Annales Univ. Sci. Budapest., Sect. Comp. 20 (2001) 237-253

AN OPTIMIZATION BASED ALGORITHM FOR DETERMINING EIGENPAIRS OF LARGE REAL PAIRS

A. Kocsor, J. Dombi and I. Bálint
(Szeged, Hungary)

Abstract. The determination of the eigenpairs of real matrices is ascribed to a local optimization problem providing primarily the eigenvectors of the matrix. Suitable non-negative functions are constructed with coinciding local and global minima, which are located at the eigenvectors of the underlying matrix. Some properties of these eiegenvector-functions are investigated and proved.

1. Introduction

The detemination of the eigenvectors and eigenvalues of large real matrices is of considerable inportance in various fields of science and technology. The machinery for the solution of the problem is worked out quite well, and we do not attempt to give an overview of the referred literature. Generally the methods are devised either for determining all eingenvectors of the matrix simultaneously or only one-by-one succesively. Some problems provide large matrices the sizes of which are beyond the possibilities of simultaneous determination of all eigenvectors and eigenvalues, the more the nature of the problem often requires only a part of the eigenpairs.

The proposed novel algorithm belongs to the iterative class. Non-negative, homogeneous functions have been established with coinciding local and global optima, which are located exactly at the eigenvectors of the underlying matrix. Therefore the eigenvectors of the matrix can be found by well behaving optimization algorithms, as the minima of the associated 'eigenvector-function'. There is a unique property of the algorithm that the convergence to a selected eigenvector is not guided by the magnitude of the associated eigenvalue as generally is the case among the iterative methods [1,2,4].

The paper consists of four sections and an Appendix: the subsequent one deals with the specification and some properties of the eigenvector-function, the third section is devoted to the optimization problems, numerical results and an example for the optimization with closely packed eigenvalues and the last one contains the conclusions, while the Appendix collects technical results. In this report the discussion is restricted to real eigenvectors of real matrices, although the algorithm can be applied to complex cases, as well. It is stressed that non-symmetrical matrices with complex eigenpairs are not excluded from the investigation, only the scope of the present paper is restricted to the determination of the real eigenpairs and a separate report will be dealing with the complex cases. In the whole paper, the $\|\cdot\|$ notation will be used for the Euclidean norm, ∇f denotes the gradient of a function $f:\mathbb{R}^n\to\mathbb{R}$ and $H(f)$ denotes its Hessian.

2. The eigenvector-functions

Let us consider the function $f_A(\mathbf{x}):\mathbb{R}^n\to\mathbb{R}$ generated by matrix $A = [a_{ij}]$,

$$f_A(\mathbf{x}) := \mathbf{x}^\top\mathbf{x}(A\mathbf{x})^\top(A\mathbf{x}) - (\mathbf{x}^\top A\mathbf{x})^2, \tag{2.1}$$

where $f_A(\mathbf{x})$ is an n-variable, 4-degree polynomial over the reals ($\mathbb{R}$) with variables $x_1,\dots,x_n$ and coefficients a_{ij} ($i,j\in\{1,\dots,n\}$).

Proposition 2.1. *Function $f_A(\mathbf{x})$ exhibits the properties:*

(i) By construction consists exclusively of 4-degree terms.

$$f_A(\mathbf{x}) = \sum_{1\le i\le j\le k\le l\le n} c_{ijkl}x_ix_jx_kx_l,$$

where the real coefficients c_{ijkl} are determined by matrix A.

(ii) $f_A(\mathbf{x})$ can be differentiated 4-times continuously.

(iii) $f_A(\mathbf{x})$ is a 4-degree homogeneous function, i.e.

$$f_A(k\mathbf{x}) = k^4 f_A(\mathbf{x}),\quad k\in\mathbb{R}.$$

Proof. The proof is quite trivial.

Another eigenvector-function (among the many possible) is

$$g_A(\mathbf{x}) = \sqrt{\mathbf{x}^\top\mathbf{x}}\sqrt{(A\mathbf{x})^\top(A\mathbf{x})} \pm \mathbf{x}^\top A\mathbf{x}, \tag{2.2}$$

where $\pm$ accounts for the sign of the eigenvalue selecting the negative for positive eigenvalues and the positive for the negative ones. For both of the functions, the normalized versions will also be applied.

Definition 2.2. For any real number $0 \leq w \leq 1$

$$f_A^{(w)}(\mathbf{x}) := \begin{cases} \frac{f_A(\mathbf{x})}{\|\mathbf{x}\|^{4w}}, & 0 < w \leq 1, \\ f_A(\mathbf{x}), & w = 0, \end{cases} \tag{2.3}$$

$$g_A^{(w)}(\mathbf{x}) := \begin{cases} \frac{g_A(\mathbf{x})}{\|\mathbf{x}\|^{2w}}, & 0 < w \leq 1, \\ g_A(\mathbf{x}), & w = 0 \end{cases} \tag{2.4}$$

with $\mathbf{x} \neq \mathbf{0}$ for $0 < w \leq 1$. In the forthcoming discussion eigenvector-functions (2.2) and (2.3) will be dealt with For illustration, the graphs of (2.2) and (2.3) relating to the 2×2 symmetric A_0 and non-symmetric A_1 matrices

$$A_0 = \begin{pmatrix} \frac{3}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{3}{2} \end{pmatrix}, \quad A_1 = \begin{pmatrix} \frac{3}{2} & 1 \\ \frac{1}{4} & \frac{3}{2} \end{pmatrix}$$

are displayed on Figure 2.1. The following statement is directly formulated for (2.3), but can be formulated for (2.2), too.

Lemma 2.3. *Function $f_A^{(w)}(\mathbf{x})$, $0 \leq w \leq 1$ exhibits the following properties:*

(i) $f_A^{(w)}(\mathbf{x}) \geq 0$ and $f_A^{(w)}(\mathbf{x}) = 0$, if and only if $\mathbf{x}$ is a (real) eigenvector of matrix A, or $\mathbf{x} = \mathbf{0}$ with $w = 0$.

(ii) $\nabla f_A^{(w)}(\mathbf{x}) = \mathbf{0} \Rightarrow f_A^{(w)}(\mathbf{x}) = 0$.

(iii) $\nabla f_A^{(w)}(\mathbf{x}) = \mathbf{0} \Leftarrow f_A^{(w)}(\mathbf{x}) = 0$.

Proof. (*i*) If applying the Cauchy-Schwartz inequality for the vectors $\mathbf{x}$ and $A\mathbf{x}$,

$$\sqrt{\mathbf{x}^\top\mathbf{x}}\sqrt{(A\mathbf{x})^\top(A\mathbf{x})} \geq (\mathbf{x}^\top A\mathbf{x}),$$

which results in

$$f_A^{(0)}(\mathbf{x}) = \mathbf{x}^\top \mathbf{x}(A\mathbf{x})^\top (A\mathbf{x}) - (\mathbf{x}^\top A\mathbf{x})^2 \geq 0.$$

The equality occurs, if and only if the vectors $\mathbf{x}$ and $A\mathbf{x}$ are linearly dependent. This statement is valid also for (2.3).

An optimization based algorithm for determining eigenpairs of large real matrices

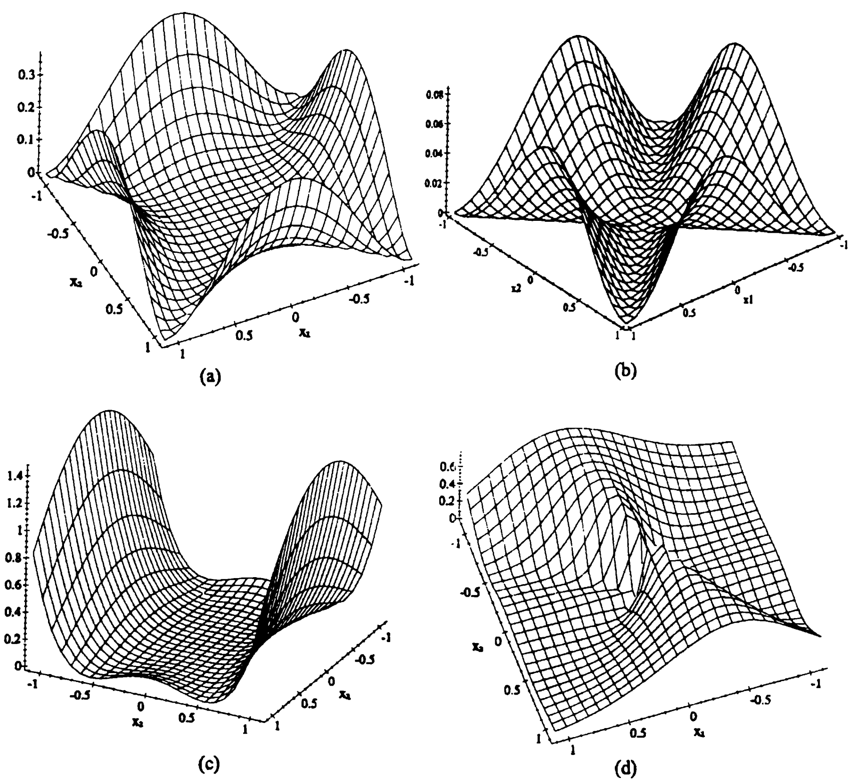

Fig.2.1. The graphs $f_{A_0}^{(0)}(\mathbf{x})$, $g_{A_0}(\mathbf{x})$, $f_{A_1}^{(1)}(\mathbf{x})$, $f_{A_1}^{(1)}(\mathbf{x})$

(*ii*) Since $f_A^{(w)}(\mathbf{x})$ is a homogeneous function of degree $4 - 4w$, Euler's theorem ensures that

$$\mathbf{x}^\top \nabla f_A^{(w)}(\mathbf{x}) = (4 - 4w) f_A^{(w)}(\mathbf{x}), \Longrightarrow \frac{\mathbf{x}^\top \nabla f_A^{(w)}(\mathbf{x})}{(4 - 4w)} = f_A^{(w)}(\mathbf{x}).$$

In case of $4-4w \neq 0$ the statement follows immediately, for $w = 1$ L'Hospital's rule is to be applied.

(*iii*) This implication follows from the nonnegative and continuous nature of $f_A^{(w)}(\mathbf{x})$ ($\mathbf{x} \neq \mathbf{0}$ in case of $0 < w \leq 1$).

2.1. Error bounds

By approximating the true eigenpairs $(\mathbf{u}, \lambda)$ of the considered matrices with $(\mathbf{x}, \sigma)$, various measures can be used for the pairwise distances of the true and approximate eigenvalues, as well as those of true and approximate eigenvectors. It will be shown that the widely used norm [8, 9], $||A\mathbf{x} - \mathbf{x}\sigma||/||\mathbf{x}||$, is closely relating to $f_A^{(w)}(\mathbf{x})$ and the subsequent three propositions.

Proposition 2.4. *For any non-zero* $\mathbf{x} \in \mathbb{R}^n$ *and for any considered matrix* A,

$$\frac{\sqrt{f_A^{(w)}(\mathbf{x})}}{||\mathbf{x}||^{2-2w}} = \frac{||A\mathbf{x} - \mathbf{x}\sigma||}{||\mathbf{x}||}; \quad \sigma = \frac{\mathbf{x}^\mathsf{T} A\mathbf{x}}{\mathbf{x}^\mathsf{T}\mathbf{x}}, \tag{2.5}$$

where σ *is the Rayleigh-quotient.*

Proof. Since $f_A^{(w)}(\mathbf{x}) = f_A^{(0)}(\mathbf{x})/||x||^{4w}$, it is enough to show that

$$\frac{\sqrt{f_A^{(0)}(\mathbf{x})}}{||\mathbf{x}||^2} = \frac{||A\mathbf{x} - \mathbf{x}\sigma||}{||\mathbf{x}||}.$$

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

$$||A\mathbf{x} - \mathbf{x}\sigma||^2(\mathbf{x}^\mathsf{T}\mathbf{x})^2 = \left\|A\mathbf{x} - \mathbf{x}\frac{\mathbf{x}^\mathsf{T} A\mathbf{x}}{\mathbf{x}^\mathsf{T}\mathbf{x}}\right\|^2 (\mathbf{x}^\mathsf{T}\mathbf{x})^2 =$$
$$= \mathbf{x}^\mathsf{T}\mathbf{x}((\mathbf{x}^\mathsf{T}\mathbf{x})(A\mathbf{x})^\mathsf{T}(A\mathbf{x}) - (\mathbf{x}^\mathsf{T} A^\mathsf{T}\mathbf{x})(\mathbf{x}^\mathsf{T} A\mathbf{x})) = \mathbf{x}^\mathsf{T}\mathbf{x} f_A^{(0)}(\mathbf{x}).$$

Proposition 2.5. *By Wilkinson's result, for any non-zero* $\mathbf{x} \in \mathbb{R}^n$ *and associated* σ, *there is an eigenvalue* λ *of a symmetric matrix* A, *which satisfies the inequality*

$$|\lambda - \sigma| \leq \frac{||A\mathbf{x} - \mathbf{x}\sigma||}{||\mathbf{x}||}. \tag{2.6}$$

Proof. The proof is Wilkinson's result [9].

As a corollary of the previous propositions, for symmetric matrices the value of (2.3) provides a suitable bound for the possible difference of a true eigenvalue and the Rayleigh-quotient.

Corollary 2.1. *For any non-zero* $\mathbf{x} \in \mathbb{R}^n$ *and symmetric matrix* A, *if*

$$\frac{\sqrt{f_A^{(w)}(\mathbf{x})}}{||\mathbf{x}||^{2-2w}} \leq \epsilon, \tag{2.7}$$

then an eigenvalue λ *exists, which satisfies* $|\lambda - \sigma| \leq \epsilon$, *where* σ *is the Rayleigh-quotient and* ϵ *is a suitable bound.*

The forthcoming theorem refer to bound for the approximation of the eigenvectors. Let be denoted the eigenvalues of the symmetric matrix A as $\lambda_1, \lambda_2, \ldots, \lambda_n$, the associated normalized eigenvectors as $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ and the angles between the eigenvectors and vector $\mathbf{x}$ as $\alpha_1, \alpha_2, \ldots, \alpha_n$. The least upper bound of the angle between the eigenvector $\mathbf{u}_m$, which is associated with the best approximated eigenvalue λ_m (by Corollary 2.1) and $\mathbf{x}$ provides a measure for the accuracy of approximation.

Proposition 2.6. *For arbitrary non-zero* $\mathbf{x} \in \mathbb{R}^n$ *and real, symmetric matrix* A,

$$\sqrt{f_A^{(w)}(\mathbf{x})}/||\mathbf{x}||^{2-2w} \leq \epsilon \Longrightarrow |\sin\alpha_m| \leq \sqrt{\frac{\epsilon}{\left||\lambda_m|\left(1 - \frac{|\tilde{\lambda}^{(1)}|}{|\lambda_m|}\right)\right|}}, \tag{2.8}$$

where $\lambda_m \neq 0$ *is the eigenvalue nearest to* σ *(Rayleigh quotient) and* $\tilde{\lambda}^{(1)}$ *is the convex linear combination of the complementer part of the spectrum*

$$\tilde{\lambda}^{(1)} = \frac{\sum_{j,j\neq m} \lambda_j \cos^2\alpha_j}{\sum_{j,j\neq m} \cos^2\alpha_j}. \tag{2.9}$$

Proof. Taking into account $|\lambda_m - \sigma| = \min_i |\lambda_i - \sigma|$, the bound $|\lambda_m - \sigma| \leq \epsilon$ and $\cos\alpha_j = \mathbf{u}_j^\top \frac{\mathbf{x}}{||\mathbf{x}||}$,

(2.10)

$$|\lambda_m - \sigma| = \left|\lambda_m - \left(\left(\frac{\mathbf{x}}{||\mathbf{x}||}\right)^\top \sum_j \lambda_j \mathbf{u}_j \mathbf{u}_j^\top \frac{\mathbf{x}}{||\mathbf{x}||}\right)\right| =$$

$$= \left|\lambda_m - \lambda_m \cos^2\alpha_m - \tilde{\lambda}^{(1)} \sum_{j,j\neq m} \cos^2\alpha_j\right| = \left|\left(\lambda_m - \tilde{\lambda}^{(1)}\right)\sin^2\alpha_m\right| \leq \epsilon.$$

The relation

$$\left| |\lambda_m| - \left|\tilde{\lambda}^{(1)}\right| \right| \sin^2 \alpha_m \leq \left|\lambda_m - \tilde{\lambda}^{(1)}\right| \sin^2 \alpha_m \leq \epsilon \tag{2.11}$$

provides a little sharpening of the statement.

The essences of the proposition is that the accuracy of the approximation of an eignvector depends on the bound ϵ, on the magnitude of the associated eigenvalue λ_m, but also on the position-dependent 'effective degeneracy' of the spectrum encoded in the convex linear combination $\tilde{\lambda}^{(1)}$ of the complementer part of the spectrum. This entanglement makes the error estimation for the eigenvectors basically different from the error estimation of the approximation of eigenvalues. If the spectrum is near-degenerate, or the position vector x (and consequently $\cos^2(\alpha_j)$) make it effectively near-degenerate, the bound loses its power. If an estimation for $|\tilde{\lambda}^{(1)}|/|\lambda_m|$ and λ_m is available from some independent information on the spectrum, the relation offers a powerful bound for the accuracy of actual calculations.

3. The optimization of eigenvector-functions

As discussed previously, the non-negative eigenvector-function (2.2) and (2.3) have their zeros at the (real) eigenvectors of matrix A. These local optimum points are simultaneously global optima of these eigenvector-functions and determining the eigenvectors goes back to determining the optimum points of (2.2) or (2.3). At the local optima of the eigenvector-functions the function value is 0 (Lemma 2.3.ii) and a zero function value implies a zero-vector gradient (Lemma 2.3.iii). Since the normalization restricts the function to the unit sphere, it is of importance to investigate the properties of local minima on the unit sphere. The propositions will be stated only for (2.3) with $w = 0$, but analogous statements refer also to other w's and (2.2).

Proposition 3.1. *For non-singular, symmetrical matrix A with non-degenerate spectrum, $\nabla f_A^{(0)}(\mathbf{x})$ and $\mathbf{x}$ are linearly dependent if $\mathbf{x}$ is in the linear space of two eigenvectors,*

$$\mathbf{x} \in \{c_1 \mathbf{u}_i \pm c_2 \mathbf{u}_j \quad c_1, c_2 \in \mathbb{R},\ 1 \leq i, j \leq n\}.$$

Proof. By Lemma 5.1.ii in the Appendix the gradient is in the subspace of three vectors $\mathbf{x}, A\mathbf{x}, A^2\mathbf{x}$, therefore $\mathrm{Rank}(\mathbf{x}, A\mathbf{x}, A^2\mathbf{x}) < 3$. In the space of eigenvectors this is of the form

$$\text{(3.1)} \qquad \mathrm{Rank}\begin{pmatrix} c_1 & \lambda_1 c_1 & \lambda_1^2 c_1 \\ \vdots & \vdots & \vdots \\ c_n & \lambda_n c_n & \lambda_n^2 c_n \end{pmatrix} < 3.$$

Proceeding indirectly, let us assume that $c_i c_j c_k \neq 0$ for a triplet of indices providing $\mathbf{x} = c_1\mathbf{u}_1 + \ldots + c_n\mathbf{u}_n$ with at least three non-zero components. This leads however to a contradiction, because

$$\det\begin{pmatrix} c_i & \lambda_i c_i & \lambda_i^2 c_i \\ c_j & \lambda_j c_j & \lambda_j^2 c_j \\ c_k & \lambda_k c_k & \lambda_k^2 c_k \end{pmatrix} = c_i c_j c_k (\lambda_k - \lambda_i)(\lambda_k - \lambda_j)(\lambda_j - \lambda_i) \neq 0$$

and the rank is 3.

Theorem 3.2. *For the symmetric, non-singular matrix A with a non-degenerate spectrum $\nabla f_A^{(0)}(\mathbf{x})$ and $\mathbf{x}$ are linearly dependent, if and only if (i) or (ii) is valid.*

(i) $\mathbf{x}$ is in the linear space of an eigenvector: $\mathbf{x} \in \{c\mathbf{u}_i \ : \ c \in \mathbb{R},\ 1 \leq i \leq n\}$.

(ii) $\mathbf{x}$ is the bisectrix of two eigenvectors: $\mathbf{x} \in \{c(\mathbf{u}_i \pm \mathbf{u}_j) \quad c \in \mathbb{R},\ 1 \leq i, j \leq n\}$.

Proof. Without restricting the generality it can be assumed by Proposition 3.1 that $\mathbf{x} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2$. An orthogonal vector $\mathbf{d}$ is of the form $\pm(-c_2\mathbf{u}_1 + c_1\mathbf{u}_2 + d_3\mathbf{u}_3 + \ldots + d_n\mathbf{u}_n)$, where $d_3, \ldots, d_n$ are free parameters. It is enough to check the fulfilment of equality $\mathbf{d}^\top \nabla f_A^{(0)}(\mathbf{x}) = 0$. By 5.1.iii in the Appendix

$$\mathbf{d}^\top \nabla f_A^{(0)}(\mathbf{x}) = 2\mathbf{d}^\top \sum_{i=1}^{2}\sum_{j=1}^{2}(\lambda_j - \lambda_i)^2 \mathbf{u}_i c_i c_j^2 = 2c_1c_2(c_1^2 - c_2^2)(\lambda_1 - \lambda_2)^2.$$

Since $\lambda_1 \neq \lambda_2$, if $c_1 = 0$ or $c_2 = 0$, the point is at an eigenvector (i), if $c_1^2 = c_2^2$, the bisectrix of two eigenvectors is obtained (ii). (The (i) part of the proof could be reached also by Lemma 2.3.)

Theorem 3.3. *If matrix A is non-singular, symmetric with a non-degenerate spectrum, then:*

(i) $\mathbf{d}^\top H(f_A^{(0)}(\mathbf{x}))\mathbf{d} \geq 0$ if $\mathbf{d} \perp \mathbf{x}$ and $\mathbf{x}$ is in the linear space of an eigenvector.

(ii) $\mathbf{d}^{\top} H(f_A^{(0)}(\mathbf{x}))\mathbf{d} \leq 0$ *if* $\mathbf{x}$ *is the bisectrix of two eigenvectors,* $\mathbf{d} \perp \mathbf{x}$ *and* $\mathbf{d}$ *is in the linear space of the same two eigenvectors.*

(iii) $\mathbf{d}^{\top} H(f_A^{(0)}(\mathbf{x}))\mathbf{d} \leq 0$ *if* $\mathbf{x}$ *is the bisectrix of two eigenvectors,* $\mathbf{d} \perp \mathbf{x}$ *and* $\mathbf{d}$ *does not contain components in the linear space of the same two eigenvectors.*

(iv) $\mathbf{x}^{\top} H(f_A^{(0)}(\mathbf{x}))\mathbf{x} = 0$ *if* $\mathbf{x}$ *is in the linear space of an eigenvector.*

(v) $\mathbf{x}^{\top} H(f_A^{(0)}(\mathbf{x}))\mathbf{x} > 0$ *if* $\mathbf{x}$ *is not in the linear space of an eigenvector.*

SYMMETRIC	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
n=25	5	27	37	39	40	41	42
n=50	7	53	77	83	85	85	86
n=100	18	102	141	162	163	163	164
n=200	33	140	209	222	227	230	232

Table 1. Convergence data for $f_A^{(w)}(\mathbf{x}),\ w = 0$

Proof. $(i - iii)$ It is assumed again that $\mathbf{x} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2$ and $\mathbf{d} = d_1\mathbf{u}_1 + d_2\mathbf{u}_2 + d_3\mathbf{u}_3 + \ldots + d_n\mathbf{u}_n$. Taking into account Lemma 5.1.v in the Appendix,

$$\mathbf{d}^{\top} H(f_A^{(0)}(\mathbf{x}))\mathbf{d} = \mathbf{d}^{\top} \times \tag{3.2}$$

$$\left[4(\lambda_1 - \lambda_2)^2 c_1 c_2 \left(\mathbf{u}_1\mathbf{u}_2^{\top} + \mathbf{u}_2\mathbf{u}_1^{\top}\right) + \sum_{i=1}^{n} 2((\lambda_i - \lambda_1)^2 c_1^2 + (\lambda_i - \lambda_2)^2 c_2^2)\mathbf{u}_i\mathbf{u}_i^{\top}\right]\mathbf{d}$$

$$= 4(\lambda_1 - \lambda_2)^2 c_1 c_2 d_1 d_2 + \sum_{i=1}^{n} 2d_i^2((\lambda_i - \lambda_1)^2 c_1^2 + (\lambda_i - \lambda_2)^2 c_2^2).$$

If considering the statements under points $(i - iii)$, the special cases of (3.2) imply

$$(i) \qquad \mathbf{d}^{\top} H(f_A^{(0)}(\mathbf{x}))\mathbf{d} \geq 0 \quad \text{if} \quad (c_1 d_2 = 0) \ \text{or} \ (c_2 d_1 = 0).$$

$$(ii) \qquad \mathbf{d}^{\top} H(f_A^{(0)}(\mathbf{x}))\mathbf{d} \leq 0 \quad \text{if} \quad \begin{cases} c_1 = -c_2 = -d_1 = d_2,\ d_3 \ldots d_n = 0, \\ c_1 = c_2 = -d_1 = d_2,\ d_3 \ldots d_n = 0, \\ c_1 = c_2 = d_1 = -d_2,\ d_3 \ldots d_n = 0. \end{cases}$$

$$(iii) \qquad \mathbf{d}^{\top} H(f_A^{(0)}(\mathbf{x}))\mathbf{d} \geq 0 \quad \text{if} \quad (c_1^2 = c_2^2,\ d_1 d_2 = 0).$$

$(iv - v)$ Notice that by points $i.$ and $v.$ of Lemma 5.1 in the Appendix

$$\mathbf{x}^\top H(f_A^{(0)}(\mathbf{x}))\mathbf{x} = 6\sum_{i=1}^{n}\sum_{j=1}^{n}(\lambda_i - \lambda_j)^2 c_i^2 c_j^2 = 6 f_A^{(0)}(\mathbf{x}),$$

which together with point i of Lemma 2.3 provides the proof.

By Propositions 3.2 and 3.3 on the unit sphere no other local minima occur than those associated with eigenvectors. The ridges, which separate the valleys of eigenvectors are at the bisectrices, the quadratics are negative in the direction othogonal to the bisectrix, but positive in such directions, which have only non-zero components in the $(n-2)$-dimensional complementary subspace.

3.1. Results

The results refer to numerical performance tests computed on symmetric and non-symmetric matrices of various sizes with uniformly distributed random elements in the interval $[-1, 1]$. For optimization of the eigenvector functions the BFGS [3, 5, 7] algorithm was used, which was started with a random trial vector and it was terminated (by Corollary 2.1) if condition

$$\frac{\sqrt{f_A^{(w)}(\mathbf{x})}}{||\mathbf{x}||^{2-2w}} < \varepsilon \tag{3.3}$$

was met.

3.1.1. Numerical result for the symmetric case. The lines of the tables are arranged by the sizes of matrices, the columns by the referred accuracies of approximation. The data are the average numbers of iteration steps necessary to reach the displayed accuracies obtained by averaging over 100 random matrices. Tables 1 ($w = 0$), 2 ($w = 0.25$), 3 ($w = 0.5$) and 4 ($w = 0.75$) refer to $f_A^{(w)}(\mathbf{x})$ (2.3), Table 5 refers to $g_A(\mathbf{x})$ (2.2).

SYMMETRIC	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
n=25	5	21	34	39	41	43	43
n=50	13	45	58	62	65	67	68
n=100	24	80	119	132	138	141	143
n=200	41	131	220	249	261	268	269

Table 2. Convergence data for $f_A^{(w)}(\mathbf{x})$, $w = 0.25$

SYMMETRIC	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
n=25	4	17	25	28	29	30	31
n=50	7	30	50	52	54	55	56
n=100	10	46	92	100	103	104	105
n=200	13	69	171	196	200	203	204

Table 3. Convergence data for $f_A^{(w)}(\mathbf{x})$, $w = 0.5$

SYMMETRIC	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
n=25	4	16	25	28	30	31	32
n=50	5	27	47	52	55	57	59
n=100	7	42	94	105	110	112	114
n=200	10	66	171	197	203	207	210

Table 4. Convergence data for $f_A^{(w)}(\mathbf{x})$, $w = 0.75$

SYMMETRIC	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
n=25	3	17	25	27	29	30	30
n=50	5	31	47	52	54	55	56
n=100	8	55	93	102	105	107	107
n=200	12	92	181	200	206	208	209

Table 5. Convergence data for $g_A(\mathbf{x})$

3.1.2. Numerical results for the non-symmetric case. As mentioned previously, because of the homogeneity of $f_A^{(0)}$, the zero-vector is an optimum point. Although in case of symmetric matrices eigenvector-function (2.1) converges only rarely to the zero-vector, this occurs more frequently with non-symmetric matrices. To avoid these pitfalls, for searching non-symmetric matrices the normalized eignvector-function (2.3) should be used. Tables 6 ($w = 0.25$), 7 ($w = 0.5$), 8 ($w = 0.75$) display the formerly specified data obtained with non-symmetric matrices. However the unnormalized eigenvector-function (2.2) proved to be generally 'zero-vector safe' and highly effective also in non-symmetric cases. Table 9 contains the referred data obtained with $g_A(\mathbf{x})$ (2.2).

3.1.3. Discussion of the results. As the tables show, increasing the size of the matrix the convergence speeds up in both of symmetric and non-symmetric cases. Also the degree of homogeneity affects the convergence speed and $w = 0.5$ proved to be the best parameter. In both cases the average numbers of BFGS iteration steps were quite low and it was close to n for symmetric matrices, while about $1.5n$ for non-symmetric matrices (n is the dimension). If using eigenvector-function (2.2) for symmetric matrices the results were similar,

NONSYMMETRIC	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
n=25	3	31	64	71	72	74	74
n=50	12	57	108	113	116	118	119
n=100	15	71	168	186	192	196	198
n=200	25	110	274	318	322	324	325

Table 6. Convergence data for $f_A^{(w)}(\mathbf{x})$, $w = 0.25$

NONSYMMETRIC	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
n=25	3	20	48	57	59	60	61
n=50	4	34	82	112	115	117	119
n=100	6	52	136	172	172	173	173
n=200	9	83	224	289	293	294	295

Table 7. Convergence data for $f_A^{(w)}(\mathbf{x})$, $w = 0.5$

NONSYMMETRIC	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
n=25	3	20	42	48	51	52	53
n=50	4	33	69	80	81	84	85
n=100	6	52	128	147	150	154	154
n=200	9	83	238	288	300	305	307

Table 8. Convergence data for $f_A^{(w)}(\mathbf{x})$, $w = 0.75$

NONSYMMETRIC	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
n=25	3	18	39	40	42	43	44
n=50	4	32	61	67	68	68	69
n=100	7	52	114	128	130	131	132
n=200	10	80	207	228	235	238	239

Table 9. Convergence data for $g_A(\mathbf{x})$

however for non-symmetric matrices (2.2) proved to be significantly better and no clear tendency appeared for converging to the zero-vector. Since the structure of (2.2) is also simpler than that of (2.3), evaluation of the function value, as well as the gradient required less operations and (2.2) seemed to be the most efficient for the selective determination of eigenvectors.

3.1.4. An example for closely packed eigenvalues. Closely packed eigenvalues provide hardly surmountable problems for most of the iterative methods. The example of a 3×3 symmetric matrix illustrates the power of the algorithm in determining eigenvectors belonging to closely packed eigenvalues. The target matrix is constructed in factorized form with eigenvalues 1.002, 1.001, 1, and

$$A_3 = 1.002\mathbf{u}_1\mathbf{u}_1^T + 1.001\mathbf{u}_2\mathbf{u}_2^T + 1\mathbf{u}_3\mathbf{u}_3^T, \tag{3.4}$$

$\mathbf{x}_0$	λ	*steps*	*accurancy*(λ)
$\left[\frac{2}{\sqrt{29}}, -\frac{3}{\sqrt{29}}, \frac{4}{\sqrt{29}}\right]$	1	7	10^{-11}
$\left[\frac{2}{\sqrt{29}}, \frac{3}{\sqrt{29}}, \frac{4}{\sqrt{29}}\right]$	1.001	7	10^{-11}
$\left[\frac{2}{\sqrt{29}}, \frac{3}{\sqrt{29}}, -\frac{4}{\sqrt{29}}\right]$	1.002	8	10^{-9}

Table 10. Convergence data for the approximation of eigenvectors belonging to closely packed eigenvalues

the convergence data for the optimization of (2.2) are listed in Table 10. The first column dispalys the various starting vectors, the second the eigenvalues, the third the necessary steps to reach the required accuracy in the approximation of the eigenvalue, the fourth displays the accuracy of approximation of the eigenvalue. As the table shows, the closely packed eigenvalues did not provide any problem for the searching procedure, which is a valuable property of the method especially, because it is known that the convergence speed of methods based on the QR transformation break down if the ratio of eigenvalues $|\lambda_i|/|\lambda_{i-1}|$ is close to 1.

4. Conclusions

This short paper presents a novel algorithm for determining eigenvectors and eigenvalues of large real matrices. Although complex matrices and vectors are not discussed here but will be dealt with separately, the method can be applied to complex matrices or complex eigenpairs of non-symmetric matrices, as well. Eigenvector-functions have been established with various degree of homogeneity. It was shown that the local optima of these functions are also global optima, which coincide with the eigenvectors of the underlying matrix. Since the eigenvector-functions are well-behaving, the known optimization procedures [3,5,6,7] can efficiently determine their minima, as the numerical investigation with the BFGS algorithm shows. The selection of the approximated eigenvector is independent of the distribution of the eigenvalues. The algorithm behaves well also in the case of closely packed eigenvalues, which is generally a hard nut for most of the methods. The procedure does not assume the storage of the whole matrix in the core and requires only one matrix-vector, vector-vector and some scalar multiplications per step. The presented statements deliver bounds for the accuracies of the obtained approximate eigenvalues and eigenvectors.

Acknowledgements. The reading of the manusript and comments on the topic are gratefully acknowledged to Professors János Csirik and László Stachó.

References

[1] **Francis J.G.F.**, The QR transformation - A unitary analogue to LR transformation, *Comput. J.*, **4** (1961), 265-271, 332-345.

[2] **Young D.M. and Gregory R.T.**, *A survey of numerical mathematics I-II.*, Addison-Wesley, Reading, 1972.

[3] **Gill P.E., Murray W. and Pitfield P.A.**, *The implementation of two revised quasi-Newton algorithms for unconstrained optimization*, Report NAC-11, National Physical Lab., 1972.

[4] **Lánczos C.**, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, *J. Res. Nat. Bur. Standards.*, **45** (1950), 255-282.

[5] **Mokhtar S. Bazaraa, Hanif D. Sherali and Shetty C.M.**, *Nonlinear programming theory and algorithms*, John Wiley & Sons, 1993.

[6] **Nazareth J.L.**, Conjugate gradient methods less dependency on conjugacy, *SIAM Review*, **28** (4) (1986), 501-511.

[7] **Nocedal J.**, The performance of several algorithms for large scale unconstrained optimization, *Large-scale numerical optimization*, eds. T.F. Coleman and Y. Li, SIAM, Philadelphia, 1990, 138-151.

[8] **Parlett B.**, *The symmetric eigenvalue problem*, Prentice-Hall, Englewood Cliffs, NJ, 1980.

[9] **Wilkinson J.H.**, *The algebraic eigenvalue problem*, Oxford University Press, Oxford, 1965.

5. Appendix

Some technical points are collected into the following lemma, which are referred in the text. The eigenvalues of the real symmetric matrix A are denoted $\lambda_1, \ldots, \lambda_n$ the corresponding, orthonormal eigenvectors, $\mathbf{u}_1, \ldots, \mathbf{u}_n$.

Lemma 5.1. *For the symmetric matrix A and $\mathbf{x} = c_1\mathbf{u}_1 + \ldots + c_n\mathbf{u}_n$,*

$$(i) \qquad f_A^{(0)}(\mathbf{x}) = \sum_{i=1}^{n}\sum_{j=1}^{n}(\lambda_i - \lambda_j)^2 c_i^2 c_j^2,$$

$$(ii) \qquad \nabla(f_A^{(0)}(\mathbf{x})) = 2(\mathbf{x}^\top\mathbf{x}A^2\mathbf{x} + \mathbf{x}(A\mathbf{x})^\top A\mathbf{x} - 2\mathbf{x}^\top A\mathbf{x}(A\mathbf{x})),$$

$$(iii) \qquad \nabla(f_A^{(0)}(\mathbf{x})) = 2\sum_{i=1}^{n}\sum_{j=1}^{n}(\lambda_j - \lambda_i)^2 u_i c_i c_j^2,$$

$$(iv) \qquad H(f_A^{(0)}(\mathbf{x})) = 2\mathbf{x}^\top\mathbf{x}A^2 + 4\mathbf{x}\mathbf{x}^\top A^2 + A^2\mathbf{x}\mathbf{x}^\top + 2I(A\mathbf{x})^\top A\mathbf{x} - \\ - 4\mathbf{x}^\top A\mathbf{x}A - 8A\mathbf{x}(A\mathbf{x})^\top,$$

$v.$ $$H(f_A^{(0)}(\mathbf{x})) = \sum_{i=1}^{n}\sum_{j=1}^{n}(\lambda_i - \lambda_j)^2(2c_j^2\mathbf{u}_i\mathbf{u}_i^\top + 4c_ic_j\mathbf{u}_i\mathbf{u}_j^\top).$$

Proof. (i) The statement follows by a direct evaluation and subsequent rearrangement of the terms,

$$f_A^{(0)}(\mathbf{x}) = \mathbf{x}^\top\mathbf{x}(A\mathbf{x})^\top(A\mathbf{x}) - (\mathbf{x}^\top A\mathbf{x})^2 = \left(\sum_{i=1}^{n} c_i^2\right)\left(\sum_{j=1}^{n}\lambda_j^2 c_j^2\right) - \left(\sum_{i=1}^{n}\lambda_i c_i^2\right)^2 =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n}(\lambda_i - \lambda_j)^2 c_i^2 c_j^2.$$

(ii) By a direct evaluation the terms are

$$\begin{aligned}\nabla f_A^{(0)}(\mathbf{x}) &= \nabla[\mathbf{x}^\top\mathbf{x}(A\mathbf{x})^\top(A\mathbf{x})] - \nabla(\mathbf{x}^\top A\mathbf{x})^2,\\ \nabla[\mathbf{x}^\top\mathbf{x}(A\mathbf{x})^\top(A\mathbf{x})] &= 2(\mathbf{x}^\top\mathbf{x}A^\top(A\mathbf{x}) + \mathbf{x}(A\mathbf{x})^\top(A\mathbf{x})),\\ \nabla(\mathbf{x}^\top A\mathbf{x})^2 &= 2(\mathbf{x}^\top A\mathbf{x})(A^\top\mathbf{x} + A\mathbf{x}).\end{aligned}$$

(iii) The statement follows by a direct evaluation and subsequent rearrangement of the terms,

$$1/2\nabla(f_A^{(0)}(\mathbf{x})) = (\mathbf{x}^\top\mathbf{x})(A^2\mathbf{x}) + (\mathbf{x})((A\mathbf{x})^\top A\mathbf{x}) - 2(\mathbf{x}^\top A\mathbf{x})(A\mathbf{x}) =$$

$$= \left(\sum_{j=1}^{n} c_j^2\right)\left(\sum_{i=1}^{n}\lambda_i^2 c_i\mathbf{u}_i\right) + \left(\sum_{i=1}^{n} c_i\mathbf{u}_i\right)\left(\sum_{j=1}^{n}\lambda_j^2 c_j^2\right) -$$

$$-2\left(\sum_{j=1}^{n}\lambda_j c_j^2\right)\left(\sum_{i=1}^{n}\lambda_i c_i\mathbf{u}_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}(\lambda_j - \lambda_i)^2\mathbf{u}_i c_i c_j^2.$$

(iv) To avoid unnecessary technicalities, only two steps of the evaluation are presented,

$$\begin{aligned}H(\mathbf{x}^\top\mathbf{x}(A\mathbf{x})^\top A\mathbf{x}) &= 2\mathbf{x}^\top\mathbf{x}A^2 + 4\mathbf{x}\mathbf{x}^\top A^2 + 4A^2\mathbf{x}\mathbf{x}^\top + 2I(A\mathbf{x})^\top A\mathbf{x},\\ H((\mathbf{x}^\top A\mathbf{x})^2) &= 4\mathbf{x}^\top A\mathbf{x}A + 8A\mathbf{x}(A\mathbf{x})^\top.\end{aligned}$$

(v) The proof is obtained by simple algebraic rearrangements.

(Received January 10, 2000)

A. Kocsor
Research Group on Artificial Intelligence
József Attila University
Árpád t. 2.
H-6720 Szeged, Hungary
kocsor@inf.u-szeged.hu

J. Dombi
Dept. of Applied Informatics
József Attila University
Árpád t. 2.
H-6720 Szeged, Hungary
dombi@inf.u-szeged.hu

I. Bálint
Dept. of Theoretical Physics
József Attila University
Tisza L. krt. 84-86.
H-6720 Szeged, Hungary

Dept. of Pharmaceutical Analysis
Szent-Györgyi Albert Med. Univ.
Somogyi u. 4.
H-6720 Szeged, Hungary

Dept. of Natural Sciences
Dunaújváros Polytechnic
Táncsics M. u. 1.
H-2400 Dunaújváros, Hungary
ibalint@physx.u-szeged.hu