Electronic Transactions on Numerical Analysis.
Volume 31, pp. 295-305, 2008.

ISSN 1068-9613.

ETNA
Kent State University
http://etna.math.kent.edu

A COUNTEREXAMPLE FOR CHARACTERIZING AN INVARIANT SUBSPACE OF A MATRIX*

HUBERT SCHWETLICK† AND KATHRIN SCHREIBER‡

Abstract. As an alternative to Newton's method for computing a simple eigenvalue and corresponding eigenvectors of a nonnormal matrix in a stable way, an approach based on singularity theory has been proposed by Schwetlick/Lösche [Z. Angew. Math. Mech., 80 (2000), pp. 9–25]. In this paper, by constructing a counterexample with a singular linear block operator, it is shown that a straightforward extension of this technique to the computation of invariant subspaces of dimension $p > 1$ will not work, in general. Finding this counterexample required a detailed study of the linear block operator.

Key words. Eigenvalue problem, simple invariant subspace, block Newton method, block Rayleigh quotient iteration.

AMS subject classifications. 65F15.

1. Introduction.

Consider the eigenvalue problem

$$Ax = \lambda x \tag{1.1}$$

with an arbitrary matrix $A \in \mathbb{C}^{n\times n}$ and suppose that $\lambda_1 \in \mathbb{C}$ is an algebraically simple eigenvalue with normalized right and left eigenvectors x_1 and y_1, resp., i.e., $\ker(A-\lambda_1 I) = \operatorname{span}\{x_1\}$, $\ker(A-\lambda_1 I)^H = \operatorname{span}\{y_1\}$, $\|x_1\| = \|y_1\| = 1$, $y_1^H x_1 \neq 0$. Here $\|\cdot\|$ denotes the Euclidean norm, and the spectral norm in case of matrices. In order to compute the eigenpair (x_1, λ_1), the normalization condition $w^H x = 1$ is added to the invariance condition (1.1) where $w \in \mathbb{C}^n$ is a normalizing vector with $\|w\| = 1$ and $w^H x_1 \neq 0$. This leads to the extended system

$$F_w(x,\lambda) = \begin{bmatrix} (A-\lambda I)x \\ w^H x - 1 \end{bmatrix} = 0, \tag{1.2}$$

which is solved by the pair $(x_*, \lambda_*) = (x_1/(w^H x_1), \lambda_1)$. The Jacobian

$$\partial F_w(x,\lambda) = \begin{bmatrix} A-\lambda I & -x \\ w^H & 0 \end{bmatrix} \tag{1.3}$$

is nonsingular at the solution if and only if λ_1 is algebraically simple. Hence, under the assumption that λ_1 is simple, Newton's method can be applied to (1.2).

Typically the basic Newton step (cf. [16]) is modified in that only the normalized new x-part is used, whereas the new λ-part is computed as Rayleigh quotient [11]. This yields the locally quadratically, for Hermitian A even cubically, convergent Rayleigh quotient iteration; cf. [11, 12]. In the limit, with the optimal normalizing vector $w = x_1$, the inverse Jacobian (1.3) is bounded from below as follows

$$\|\partial F_{x_1}(x_1,\lambda_1)^{-1}\| = \left\| \begin{bmatrix} A-\lambda_1 I & -x_1 \\ x_1^H & 0 \end{bmatrix}^{-1} \right\| \geq \frac{1}{|y_1^H x_1|};$$

*Received November 30, 2007. Accepted January 8, 2009. Published online on June 8, 2009. Recommended by Daniel Kressner.

†Technische Universität Dresden, Fakultät Mathematik und Naturwissenschaften, Institut für Numerische Mathematik, 01062 Dresden, Germany (hubert.schwetlick@tu-dresden.de).

‡Technische Universität Berlin, Institut für Mathematik, MA 4-5, Straße des 17. Juni 136, 10623 Berlin, Germany (schreibe@math.tu-berlin.de). The research of this author is supported by the DFG Research Center MATHEON in Berlin.

ETNA
Kent State University
http://etna.math.kent.edu

cf. [13]. Hence, if x_1 and y_1 are almost orthogonal, i.e., if λ_1 is strongly ill-conditioned (which can occur if A is strongly nonnormal), then the norm of the inverse Jacobian may be arbitrarily large. The reason lies in the bordering vector $-x_1$ in the upper block, which is then almost orthogonal to the missing direction $\operatorname{span}\{y_1\}$ in the range $\operatorname{Im}(A - \lambda_1 I) = \operatorname{span}\{y_1\}^\perp$.

In order to circumvent this possible growth of the inverse Jacobian, in [13] an alternative approach has been introduced using techniques from singularity theory of nonlinear equations. There, the eigenvalue λ_1 is characterized by the following system

$$C(\lambda, u, v)\begin{bmatrix} x \\ \mu \end{bmatrix} = \begin{bmatrix} A - \lambda I & v \\ u^H & 0 \end{bmatrix}\begin{bmatrix} x \\ \mu \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \Leftrightarrow \begin{cases} (A - \lambda I)x + v\mu = 0, \\ \qquad\qquad u^H x = 1, \end{cases} \tag{1.4}$$

where u, v are normalized approximations to x_1, y_1, respectively. Since $C(\lambda_1, u, v)$ is nonsingular if $x_1^H u \neq 0$ and $y_1^H v \neq 0$ [13], it is also nonsingular for λ close to λ_1. Hence, for such λ, the system (1.4) uniquely defines $x = x(\lambda)$, $\mu = \mu(\lambda)$ as functions of λ. Moreover, we have $\mu(\lambda_1) = 0$ and $x(\lambda_1) = x_1/(u^H x_1)$. Applying one Newton step $\theta \mapsto \theta_+ = \theta - \mu(\theta)/\dot{\mu}(\theta)$ to $\mu(\lambda) = 0$ for given u, v from the current approximation $\lambda = \theta$, yields the generalized Rayleigh quotient

$$\theta_+ = \frac{y(\theta)^H A x(\theta)}{y(\theta)^H x(\theta)}$$

(cf. [13]), where $x = x(\theta)$ comes from the primal system (1.4), whereas $y = y(\theta)$ is part of the solution of the dual system

$$C(\lambda, u, v)^H\begin{bmatrix} y \\ \nu \end{bmatrix} = \begin{bmatrix} (A - \lambda I)^H & u \\ v^H & 0 \end{bmatrix}\begin{bmatrix} y \\ \nu \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \Leftrightarrow \begin{cases} (A - \lambda I)^H y + u\nu = 0, \\ \qquad\qquad v^H y = 1, \end{cases}$$

which defines $y = y(\lambda)$, $\nu = \nu(\lambda)$ as functions of λ. Note that $\mu(\lambda) \equiv \bar{\nu}(\lambda)$, hence $\nu(\lambda_1) = 0$ and $y(\lambda_1) = y_1/v^H y_1$. New bordering right and left eigenvector approximations are given by $u_+ = x(\theta)/\|x(\theta)\|$, $v_+ = y(\theta)/\|y(\theta)\|$. This generalized Rayleigh quotient iteration (GRQI), which is similar to Parlett's alternating RQI [12], converges quadratically; cf. [13].

The so defined matrix $C(\lambda, u, v)$ is independent of the conditioning of the eigenvalue, since we have in the limit

$$\left\|C(\lambda_1, x_1, y_1)^{-1}\right\| = \left\|\begin{bmatrix} A - \lambda_1 I & y_1 \\ x_1^H & 0 \end{bmatrix}^{-1}\right\| = \max\left\{\frac{1}{\sigma_{n-1}(A - \lambda_1 I)}, 1\right\},$$

where $\sigma_{n-1}(A - \lambda_1 I)$ denotes the smallest positive singular value of $A - \lambda_1 I$. This result can be derived by forming a unitary decomposition of $C(\lambda_1, x_1, y_1)$ using the singular value decomposition

$$A - \lambda_1 I = [\,Y_1\, y_1\,]\begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix}[\,X_1\, x_1\,]^H,$$

where $\Sigma_1 = \operatorname{diag}(\sigma_1(A - \lambda_1 I), \ldots, \sigma_{n-1}(A - \lambda_1 I))$ is nonsingular. Hence, the norm does not grow like $1/|y_1^H x_1|$ as in the direct Newton approach. In this sense the matrices of the alternative method are optimally bordered. When, for large scale problems, the arising linear systems are solved by Krylov subspace methods, then in case of very ill-conditioned eigenvalues, the better behavior of C^{-1} compared to ∂F_u^{-1} leads to a better performance of the alternative approach.

ETNA
Kent State University
http://etna.math.kent.edu

If λ_1 is not simple, then both approaches lead to systems which are singular at the solution, and if λ_1 belongs to a cluster the inverses will have large norms. In these cases, block methods which determine invariant subspaces belonging to a subset of the spectrum $\lambda(A)$ are the methods of choice. Such block methods for computing invariant subspaces have been considered already in [5, 7, 14], and all three are discussed in [6]. Whereas the standard Newton approach can be generalized in a straightforward way to the block case (cf. [4, 8, 10] and [2, 1] for block Rayleigh quotient iterations on Grassmann manifolds) it was not clear whether this is possible for the alternative, singularity theory based approach. In this paper we show that this is not the case by constructing a counterexample where the block system is singular at the solution, so it does not uniquely define the singularity function.

The next subsection gives a short review on the standard block Newton approach, Section 2 introduces the block formulation of the alternative ansatz (1.4). Section 3 characterizes the so defined operator and its matrix representation. Section 4 develops conditions on the nonsingularity of the operator, and Section 5 provides the counterexample, i.e., a singular operator.

In what follows the spectrum of M will be denoted by $\lambda(M)$.

1.1. The standard block Newton approach. We start with some facts and notation from [15]. Recall that finding an invariant subspace $\mathcal{X} = \operatorname{Im} X$ of dimension p to A means finding a matrix $X \in \mathbb{C}^{n\times p}$ with $\operatorname{rank} X = p$ and a square matrix $L \in \mathbb{C}^{p\times p}$ such that

$$AX = XL. \tag{1.5}$$

Now let $X = X_1$ with $X_1^H X_1 = I_p$ define a right invariant subspace $\operatorname{Im} X_1$ of A, i.e., $AX_1 = X_1 L_1$, hence, $L = L_1 = X_1^H A X_1$. If $Y_2 \in \mathbb{C}^{n\times q}$, $p + q = n$, is chosen such that $[X_1\ Y_2]$ is unitary, then we obtain the block Schur decomposition

$$A = [X_1\ Y_2] \begin{bmatrix} L_1 & H \\ 0 & L_2 \end{bmatrix} [X_1\ Y_2]^H. \tag{1.6}$$

Obviously, we have $\lambda(L_1) \cup \lambda(L_2) = \lambda(A)$. If $\lambda(L_1) \cap \lambda(L_2) = \emptyset$, then $\operatorname{Im} X_1$ is called a *simple* invariant subspace, and $\operatorname{Im} X_1$ is simple if and only if the linear mapping

$$T = T(L_1, L_2):\ Z \in \mathbb{C}^{p\times q} \longmapsto T[Z] = L_1 Z - Z L_2 \in \mathbb{C}^{p\times q}$$

is nonsingular. In what follows we suppose that $\operatorname{Im} X_1$ is a simple invariant subspace.

A normalization condition $W^H X = I_p$, with $W \in \mathbb{C}^{n\times p}$ and $W^H W = I_p$, in the direct block Newton approach is added to the invariance condition (1.5), which yields the extended block system

$$F_W(X, L) = \begin{bmatrix} AX - XL \\ W^H X - I_p \end{bmatrix} = \begin{bmatrix} 0_{n\times p} \\ 0_{p\times p} \end{bmatrix} \tag{1.7}$$

for (X, L) as the natural generalization of (1.2). If $X_1^H W \in \mathbb{C}^{p\times p}$ is nonsingular, which is equivalent to $\measuredangle(\operatorname{Im} W, \operatorname{Im} X_1) < \pi/2$, then (1.7) is solved by the pair

$$(X_*, L_*) = (X_1\Phi, \Phi^{-1} L_1 \Phi),$$

where $\Phi = (X_1^H W)^{-H}$. The derivative of F_W is given by

$$\partial F_W(X, L)[S, M] = \begin{bmatrix} AS - SL - XM \\ W^H S \end{bmatrix}.$$

One can show that $\partial F_W(X_*, L_*)$ is a nonsingular linear operator on $\mathbb{C}^{n\times p} \times \mathbb{C}^{p\times p}$ if and only if $\operatorname{Im} X_1$ is simple; cf., for instance, [4]. In this case Newton's method can be applied. The Newton step $(U, \Theta) \mapsto (U_+, \Theta_+)$, where $U_+ = U + S$, $\Theta_+ = \Theta + M$ with $W = U$, is defined by the linearized block system $F_U(U, \Theta) + \partial F_U(U, \Theta)[S, M] = 0$, i.e.,

$$\begin{bmatrix} AS - S\Theta - UM \\ U^H S \end{bmatrix} = - \begin{bmatrix} AU - U\Theta \\ 0 \end{bmatrix}, \tag{1.8}$$

for the Newton correction (S, M). The upper block of this system is a Sylvester equation. In order to simplify the solution of this system, the plain Newton step is modified in that (U, Θ) is chosen as a Schur pair, i.e., such that $\Theta = U^H AU$ is upper triangular (diagonal in the Hermitian case), so that the Bartels-Stewart algorithm [3] can be applied. The basis U_+ obtained by Newton's method is then orthonormalized by the modified Gram-Schmidt process, $U_+ = \mathtt{mgs}(U_+)$, and undergoes a Rayleigh-Schur process (Rayleigh-Ritz in the Hermitian case) to deliver the final Schur (Ritz) pair $(U_+, \Theta_+) = \mathtt{rs}(U_+)$; for details see [10] for the case $A = A^T$, which extends readily to $A = A^H$, [4] for general A, and [8] for the real symmetric case.

2. The generalized block system. In this section we want to extend the alternative system (1.4) for the characterization of an invariant subspace. The straightforward block generalization of (1.4) yields

$$C(L, U, V)[X, M] = \begin{bmatrix} AX - XL + VM \\ U^H X \end{bmatrix} = \begin{bmatrix} 0_{n\times p} \\ I_p \end{bmatrix}, \tag{2.1}$$

with given $L \in \mathbb{C}^{p\times p}$ and bordering matrices $U, V \in \mathbb{C}^{n\times p}$, with $U^H U = V^H V = I_p$, and unknowns $X \in \mathbb{C}^{n\times p}$, $M \in \mathbb{C}^{p\times p}$. The linear operator $C(L, U, V)$ of (2.1) differs from the operator in (1.8) in that the matrix $-U$ in the upper block is replaced by V. If, for a certain $L = L_*$, there is a solution $(X, M) = (X_*, M_*)$ with $M_* = 0$, then $X = X_*$ satisfies

$$R(L_*)[X] = AX - XL_* = 0, \tag{2.2}$$

i.e., X_* then defines an invariant subspace $\mathcal{X} = \operatorname{Im} X_*$ of A with spectrum $\lambda(L_*)$. Due to $U^H X_* = I_p$, this subspace has dimension p.

Under the assumption that $X_1^H U \in \mathbb{C}^{p\times p}$ is nonsingular, where $\operatorname{Im} X_1$ is the simple invariant subspace we are looking for, the block singularity system has the solution $(X_*, M_*) = (X_1\Phi, 0)$, where $\Phi = (X_1^H U)^{-H}$ and $L = L_* = \Phi^{-1} L_1 \Phi$. Thus, if the linear operator $C(L_*, U, V)$ defined in (2.1) is nonsingular, then this solution is unique, and for L close to L_* system (2.1) uniquely defines $X = X(L)$, $M = M(L)$ as functions of L, where $X(L_*) = X_*$, $M(L_*) = 0$. In this case, the Newton method can be applied to the equation $M(L) = 0$ in order to find its solution L_*, where the p^2-dimensional singularity function $M : L \in \mathbb{C}^{p\times p} \mapsto M(L) \in \mathbb{C}^{p\times p}$ is implicitly defined by the block system (2.1).

3. The operator $C(L, U, V)$. We want to have a closer look at the operator $C(L, U, V)$ defined in (2.1) at the solution, and start with characterizing the corresponding invariant subspaces. Following [15] we construct a similarity transformation which brings A to block diagonal form $\operatorname{diag}(L_1, L_2)$ with L_1, L_2 from the Schur decomposition (1.6), i.e.,

$$A = [X_1\ X_2] \begin{bmatrix} L_1 & 0 \\ 0 & L_2 \end{bmatrix} [X_1\ X_2]^{-1} \ \Leftrightarrow\ A^H = [Y_1\ Y_2] \begin{bmatrix} L_1^H & 0 \\ 0 & L_2^H \end{bmatrix} [Y_1\ Y_2]^{-1},$$

where $Y_1 = X_1 - Y_2 Q^H$, $X_2 = Y_2 + X_1 Q$, such that $[X_1\ X_2]$ is nonsingular and $[X_1\ X_2] = [Y_1\ Y_2]^{-H}$. It is easy to see that the block diagonal form is achieved if and only if Q satisfies

$$T(L_1, L_2)[Q] = L_1 Q - Q L_2 = -H = -X_1^H A Y_2.$$

ETNA
Kent State University
http://etna.math.kent.edu

Recall that this equation is uniquely solvable, since T is nonsingular due to the simplicity of $\operatorname{Im} X_1$.

We also need a block Schur decomposition

$$A^H = [\hat{Y}_1\ \hat{X}_2] \begin{bmatrix} \hat{L}_1^H & \hat{H}^H \\ 0 & \hat{L}_2^H \end{bmatrix} [\hat{Y}_1\ \hat{X}_2]^H, \tag{3.1}$$

of A^H, with unitary $[\hat{Y}_1\ \hat{X}_2]$. Some straightforward computations give

$$\begin{aligned} \hat{Y}_1 &= Y_1 D_1^{-1/2} = (X_1 - Y_2 Q^H) D_1^{-1/2}, &\qquad \hat{L}_1 &= D_1^{-1/2} L_1 D_1^{1/2}, \\ \hat{X}_2 &= X_2 D_2^{-1/2} = (Y_2 + X_1 Q) D_2^{-1/2}, &\qquad \hat{L}_2 &= D_2^{1/2} L_2 D_2^{-1/2}, \end{aligned} \tag{3.2}$$

and $\hat{H} = D_2^{-1/2} Q^H L_1 D_1^{1/2} - D_2^{1/2} L_2 Q^H D_1^{-1/2}$, where $D_1 = I + QQ^H$, $D_2 = I + Q^H Q$.

We can now start investigating the operator C at the solution $L_* = \Phi^{-1} L_1 \Phi$, where we assume that U, V satisfy $U^H U = V^H V = I_p$ as above and, additionally, that $X_1^H U$ and $\hat{Y}_1^H V$ are nonsingular. More precisely, we wish to know whether the system

$$C(L_*, U, V)[X, M] = \begin{bmatrix} AX - XL_* + VM \\ U^H X \end{bmatrix} = \begin{bmatrix} 0 \\ I_p \end{bmatrix} \tag{3.3}$$

has other solutions than $(X_*, M_*) = (X_1\Phi, 0)$, $\Phi = (X_1^H U)^{-H}$, i.e., whether the operator $C(L_*, U, V)$ is nonsingular. Recall that the mapping $[X, M] \mapsto C(L, U, V)[X, M]$ is linear in the arguments.

3.1. The mapping $X \mapsto R(L_*)[X] = AX - XL_*$. In order to gain some insight into (3.3) we want to examine the nullspace of $R(L_*)$ and its dimension, i.e., we wish to find all solutions X of the Sylvester equation (2.2). Considering $L_* = \Phi^{-1} L_1 \Phi$, we obtain

$$R = R(L_*)[X] = AX - X\Phi^{-1} L_1 \Phi = \{A(X\Phi^{-1}) - (X\Phi^{-1}) L_1\}\Phi = 0.$$

Writing $X\Phi^{-1}$ as

$$X\Phi^{-1} = [X_1\ X_2][X_1\ X_2]^{-1} X\Phi^{-1} = [X_1\ X_2][Y_1\ Y_2]^H X\Phi^{-1} = X_1 B + X_2 Z_2$$

with $B = Y_1^H X\Phi^{-1}$, $Z_2 = Y_2^H X\Phi^{-1}$, and considering $AX_j = X_j L_j$, $j = 1, 2$, we end up with

$$R\Phi^{-1} = A(X_1 B + X_2 Z_2) - (X_1 B + X_2 Z_2) L_1 = [X_1\ X_2] \begin{bmatrix} L_1 B - B L_1 \\ L_2 Z_2 - Z_2 L_1 \end{bmatrix} = 0.$$

Thus, we have $R = 0$ if and only if $L_1 B - B L_1 = 0$ and $L_2 Z_2 - Z_2 L_1 = 0$. Since $\operatorname{Im} X_1$ is simple, the latter condition implies $Z_2 = 0$. Hence, the kernel is given by

$$\ker R(L_*) = \{X :\ AX - XL_* = 0\} = \{X = X_1 B \Phi :\ L_1 B - B L_1 = 0\}. \tag{3.4}$$

However, the dimension $\hat{p} := \dim \ker R(L_*)$ seems to be less clear. Using the Jordan decompositions of L_1 we obtain the bounds $p \le \hat{p} \le p^2$, where $\hat{p} = p$ if L_1 has p different eigenvalues which is the generic case, and $\hat{p} = p^2$ if $L_1 = \lambda_1 I_p$. In the latter case $L_1 B - B L_1 = 0$ is solved by any $B \in \mathbb{C}^{p \times p}$. Gantmacher [9] gives an exact formula for $\hat{p}$, namely

$$\hat{p} = \sum_{\alpha=1}^{u} \sum_{\beta=1}^{u} \delta_{\alpha\beta},$$

when L_1 has the elementary divisors $(\lambda-\lambda_1)^{s_1},\ (\lambda-\lambda_2)^{s_2},\ \ldots,(\lambda-\lambda_u)^{s_u}$, where $s_1+s_2+\ldots+s_u=p$, and $\delta_{\alpha\beta}$ is defined as the degree of the greatest common divisor of the polynomials $(\lambda-\lambda_\alpha)^{s_\alpha}$ and $(\lambda-\lambda_\beta)^{s_\beta}$.

With $\hat{p}$ as above we obtain

$$\ker R(L_*)=\{X:\ AX-XL_*=0\}=\operatorname{span}\{X^{(1)},\ldots,X^{(\hat{p})}\},$$

with $X^{(i)}=X_1B_i\Phi$, and $\{B_i\}_{i=1}^{\hat{p}}$ are linearly independent solutions of $L_1B-BL_1=0$. Note that $B_1=I_p$ is always a solution and defines $X_*=X_1\Phi$.

3.2. Unitary decomposition. For the analysis of the operator $C(L_*,U,V)$ it turned out to be more promising to treat system (3.3) in terms of vector notation using Kronecker products. This will lead to an SVD-like decomposition. As usual, for a matrix $A=(a_{ij})\in\mathbb{C}^{m\times n}$, the vector $\operatorname{vec}(A)=(a_{11},\ldots,a_{m1},a_{12},\ldots,a_{1n},\ldots,a_{mn})^T\in\mathbb{C}^{mn}$ contains the entries of A in columnwise order.

Writing $\operatorname{vec}(R(L_*)[X])=\mathcal{R}(L_*)\cdot\operatorname{vec}(X)$, $\operatorname{vec}(VM)=\mathcal{V}\cdot\operatorname{vec}(M)$, $\operatorname{vec}(U^HX)=\mathcal{U}^H\cdot\operatorname{vec}(X)$, where $\mathcal{R}(L_*)=I_p\otimes A-L_*^T\otimes I_n$, $\mathcal{V}=I_p\otimes V$, $\mathcal{U}^H=I_p\otimes U^H$, equation (3.3) can equivalently be written as the standard linear system

$$\begin{bmatrix}\mathcal{R}(L_*)\operatorname{vec}(X)+\mathcal{V}\operatorname{vec}(M)\\ \mathcal{U}^H\operatorname{vec}(X)\end{bmatrix}=\begin{bmatrix}\mathcal{R}(L_*) & \mathcal{V}\\ \mathcal{U}^H & 0\end{bmatrix}\begin{bmatrix}\operatorname{vec}(X)\\ \operatorname{vec}(M)\end{bmatrix}=\begin{bmatrix}\operatorname{vec}(0)\\ \operatorname{vec}(I)\end{bmatrix}.$$

Actually, the matrix

$$\mathcal{C}(L_*)=\mathcal{C}(L_*,U,V)=\begin{bmatrix}\mathcal{R}(L_*) & \mathcal{V}\\ \mathcal{U}^H & 0\end{bmatrix}\in\mathbb{C}^{(np+p^2)\times(np+p^2)}\tag{3.5}$$

forms the matrix representation of the operator $C(L_*,U,V)$, and it has the same structure as the matrix in system (1.4) which characterizes the algorithm GRQI. Thus, the skew non-Hermitian bordering of the eigenvalue problem designed to yield optimal condition numbers is transferred to the block case.

We are interested in a unitary block SVD-like reduction of the large matrix $\mathcal{C}$. Therefore, from now on we choose $\{X^{(i)}\}_{i=1}^{\hat{p}}$ with $X^{(i)}=X_1B_i\Phi$ as orthonormal basis of the kernel of $R(L_*)$ with respect to the scalar product in the $\operatorname{vec}(X)$-space, i.e., we choose B_i such that $\operatorname{vec}(X^{(i)})^H\operatorname{vec}(X^{(j)})=\operatorname{trace}(\Phi^HB_i^HB_j\Phi)=\delta_{i,j}\ (i,j=1,\ldots,\hat{p})$.

It is easy to verify that a basis for $\ker R(L_*)^H$, where $R(L_*)^H$ is defined by

$$R(L_*)^H[Y]=A^HY-YL_*^H,$$

is given by $\{\tilde{Y}^{(i)}\}_{i=1}^{\hat{p}}$, with $\tilde{Y}^{(i)}=\hat{Y}_1D_1^{1/2}B_i^H\Phi^{-H}$; cf. (3.1) and (3.2). The matrix representation of this operator is just $\mathcal{R}(L_*)^H=I_p\otimes A^H-\bar{L}_*\otimes I_n$. Then one can find $C_i\in\mathbb{C}^{p\times p}$ such that $Y^{(i)}=\{Y_1D_1^{1/2}C_i^H\Phi^{-H}\}_{i=1}^{\hat{p}}$ is an orthonormal basis of this nullspace. Finally, the $\hat{p}$-dimensional bases constructed above can be extended such that $\{X^{(i)}=X_1B_i\Phi\}_{i=1}^{p^2}$ and $\{Y^{(i)}=Y_1D_1^{1/2}C_i^H\Phi^{-H}\}_{i=1}^{p^2}$ are orthonormal bases of the p^2-dimensional linear spaces

$$\{X=X_1B\Phi:B\in\mathbb{C}^{p\times p}\},\quad\{Y=\hat{Y}_1D_1^{1/2}C^H\Phi^{-H}:C\in\mathbb{C}^{p\times p}\},$$

respectively.

Let $f^{(i)}=\operatorname{vec}(B_i\Phi)$, $h^{(i)}=\operatorname{vec}(D_1^{1/2}C_i^H\Phi^{-H})$. Then, introducing the vectors

$$x^{(i)}=\operatorname{vec}(X^{(i)})=(I\otimes X_1)f^{(i)},\qquad y^{(i)}=\operatorname{vec}(Y^{(i)})=(I\otimes\hat{Y}_1)h^{(i)},\tag{3.6}$$

ETNA
Kent State University
http://etna.math.kent.edu

we finally define the matrices $\tilde{X} = \left[\tilde{X}_1 \,|\, \tilde{X}_{21} \,|\, \tilde{X}_{22}\right]$, $\tilde{Y} = \left[\tilde{Y}_1 \,|\, \tilde{Y}_{21} \,|\, \tilde{Y}_{22}\right]$, where

$$\tilde{X} = \left[x^{(np)} \dots x^{(p^2+1)} \,|\, x^{(p^2)} \dots x^{(\hat{p}+1)} \,|\, x^{(\hat{p})} \dots x^{(1)}\right],$$

and $\tilde{Y}$ is decomposed analogously. Here, $\tilde{X}_1$ and $\tilde{Y}_1$ are chosen such that $\tilde{X}$, $\tilde{Y}$ are unitary. Note that $\ker \mathcal{R}(L_*) = \operatorname{Im}(\tilde{X}_{22})$, $\ker \mathcal{R}(L_*)^H = \operatorname{Im}(\tilde{Y}_{22})$, i.e., $\mathcal{R}\tilde{X}_{22} = 0$ and $\mathcal{R}^H\tilde{Y}_{22} = 0$. Thus, in the style of the singular value decomposition, we end up with

$$\tilde{Y}^H\mathcal{R}(L_*)\tilde{X} = \begin{bmatrix} \tilde{Y}_1^H\mathcal{R}\tilde{X}_1 & \tilde{Y}_1^H\mathcal{R}\tilde{X}_{21} & \tilde{Y}_1^H\mathcal{R}\tilde{X}_{22} \\ \tilde{Y}_{21}^H\mathcal{R}\tilde{X}_1 & \tilde{Y}_{21}^H\mathcal{R}\tilde{X}_{21} & \tilde{Y}_{21}^H\mathcal{R}\tilde{X}_{22} \\ \tilde{Y}_{22}^H\mathcal{R}\tilde{X}_1 & \tilde{Y}_{22}^H\mathcal{R}\tilde{X}_{21} & \tilde{Y}_{22}^H\mathcal{R}\tilde{X}_{22} \end{bmatrix} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & 0 \\ \Sigma_{21} & \Sigma_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix} =: \Sigma, \tag{3.7}$$

and $\mathcal{C}(L_*)$ can be decomposed in the following way

$$\mathcal{C}(L_*, U, V) = \begin{bmatrix} \mathcal{R} & \mathcal{V} \\ \mathcal{U}^H & 0 \end{bmatrix} = \begin{bmatrix} \tilde{Y}\Sigma\tilde{X}^H & \mathcal{V} \\ \mathcal{U}^H & 0 \end{bmatrix} = \begin{bmatrix} \tilde{Y} & 0 \\ 0 & I \end{bmatrix} \tilde{\mathcal{C}}(L_*, U, V) \begin{bmatrix} \tilde{X} & 0 \\ 0 & I \end{bmatrix}^H,$$

where

$$\tilde{\mathcal{C}}(L_*, U, V) := \begin{bmatrix} \Sigma & \tilde{Y}^H\mathcal{V} \\ \mathcal{U}^H\tilde{X} & 0 \end{bmatrix} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & 0 & \mathcal{V}_1 \\ \Sigma_{21} & \Sigma_{22} & 0 & \mathcal{V}_{21} \\ 0 & 0 & 0 & \mathcal{V}_{22} \\ \mathcal{U}_1^H & \mathcal{U}_{21}^H & \mathcal{U}_{22}^H & 0 \end{bmatrix}. \tag{3.8}$$

4. Conditions for nonsingularity. In what follows we choose the orthonormal bordering matrices $U = X_1$, $V = \hat{Y}_1$, which are optimal in that they minimize $\|(U^HX_1)^{-1}\|$, $\|(V^H\hat{Y}_1)^{-1}\|$. This choice implies $\Phi = (U^HX_1)^{-1} = I$, hence, $X_* = X_1$, $L_* = L_1$, and $\mathcal{U}_1 = \mathcal{V}_1 = 0$ in (3.8). It also follows that $[\,\mathcal{U}_{21}^H \,|\, \mathcal{U}_{22}^H\,]$ and $[\,\mathcal{V}_{21}^H \,|\, \mathcal{V}_{21}^H\,]$ are unitary.

This section starts with equivalent statements on the nonsingularity of the block matrix $C(L_*, U, V) = C(L_1, X_1, \hat{Y}_1)$ and provides a theorem on the general feasibility of matrix borderings.

THEOREM 4.1. *Let $X_1 \in \mathbb{C}^{p\times n}$ with $X_1^HX_1 = I_p$, $AX_1 = X_1L_1$, define a simple invariant subspace $\operatorname{Im} X_1$ of $A \in \mathbb{C}^{n\times n}$. Moreover, let $\hat{Y}_1 \in \mathbb{C}^{p\times n}$, with $\hat{Y}_1^H\hat{Y}_1 = I_p$ and $A^H\hat{Y}_1 = \hat{Y}_1\hat{L}_1^H$, define a corresponding invariant subspace $\operatorname{Im} \hat{Y}_1$ of A^H; cf. (3.1). Define $\hat{X}_2$, Y_2 according to (1.6) and (3.1). Then, the following statements are equivalent:*

- *(i) The linear operator $C(L_1, X_1, \hat{Y}_1)$ defined in (2.1) (or, equivalently, its matrix representation $\mathcal{C}(L_1, X_1, \hat{Y}_1)$ defined in (3.5), or its transformed version $\tilde{\mathcal{C}}(L_1, X_1, \hat{Y}_1)$ defined in (3.8)) is nonsingular.*
- *(ii) The matrix Σ_{11} defined in (3.7) is nonsingular.*
- *(iii) We have $\lambda(P) \cap \lambda(L_1) = \emptyset$, where $P = (\hat{X}_2^HAY_2)(\hat{X}_2^HY_2)^{-1}$.*

Proof. Let us prove the equivalence of (i) and (ii). Due to the choice $U = X_1$, $V = \hat{Y}_1$, the matrix $\tilde{\mathcal{C}}(L_1, X_1, \hat{Y}_1)$ reduces to

$$\tilde{\mathcal{C}}(L_1) = \begin{bmatrix} \Sigma_{11} & \tilde{\Sigma}_{12} \\ \tilde{\Sigma}_{21}^H & S_2 \end{bmatrix},$$

with $\tilde{\Sigma}_{12} := [\,\Sigma_{12} \,|\, 0 \,|\, 0\,]$, $\tilde{\Sigma}_{21} := [\,\Sigma_{21}^H \,|\, 0 \,|\, 0\,]$, and the nonsingular matrix

$$S_2 = \begin{bmatrix} \Sigma_{22} & 0 & \mathcal{V}_{21} \\ 0 & 0 & \mathcal{V}_{22} \\ \mathcal{U}_{21}^H & \mathcal{U}_{22}^H & 0 \end{bmatrix} \quad \text{with} \quad S_2^{-1} = \begin{bmatrix} 0 & 0 & \mathcal{U}_{21} \\ 0 & 0 & \mathcal{U}_{22} \\ \mathcal{V}_{21}^H & \mathcal{V}_{22}^H & -\mathcal{V}_{21}^H\Sigma_{22}\mathcal{U}_{21} \end{bmatrix}.$$

ETNA
Kent State University
http://etna.math.kent.edu

Therefore, the Schur complement S_1 of S_2 exists, and $\tilde{\mathcal{C}}(L_1, X_1, \hat{Y}_1)$ is nonsingular if and only if S_1 is nonsingular. But $S_1 = \Sigma_{11} - \tilde{\Sigma}_{12} S_2^{-1} \tilde{\Sigma}_{21}^H = \Sigma_{11}$.

Last, we show the equivalence of (i) and (iii). Equation (3.7) provides $\Sigma_{11} = \tilde{Y}_1^H \mathcal{R} \tilde{X}_1$. Now decompose the unitary matrix $\tilde{X}$ in two instead of three parts

$$\tilde{X} = [\tilde{X}_1 \,|\, \tilde{X}_2] = [x^{(np)} \cdots x^{(p^2+1)} \,|\, x^{(p^2)} \cdots x^{(1)}],$$

i.e., $\tilde{X}_2 = [\tilde{X}_{21}\, \tilde{X}_{22}]$. Since ${x^{(i)}}^H x^{(j)} = {x^{(i)}}^H (I \otimes X_1) f^{(j)} = 0$ for $i = p^2+1, \ldots, np$, $j = 1, \ldots, p^2$ (cf. (3.6)), the orthogonality conditions yield $X_1^H X^{(i)} = 0$ and, consequently, ${x^{(i)}}^H (I \otimes X_1) = 0$ which is equivalent to $(I \otimes X_1^H) x^{(i)} = \mathrm{vec}(X_1^H X^{(i)}) = 0$. Hence, and because we have $X_1^H Y_2 = 0$, the existence of matrices K_i of size $q \times p$ follows, such that $X^{(i)} = Y_2 K_i$. This gives $x^{(i)} = (I \otimes Y_2) k^{(i)}$ where $k^{(i)} := \mathrm{vec}(K_i)$.

Setting $\tilde{K}_1 = [k^{(np)} \ldots k^{(p^2+1)}]$ leads to the representation $\tilde{X}_1 = (I \otimes Y_2)\tilde{K}_1$. Because of

$${x^{(i)}}^H x^{(j)} = {k^{(i)}}^H (I \otimes Y_2^H)(I \otimes Y_2) k^{(j)} = {k^{(i)}}^H k^{(j)} = \delta_{ij},$$

the matrix $\tilde{K}_1$ is unitary. Analogously, one obtains $\tilde{Y}_1 = (I \otimes \hat{X}_2)\tilde{N}_1$, where the matrix $\tilde{N}_1 = [n^{(np)} \ldots n^{(p^2+1)}]$ is unitary. Applying both results gives

$$\begin{aligned}
\Sigma_{11} &= \tilde{Y}_1^H \mathcal{R} \tilde{X}_1 = \tilde{N}_1^H (I \otimes \hat{X}_2^H) \mathcal{R} (I \otimes Y_2) \tilde{K}_1 \\
&= \tilde{N}_1^H \left[I \otimes \hat{X}_2^H A Y_2 - L_1^T \otimes \hat{X}_2^H Y_2 \right] \tilde{K}_1 \\
&= \tilde{N}_1^H \left[I \otimes (\hat{X}_2^H A Y_2)(\hat{X}_2^H Y_2)^{-1} - L_1^T \otimes I \right] (I \otimes \hat{X}_2^H Y_2) \tilde{K}_1.
\end{aligned}$$

Since all other factors are nonsingular, Σ_{11} is nonsingular if and only if the inner factor $I \otimes P - L_1^T \otimes I$ is nonsingular where $P := (\hat{X}_2^H A Y_2)(\hat{X}_2^H Y_2)^{-1}$. But this is equivalent to the nonsingularity of the mapping $T(P, L_1)[Z] = PZ - ZL_1$, i.e., to the condition $\lambda(P) \cap \lambda(L_1) = \emptyset$. □

Since $\hat{X}_2^H Y_2 = D_2^{-1/2}$ and $\hat{X}_2^H A Y_2 = D_2^{-1/2}(D_2 L_2 - Q^H L_1 Q)$, we obtain a more convenient expression for P, namely

$$P = D_2^{1/2}(L_2 - Q^H L_1 Q) D_2^{1/2},$$

which made it possible to construct our counterexample; see Section 5.

4.1. Feasible borderings. The analysis of block matrices with four blocks gives the following general result on nonsingularity.

THEOREM 4.2. *Let*

$$G = \begin{bmatrix} G_{11} & G_{12} \\ G_{21}^H & G_{22}^H \end{bmatrix},$$

where $G_{11} \in \mathbb{C}^{n \times n}, G_{12}, G_{21} \in \mathbb{C}^{n \times p}, G_{22} \in \mathbb{C}^{p \times p}$ *and let*

$$\mathcal{X} = \begin{bmatrix} X \\ M \end{bmatrix} \in \mathbb{C}^{(n+p) \times p}$$

such that $\ker[G_{11}\ \ G_{12}] = \mathrm{Im}\, \mathcal{X}$. *Then, the block matrix G is nonsingular, if and only if* $\mathrm{rank}\,[G_{11}\ G_{12}\,] = n$ *and* $[G_{21}^H\ G_{22}^H]\mathcal{X}$ *is nonsingular.*

Proof. For the nonsingularity of the block matrix G it is clearly necessary that $[G_{11}\ G_{12}]$ has full row rank, i.e., the dimension of the kernel of $[G_{11}\ G_{12}]$ equals p. Suppose there exist $x \in \mathbb{R}^n, m \in \mathbb{R}^p$, such that

$$G\begin{bmatrix} x \\ m \end{bmatrix} = \begin{bmatrix} G_{11} & G_{12} \\ G_{21}^H & G_{22}^H \end{bmatrix}\begin{bmatrix} x \\ m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Longleftrightarrow \quad \begin{cases} G_{11}x + G_{12}m = 0, \\ G_{21}^H x + G_{22}^H m = 0. \end{cases} \tag{4.1}$$

Then the upper block implies

$$\begin{bmatrix} x \\ m \end{bmatrix} \in \ker[G_{11}\ G_{12}] = \operatorname{Im} \mathcal{X},$$

i.e., $x = Xs$, $m = Ms$, for some vector s. Inserting this representations into the lower block of equation (4.1) yields $G_{21}^H Xs + G_{22}^H Ms = 0$. Since $[G_{21}^H\ G_{22}^H]\mathcal{X}$ is nonsingular, we have $s = 0$, and hence $x = 0$ and $m = 0$. ☐

A similar statement holds if one considers the left part of G instead of the upper one. Note, that $p > \dim\ker(G_{11})$ is feasible. Moreover, the nonsingularity of $[G_{21}^H\ G_{22}^H]\mathcal{X}$ is equivalent to

$$\measuredangle\left(\operatorname{Im}\begin{bmatrix} G_{21} \\ G_{22} \end{bmatrix}, \operatorname{Im}\mathcal{X}\right) < \frac{\pi}{2}.$$

Applying Theorem 4.2 to our case, i.e., to the matrix $\mathcal{C}(L_1, X_1, \hat{Y}_1)$, we consider the left block $\left[\mathcal{R}^H\ \mathcal{U}\right]^H$. Then $\mathcal{R}x = 0$ holds, if and only if x is an element from the kernel, i.e., if $x = \tilde{X}_{22}\xi$. Furthermore, $\mathcal{U}^H x = 0$ yields $\mathcal{U}^H \tilde{X}_{22}\xi = \tilde{\mathcal{U}}_{22}^H \xi = 0$ implying $\xi = 0$, i.e., the full rank condition is always fulfilled. In this sense, the bordering with $\mathcal{U}$ is feasible. When looking at the conjugate transposed of $\mathcal{C}$, which is required for the computation of the left invariant subspace, the feasibility of $\mathcal{V}$ is easily verified as well.

5. The counterexample. While trying to prove the nonsingularity of the linear operator $C(L_1, X_1, \hat{Y}_1)$ we found a counterexample, i.e., a singular matrix $\mathcal{C}(L_1, X_1, \hat{Y}_1)$. Using Theorem 4.1 (iii) we eventually generated the following matrix

$$A = \left[\begin{array}{rr|rrr} 1 & 1 & -1 & 0 & -1 \\ 0 & -1 & 0 & 3 & 5 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{array}\right] = \left[\begin{array}{c|c} L_1 & H \\ \hline 0 & L_2 \end{array}\right] = [X_1\ Y_2]^H A [X_1\ Y_2],$$

where $[X_1\ \ Y_2\ ] = [e_1\ e_2\,|\,e_3\ e_4\ e_5]$, i.e., $L_* = L_1$, $X_* = X_1$ and $AX_1 = X_1L_1$, i.e., $\operatorname{Im} X_1$ is a simple invariant subspace of A with spectrum $\lambda(L_1) = \{1, -1\}$. The induced 14×14 matrix

$$\mathcal{C}(L_1, X_1, Y_1) = \begin{bmatrix} I \otimes A - L_1^T \otimes I & I \otimes Y_1 \\ I \otimes X_1^T & 0 \end{bmatrix}$$

is singular. Here, we have used the exact invariant subspace of A^T spanned by

$$V = Y_1 = \begin{bmatrix} 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & -1 \end{bmatrix}^T,$$

though it is not orthonormal. Nevertheless, the rank of $\mathcal{C}$ is the same for $V = Y_1$ and its orthonormalized version $V = \hat{Y}_1 = Y_1(Y_1^H Y_1)^{-1/2}$.

ETNA
Kent State University
http://etna.math.kent.edu

The upper block of $\mathcal{C}(L_1)$ has full row rank $n \cdot p = 10$, the left block has full column rank 10, but $\mathcal{C}(L_1)$ has rank drop 1. Examination of the structure of the involved matrices shows where the singularity results from. We check condition (iii) of Theorem 4.1. Recall, that $P = D_2^{1/2} L_2 D_2^{1/2} - D_2^{1/2} Q^H L_1 Q D_2^{1/2}$. Since the first column of L_2 is zero, the first column of $D_2^{1/2} L_2 D_2^{1/2}$ is zero, as well. Hence, whenever $D_2^{1/2} Q^H L_1 Q D_2^{1/2}$ is such that the first column equals $\lambda_i e_1$, where λ_i and $-\lambda_i$ are eigenvalues of L_1, then $-\lambda_i$ is an eigenvalue of P, too. If $p > q$, then the eigenvalues of $D_2^{1/2} Q^H L_1 Q D_2^{1/2}$ form a subset of the eigenvalues of L_1, because L_1 is $p \times p$ and $QD_2^{1/2}$ is $p \times q$. Otherwise, as holds for the example matrix where $p = 2$ and $q = 3$, no correlation in general exists. To make a general statement seems quite complicated, since the matrices D_2, Q, H, L_1 and L_2 depend on each other. The exceptional matrix arises through interaction of all of these.

6. Conclusions and outlook. The counterexample presented in this paper shows that a straightforward generalization of the singularity theory based approach for characterizing invariant subspaces from $p = 1$ to $p > 1$ does not work, in general.

One possibility to overcome this problem would be to change $\mathcal{C}(L_1)$ by introducing a bottom right block $\mathcal{T} \in \mathbb{C}^{p^2 \times p^2}$ instead of the zero matrix, i.e., to work with

$$\mathcal{C}_{mod}(L) = \begin{bmatrix} I \otimes A - L^T \otimes I & I \otimes V \\ I \otimes U^H & \mathcal{T} \end{bmatrix}.$$

Of course, there is always a matrix $\mathcal{T}$ such that $\mathcal{C}$ is nonsingular; cf. Theorem 4.2. However, in order to implement the method efficiently, the matrix $\mathcal{T}$ should have the form $\mathcal{T} = I \otimes T$, with $T \in \mathbb{C}^{p \times p}$ since then

$$\mathcal{C}_{mod}(L) = \begin{bmatrix} I \otimes A - L^T \otimes I & I \otimes V \\ I \otimes U^H & I \otimes T \end{bmatrix}$$

corresponds to the operator $C_{mod}(L)$ defined by

$$C_{mod}(L)[X, M] = \begin{bmatrix} AX - XL + VM \\ U^H X + TM \end{bmatrix}.$$

Anyway, this restriction in the choice of $\mathcal{T}$ reduces the degree of freedom from p^4 to p^2, and it is an open question whether such a T can be found.

Secondly, instead of changing the matrix we could ask for which classes of matrices the additional condition stated in Theorem 4.1, part (iii), is satisfied, but this question is not yet answered.

REFERENCES

[1] P.-A. ABSIL, R. SEPULCHRE, P. VAN DOOREN, AND R. MAHONY, *Cubically convergent iterations for invariant subspace computations*, SIAM J. Matrix Anal. Appl., 26 (2004), pp. 70–96.

[2] P.-A. ABSIL AND P. VAN DOOREN, *Two-sided Grassmann-Rayleigh quotient iteration*, Technical Report UCL-INMA-2007.024, Universite Catholique de Louvain, 2007.

[3] R. H. BARTELS AND G. W. STEWART, *Solution of the matrix equation $AX + XB = C$*, Comm. ACM, 15 (1972), pp. 820–826.

[4] W.-J. BEYN, W. KLESS, AND V. THÜMMLER, *Continuation of low-dimensional invariant subspaces in dynamical systems of large dimension*, in Ergodic Theory, Analysis and Efficient Simulation of Dynamical Systems, B. Fiedler, ed., Springer, Berlin, pp. 47–72, 2001.

[5] F. CHATELIN, *Simultaneous Newton's iteration for the eigenproblem*, Defect correction methods. Theory and applications, Comput. Suppl. 5, pp. 67–74, Springer, Vienna, 1984.

[6] J. W. DEMMEL, *Three methods for refining estimates of invariant subspaces,* Computing, 38 (1987), pp. 43–57.

[7] J. J. DONGARRA, C. B. MOLER, AND J. H. WILKINSON, *Improving the accuracy of computed eigenvalues and eigenvectors*, SIAM J. Numer. Anal., 46 (1983), pp. 46–58.
[8] J.-L. FATTEBERT, *A block Rayleigh quotient iteration with local quadratic convergence*, Electron. Trans. Numer. Anal., 7 (1998), pp. 56–74.
http://etna.math.kent.edu/vol.7.1998/pp56-74.dir/.
[9] F. R. GANTMACHER, *Matrizenrechnung. Teil I: Allgemeine Theorie*, Deutscher Verlag der Wissenschaften, Berlin, 1958.
[10] R. LÖSCHE, H. SCHWETLICK, AND G. TIMMERMANN, *A modified block Newton iteration for approximating an invariant subspace of a symmetric matrix*, Linear Algebra Appl., 275/276 (1998), pp. 381–400.
[11] A. OSTROWSKI, *On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors I–VI*, Arch. Rational Mech. Anal., Part I: 1 (1958), pp. 233–241, Part II: 2 (1959), pp. 423–428, Part III: 3 (1959), pp. 325–340, Part IV: 3 (1959), pp. 341–347, Part V: 3 (1959), pp. 472–481, Part VI: 4 (1959), pp. 153–165.
[12] B. N. Parlett. *The Rayleigh quotient iteration and some generalizations for nonnormal matrices*, Math. Comp., 28 (1974), pp. 679–693.
[13] H. SCHWETLICK AND R. LÖSCHE, *A generalized Rayleigh quotient iteration for computing simple eigenvalues of nonnormal matrices*, Z. Angew. Math. Mech., 80 (2000), pp. 9–25.
[14] G. W. STEWART, *Introduction to Matrix Computations*, Academic Press, New York, 1973.
[15] G. W. STEWART AND JI-GUANG SUN, *Matrix Perturbation Theory*, Academic Press, Boston, 1990.
[16] H. UNGER, *Nichtlineare Behandlung von Eigenwertaufgaben*, Z. Angew. Math. Mech., 30 (1950), pp. 281–282. English translation available at
http://www.math.tu-dresden.de/~schwetli/Unger.html.