1 - |\cos(\xi)|$ and

$$\begin{aligned}\cos(\xi) &= \frac{\mathbf{x}^\top A\mathbf{x}}{\|A\mathbf{x}\|\|\mathbf{x}\|} = \frac{\frac{\mathbf{x}^\top}{\|\mathbf{x}\|}\sum_j \lambda_j \mathbf{u}_j\mathbf{u}_j^\top \frac{\mathbf{x}}{\|\mathbf{x}\|}}{\|\sum_j \lambda_j \mathbf{u}_j\mathbf{u}_j^\top \frac{\mathbf{x}}{\|\mathbf{x}\|}\|} \\ &= \frac{\lambda_m \cos^2(\alpha_m) + \bar{\lambda}_m \sum_{j,\ j\neq m}\cos^2\alpha_j}{\sqrt{\lambda_m^2\cos^2(\alpha_m) + \bar{\bar{\lambda}}_m^2 \sum_{j,\ j\neq m}\cos^2\alpha_j}} \\ &= \frac{\cos^2(\alpha_m) + c_1(1-\cos^2(\alpha_m))}{\sqrt{\cos^2(\alpha_m) + c_2^2(1-\cos^2(\alpha_m))}}. \end{aligned} \tag{A2}$$

Simplifying the fraction by dividing the numerator and denominator by $\cos^2(\alpha_m)(\neq 0)$, and taking the absolute value, we obtain

$$|\cos(\xi)| = \frac{\left|1 + c_1\tan^2(\alpha_m)\right|}{\sqrt{1+c_2^2\tan^2(\alpha_m)}\sqrt{1+\tan^2(\alpha_m)}}. \tag{A3}$$

The following lower and upper bounds can be obtained by minorizing/majorizing the numerator, and by majorizing/minorizing the denominator:

$$\begin{aligned}&\frac{1-|c_1|\tan^2(\alpha_m)}{1+\frac{c_2^2+1}{2}\tan^2(\alpha_m)} \\ &\quad\leq \frac{\left|1 + c_1\tan^2(\alpha_m)\right|}{\sqrt{1+c_2^2\tan^2(\alpha_m)}\sqrt{1+\tan^2(\alpha_m)}} \\ &\quad\leq \frac{1+|c_1|\tan^2(\alpha_m)}{1+|c_2|\tan^2(\alpha_m)}. \end{aligned} \tag{A4}$$

In the case of the numerator the minorizing/majorizing task is trivial. However, in the case of the denominator the majorizing task is based on arithmetic and geometric mean relations, while the minorizing task is based on the following trigonometric inequality:

$$\begin{aligned}&1+|c_2|\tan^2(\alpha_m) \\ &\quad\leq \sqrt{1+c_2^2\tan^2(\alpha_m)}\sqrt{1+\tan^2(\alpha_m)}. \end{aligned} \tag{A5}$$

■

Received: 6 September 2001
Revised: 10 May 2002