Math. Nachr. **208** (1999), 5 – 29

Verification Algorithms for Generalized Singular Values

By GÖTZ ALEFELD of Karlsruhe, ROLF HOFFMANN of Karlsdorf
and GÜNTER MAYER of Rostock

Dedicated to Prof. Dr. G. WILDENHAIN, *Rostock, on the occasion of his* 60^{th} *birthday*

(Received January 7, 1997)
(Revised Version December 18, 1998)

Abstract. By means of interval arithmetic tools we present new algorithms for verifying and enclosing generalized singular values and corresponding vectors for a matrix pair $(A, B) \in \mathbb{R}^{p\times n} \times \mathbb{R}^{q\times n}$. To this end, we state and prove a fundamental theorem in interval analysis which shows a way how enclosures can be constructed if approximations are known. Furthermore, we perform a careful comparison of the new method with those introduced in [11].

1. Introduction

We first address the main purpose of our paper. It consists in constructing tight intervals which contain the components c, s of a generalized singular value and the components of a column of U, V and X, respectively. This means, in particular, that we provide a method which verifies a generalized singular value and the corresponding column vectors of U, V, X. Gathering these quantities into a vector z^* we will interprete z^* as a fixed point of some function t. We will expand t into a Taylor series at an approximation $\tilde{z}$ of z^*. The interval arithmetic evaluation of the resulting expression yields an interval function $[g]([z], \tilde{z})$ which forms the base of our verification algorithm. We will apply a general result from interval analysis (Theorem 3.1) in order to construct an interval vector $[z]$ such that $[g]([z], \tilde{z}) \subseteq [z]$ holds. This subset property guarantees that the last two components of $[z]$ enclose exactly one generalized singular value (c^*, s^*). We will also improve the bounds for z^* iteratively by constructing a sequence $[z]^k$ which starts with $[z]^0 := [z]$ and which converges to z^*.

1991 *Mathematics Subject Classification.* Primary: 15A18, 15A23, 65F15, 65F20, 65G10, 65H10.

Keywords and phrases. Singular values, generalized singular values, verification of generalized singular values, enclosures for generalized singular values, interval analysis, nonlinear systems of equations.

In [11] two different methods for verifying generalized singular values have already been introduced. Their convergence analysis has been performed independently of each other. We show how these methods are related to the new one. We are even able to unify the convergence analysis of the methods of [11] and that of the method of the present paper by using the same fundamental theorem mentioned already. Note that this theorem also yields the results of [1] – [4] and [6] when applied appropriately; cf. [15] for details. We also recall that simple singular values of the singular value decomposition (1.2) have first been verified and enclosed in [2]; see also [14].

Finally, we mention already at this point that the number of unknowns (and equations) of the new method is smaller by one and $n+1$, respectively, compared with the two methods from [11]. Of course, with respect to the size this is a minor improvement, at least in the first case.

We have arranged our paper as follows: To have a clear basis we continue this introduction by repeating the definition of a singular value and that of a generalized singular value, respectively. After that the basic facts about these concepts are repeated. In Section 2 we list the notation which we use throughout the paper. In Section 3 we recall a general result from interval analysis which is central for the theoretical results of Section 4. These results are illustrated by a numerical example in Section 5.

As is well-known (cf. [10] for example), the singular values σ_i of a rectangular matrix $A \in \mathbb{R}^{p\times n}$ are defined as the nonnegative square roots of the eigenvalues λ_i of the $n \times n$ matrix $A^T A$. Therefore, they are solutions of the equation

$$\det\left(A^T A - \lambda I\right) = 0\,. \tag{1.1}$$

Singular values give insight in the structure of a matrix; the following theorem holds, for example.

Theorem 1.1. ([10, p. 16].) *If $A \in \mathbb{R}^{p\times n}$ then there exist orthogonal matrices $U \in \mathbb{R}^{p\times p}$, $V \in \mathbb{R}^{n\times n}$ such that*

$$U^T A V = \Sigma \tag{1.2}$$

where $\Sigma := \operatorname{diag}\left(\sigma_1, \ldots, \sigma_{\min\{p,n\}}\right) \in \mathbb{R}^{p\times n}$ with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_{\min\{p,n\}} = 0$, $r := \operatorname{rank}(A)$.

The representation (1.2) is called the singular value decomposition of A. It is used, e. g., for solving the least squares problem

$$\|Ax - b\|_2 \longrightarrow \min\,, \quad b \in \mathbb{R}^p\,, \quad x \in \mathbb{R}^n\,, \tag{1.3}$$

such that $\|x\|_2 \to \min$. (As usual, $\|\cdot\|_2$ denotes the Euclidean norm.)

In order to consider the least squares problem (1.3) subject to the quadratic inequality constraint

$$\|Bx - d\|_2 \leq \alpha \tag{1.4}$$

with $B \in \mathbb{R}^{q\times n}$, $d \in \mathbb{R}^q$, one generalizes the concept of singular values. (Cf. [10, p. 404 ff.].) One might start with the generalized eigenvalue problem which results in

$$\det\left(A^T A - \lambda B^T B\right) = 0 \tag{1.5}$$

and one could define generalized singular values as the positive square roots of the positive eigenvalues of (1.5). This was done in [20]. In order to eliminate the preference of A against B PAIGE and SAUNDERS defined in [18] the generalized singular values of the matrix pair (A, B) as those pairs (c, s) of real numbers which form the solutions of the equation

$$\det\left(s^2 A^T A - c^2 B^T B\right) = 0. \tag{1.6}$$

Since with (c, s) each pair $(\tau c, \tau s)$, $(\tau c, -\tau s)$ with $\tau \in \mathbb{R}$ solves (1.6), one requires

$$c \geq 0, \quad s \geq 0, \quad c^2 + s^2 = 1 \tag{1.7}$$

which reduces the number of solutions drastically. In particular, $(0, 0)$ is not a generalized singular value. In our paper we will use (1.6), (1.7) as the definition of a generalized singular value.

If $\operatorname{rank}\begin{pmatrix} A \\ B \end{pmatrix} < n$ then there is a vector $x \in \mathbb{R}^n \backslash \{0\}$ with $Ax = 0$ and $Bx = 0$ simultaneously. Hence $\left(s^2 A^T A - c^2 B^T B\right)x = 0$ and $\det\left(s^2 A^T A - c^2 B^T B\right) = 0$ for all pairs $(c, s) \in \mathbb{R}^2$. Due to this fact we will assume

$$\operatorname{rank}\begin{pmatrix} A \\ B \end{pmatrix} = n \tag{1.8}$$

from now on.

We want to show how the solutions of (1.5) are related to the solutions of (1.6), (1.7). Assume first that λ^* is a solution of (1.5), i.e., there is a vector $x^* \neq 0$ such that $\left(A^T A - \lambda^* B^T B\right)x^* = 0$ holds. This implies $\|Ax^*\|_2^2 = \lambda^* \|Bx^*\|_2^2$, whence $Bx^* \neq 0$ by (1.8). Thus we get $\lambda^* \geq 0$. Therefore, each solution λ^* of (1.5) induces the generalized singular value $(c^*, s^*) = \left(\frac{\sqrt{\lambda^*}}{\sqrt{1+\lambda^*}}, \frac{1}{\sqrt{1+\lambda^*}}\right)$ with $s^* \neq 0$ and, conversely, each solution (c^*, s^*) of (1.6), (1.7) with $s^* \neq 0$ induces the eigenvalue $\lambda^* = \frac{(c^*)^2}{(s^*)^2}$ of the generalized eigenvalue problem (1.5). But note that the set

$$\begin{aligned} \mu(A, B) := \{ (c, s) \in \mathbb{R}^2 \mid & \det\left(s^2 A^T A - c^2 B^T B\right) = 0, \\ & c \geq 0,\ s \geq 0,\ c^2 + s^2 = 1 \} \end{aligned} \tag{1.9}$$

of generalized singular values may also contain the element $(1, 0)$ which is not related to a solution of (1.5) and which implies $\operatorname{rank}(B) < n$ by (1.6), or, equivalently, $Bx = 0$ for some vector $x \neq 0$.

An example, for which this is not the case is given by choosing $p = q = n$ and B nonsingular. Then the condition (1.8) is fulfilled and $\det\left(B^T B\right) \neq 0$ whence, by (1.6), $(1, 0)$ is not a generalized singular value. From

$$\det\left(s^2 A^T A - c^2 B^T B\right) = (\det B)^2 \det\left(s^2 \left(AB^{-1}\right)^T \left(AB^{-1}\right) - c^2 I\right) = 0$$

one sees at once that the generalized singular values (c, s) of A, B correspond here directly to the singular values of AB^{-1} via $\sqrt{\lambda} = \frac{c}{s}$.

Another example which illustrates the connection between singular values and generalized singular values is given by $A \in \mathbb{R}^{p \times n}$, $B := \operatorname{diag}(1, \ldots, 1) \in \mathbb{R}^{q \times n}$ with $q \geq n$.

Then $B^TB = I \in \mathbb{R}^{n\times n}$, hence $\det(B^TB) \neq 0$ and, again, $(c,s) = (1,0)$ is not a generalized singular value of (A,B). Therefore, the quotients $\frac{c}{s}$ with $(c,s) \in \mu(A,B)$ are precisely the singular values of A. This justifies the name generalized singular values.

Definition 1.2. A generalized singular value (c,s) is called *simple*, if

i) in the case $cs \neq 0$, $\lambda^* = \frac{c^2}{s^2}$ is a simple zero of $\det(A^TA - \lambda B^TB)$,

ii) in the case $(c,s) = (0,1)$, $\lambda^* = 0$ is a simple zero of $\det(A^TA - \lambda I)$ and of $\det(AA^T - \lambda I)$,

iii) in the case $(c,s) = (1,0)$, $\lambda^* = 0$ is a simple zero of $\det(B^TB - \lambda I)$ and of $\det(BB^T - \lambda I)$.

Since the multiplicity of the eigenvalue $\lambda^* = 0$ of A^TA might be different from that of AA^T, the simplicity of $\lambda^* = 0$ for both A^TA and AA^T is required in ii). The same remark applies to the definition of the simplicity of the pair $(1,0)$ in iii).

In all cases where $\min\{p,q\} < n$ but $p+q \geq n$ (the latter inequality is necessary for (1.8)), one can add $n-p$ rows $z^T = 0$ to A and $n-q$ rows $\hat{z}^T = 0$ to B, respectively, in order to fulfill $p, q \geq n$. This modification does neither change the generalized singular values of (A,B) nor does it change the rank in (1.8). Therefore, without loss of generality, we assume $p, q \geq n$ for all our subsequent considerations. Then, analogously to (1.2), the matrices A, B can also be decomposed by means of orthogonal matrices. The following theorem is contained as a particular case in Theorem 2 of [21].

Theorem 1.3. *Let $A \in \mathbb{R}^{p\times n}$, $B \in \mathbb{R}^{q\times n}$ with $p, q \geq n$. If (1.8) holds then there exist orthogonal matrices $U \in \mathbb{R}^{p\times p}$, $V \in \mathbb{R}^{q\times q}$ and a nonsingular matrix $X \in \mathbb{R}^{n\times n}$ such that*

$$U^TAX = \Sigma_A = \begin{pmatrix} C \\ O \end{pmatrix} \in \mathbb{R}^{p\times n}, \tag{1.10}$$

$$V^TBX = \Sigma_B = \begin{pmatrix} S \\ O \end{pmatrix} \in \mathbb{R}^{q\times n}, \tag{1.11}$$

where $C = \operatorname{diag}(c_1,\ldots,c_n) \in \mathbb{R}^{n\times n}$, $S = \operatorname{diag}(s_1,\ldots,s_n) \in \mathbb{R}^{n\times n}$ are diagonal matrices satisfying

$$c_i \geq 0, \quad s_i \geq 0 \quad \text{and} \quad c_i^2 + s_i^2 = 1, \quad i = 1,\ldots,n. \tag{1.12}$$

For the set $\mu(A,B)$ from (1.9) one obtains

$$\mu(A,B) = \{(c_i,s_i) \mid i = 1,\ldots,n\}. \tag{1.13}$$

Multiple occurence of the pairs (c_i,s_i) is allowed. The representation (1.10), (1.11) is called a generalized singular value decomposition of (A,B). It is certainly not unique since U, V, X can always be replaced by $-U$, $-V$ and $-X$ without changing the right-hand side of (1.10) and (1.11).

We remark that in the more general formulation of Theorem 1.3 in [21] the restrictions (1.8) and $q \geq n$ are not required. If $p < n$ or $q < n$ the right-hand sides of

(1.10), (1.11) read $(C \quad O)$ and $(S \quad O)$, respectively. In this case a representation analogously to (1.10), (1.11) may, however, not exist even if (1.8) holds. This can be seen from the simple counterexample $A = (1\ 0)$, $B = (0\ 1)$ from [20] which requires X to be singular. Extending A and B to $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ yields to (1.10), (1.11) with $U = X = I$ and $V = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, for example.

A more general definition of a generalized singular value decomposition can be found in [7], p. 22, e. g. It takes into account the cases $p < n$ and $q < n$ in full generality and is equivalent to ours in the case $p, q \geq n$; cf. [7, p. 205].

We finally mention two simple corollaries which can be deduced from Theorem 1.3.

Corollary 1.4. *With the notation and with the assumptions of Theorem* 1.3 *we get*

$$\operatorname{rank}(A) = \operatorname{rank}(\Sigma_A) = \operatorname{rank}(C), \quad \operatorname{rank}(B) = \operatorname{rank}(\Sigma_B) = \operatorname{rank}(S).$$

Corollary 1.5. *With the notation and with the assumptions of Theorem* 1.3 *we get*

$$(s_i^2 A^T A - c_i^2 B^T B) x^i = 0 \tag{1.14}$$

for any column x^i of X. In particular, if $s_i \neq 0$ then x^i is an eigenvector of the generalized eigenvalue problem $(A^T A - \lambda B^T B)x = 0$. The corresponding eigenvalue is $\lambda = \frac{c_i^2}{s_i^2}$.

Proof. From (1.10), (1.11) we get $AX = U\Sigma_A$, $U^T A = \Sigma_A X^{-1}$, $BX = V\Sigma_B$, $V^T B = \Sigma_B X^{-1}$. This implies

$$\begin{aligned} A^T AX &= A^T U \Sigma_A = (X^{-1})^T \Sigma_A^T \Sigma_A = (X^{-1})^T C^2, \\ B^T BX &= B^T V \Sigma_B = (X^{-1})^T \Sigma_B^T \Sigma_B = (X^{-1})^T S^2, \end{aligned}$$

and therefore

$$A^T AXS^2 - B^T BXC^2 = O,$$

which proves (1.14). The remaining properties are trivial consequences of (1.14). □

2. Notations

In this section we list the notations which we will use throughout the paper.

By $\mathbb{R}^n$, $\mathbb{R}^{m\times n}$, $\boldsymbol{IR}$, $\boldsymbol{IR}^n$, $\boldsymbol{IR}^{m\times n}$ we denote the set of real vectors with n components, the set of real $m \times n$ matrices, the set of intervals, the set of interval vectors with n components and the set of $m \times n$ interval matrices, respectively. By "interval" we always mean a real compact interval. Interval vectors and interval matrices are vectors and matrices, respectively, with interval entries. We write intervals in brackets with the exception of degenerate intervals (so-called point intervals) which we identify with the element being contained, and we proceed similarly with interval vectors and interval matrices. Examples are the null matrix O, the identity matrix I, the i-th

column $e^{(i)}$ of I and the vector $e := (1, 1, \ldots, 1)^T$. In order to indicate $I \in \mathbb{R}^{n\times n}$ and $e \in \mathbb{R}^n$, respectively, we sometimes write I_n instead of I and e_n instead of e. Although e_n will also denote the n-th component of e it will always be clear from the context which meaning is valid. As usual, we identify $\mathbb{R}^{n\times 1}$ and $\boldsymbol{IR}^{n\times 1}$ with $\mathbb{R}^n$ and $\boldsymbol{IR}^n$, respectively. We equip $\mathbb{R}^n$ and $\mathbb{R}^{m\times n}$, respectively, with the natural semi-ordering "$\leq$" which is defined to hold entrywise. We use the notation $[A] = \big([a]_{ij}\big) \in \boldsymbol{IR}^{m\times n}$ simultaneously without further reference, and we assume the same for the elements of $\mathbb{R}^n$, $\mathbb{R}^{m\times n}$ and $\boldsymbol{IR}^n$. For $[a] = [\underline{a}, \overline{a}] \in \boldsymbol{IR}$ we define the absolute value $\big|[a]\big|$ by $\big|[a]\big| := \max\{|\underline{a}|, |\overline{a}|\}$ and the diameter $d\big([a]\big)$ by $d\big([a]\big) := \overline{a} - \underline{a}$, and we denote the convex hull of $[a]$, $[b] \in \boldsymbol{IR}$ by $[a] \cup [b]$. For interval vectors and interval matrices, these quantities are defined entrywise. In particular, $|A| = \big(|a_{ij}|\big) \in \mathbb{R}^{m\times n}$ for point matrices $A \in \mathbb{R}^{m\times n}$. The operations for intervals etc. can be found in [5]. Based on the elementary rules

$$\begin{aligned}
\big|[a]\big| &\leq \big|[b]\big| \quad \text{for} \quad [a] \subseteq [b]\,, \\
\big|[a] \pm [b]\big| &\leq \big|[a]\big| + \big|[b]\big|\,, \\
\big|[a][b]\big| &= \big|[a]\big|\,\big|[b]\big|\,, \\
d\big([a]\big) &\leq d\big([b]\big) \quad \text{for} \quad [a] \subseteq [b]\,, \\
d\big([a] \pm [b]\big) &= d\big([a]\big) + d\big([b]\big)\,, \\
d\big([a][b]\big) &\leq \big|[a]\big|\, d\big([b]\big) + d\big([a]\big)\big|[b]\big|\,, \\
d\big(c[a]\big) &= |c|\, d\big([a]\big)\,,
\end{aligned}$$

for $[a]$, $[b] \in \boldsymbol{IR}$, $c \in \mathbb{R}$, one easily proves the same relations for interval matrices $[A]$, $[B]$ and for real matrices C of the appropriate dimensions, with the exception of the third line in which the equality sign has to be replaced by "$\leq$". We recall the subdistributivity

$$[a]\big([b] + [c]\big) \subseteq [a][b] + [a][c] \tag{2.1}$$

of the interval arithmetic which shows that algebraically $(\boldsymbol{IR}, +, \cdot)$ is neither a ring nor a field. In fact, $(\boldsymbol{IR}, +)$ and $(\boldsymbol{IR}, \cdot)$ are two commutative semi-groups with the zero elements 0 and 1, respectively. The subdistributivity (2.1) also holds if $[a]$, $[b]$, $[c]$ are replaced by interval matrices.

Convergence and continuity in $\boldsymbol{IR}$, $\boldsymbol{IR}^n$, $\boldsymbol{IR}^{m\times n}$ are understood with respect to the Hausdorff distance $q(\cdot, \cdot)$ which reads $q\big([a], [b]\big) := \max\big\{\, |\underline{a} - \underline{b}|\,,\, |\overline{a} - \overline{b}| \,\big\}$ for elements $[a] = [\underline{a}, \overline{a}]$, $[b] = \big[\underline{b}, \overline{b}\big]$ of $\boldsymbol{IR}$, and which is defined entrywise in $\boldsymbol{IR}^n$ and $\boldsymbol{IR}^{m\times n}$.

For further details on interval analysis we refer to [5] or [16].

As we already mentioned, we write $\|\cdot\|_2$ for the Euclidean norm in $\mathbb{R}^n$, and we use $\|\cdot\|_\infty$ for the maximum norm of vectors and for the row sum norm of matrices, respectively. By $A := \operatorname{diag}(\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^{m\times n}$ we denote the rectangular diagonal matrix with the entries $a_{ij} := 0$ if $i \neq j$ and $a_{ii} := \alpha_i$ in the diagonal.

3. An auxiliary result from interval analysis

In this section we provide an essential tool (Theorem 3.1) for verifying and enclosing a fixed point x^* of some given function $t : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ which is assumed to be twice continuously differentiable on a given open set D. Let t' denote the derivative of t. With some fixed vector $\tilde{x}$ from D we will consider the interval function

$$[g]([x],\tilde{x}) := t(\tilde{x}) + t'(\tilde{x})([x]-\tilde{x}) + [h]([x],\tilde{x}) \tag{3.1}$$

for $[x] \subseteq D$. For all such $[x]$ we require

$$t(x) \in [g]([x],\tilde{x}) \quad \text{for all} \quad x \in [x], \tag{3.2}$$

and

$$\| |[h]([x],\tilde{x})| \|_\infty \leq \gamma \| |[x]-\tilde{x}| \|_\infty^2 . \tag{3.3}$$

Here, the interval function $[h]$ with $[h]([x],\tilde{x}) \in \boldsymbol{IR}^n$ is supposed to be continuous with respect to the first argument and inclusion monotone (i.e., $[x] \subseteq [y]$ implies $[h]([x],\tilde{x}) \subseteq [h]([y],\tilde{x})$ for fixed $\tilde{x}$). The constant γ is nonnegative and fixed for all $[x] \subseteq D$; it may depend on $\tilde{x}$.

In the subsequent Theorem 3.1 we will indicate a way how to construct an interval vector $[x]$ such that $t(x) \in [x]$ holds for all $x \in [x]$. Hence Brouwer's fixed point theorem guarantees the existence of at least one fixed point $x^* \in [x]$ of t. In Section 4 we will choose

$$t(x) := x - Pf(x) \tag{3.4}$$

with some nonsingular matrix P. Then the fixed points of t are the zeros of f and vice versa. Thus, together with (3.4), Theorem 3.1 also provides a mechanism for verifying and enclosing zeros of f.

Theorem 3.1. *With D, $[g]$, $[h]$, t, $\tilde{x}$ as in (3.1) - (3.3) and with $\gamma > 0$ from (3.3) choose $r \in \mathbb{R}$ with $r \geq 0$ such that $[x]^0 := \tilde{x} + [-r,r]e \subseteq D$, and define α, β by*

$$\alpha := \| t(\tilde{x}) - \tilde{x} \|_\infty , \quad \beta := \| t'(\tilde{x}) \|_\infty .$$

Let $\beta < 1$, $\Delta := (1-\beta)^2 - 4\alpha\gamma \geq 0$ and let

$$r^- := (1-\beta-\sqrt{\Delta})/(2\gamma) , \quad r^+ := (1-\beta+\sqrt{\Delta})/(2\gamma) .$$

a) *If $r \geq r^-$ then t has at least one fixed point $x^* \in [x]^0$. The iteration*

$$[x]^{k+1} := [g]([x]^k,\tilde{x}) \cap [x]^k , \quad k = 0,1,\ldots ,$$

converges to some interval vector $[x]^*$ with

$$x^* \in [x]^* \subseteq [x]^k \subseteq [x]^{k-1} \subseteq \cdots \subseteq [x]^0 , \quad k \in \mathbb{N} .$$

b) *If $r \in [r^-,r^+]$ then t has at least one fixed point $x^* \in [x]^0$. In addition, $[g]([x]^0,\tilde{x}) \subseteq [x]^0$ holds and the iteration*

$$[x]^{k+1} := [g]([x]^k,\tilde{x}) , \quad k = 0,1,\ldots ,$$

converges to some interval vector $[x]^*$ *with*

$$x^* \in [x]^* \subseteq [x]^k \subseteq [x]^{k-1} \subseteq \cdots \subseteq [x]^0\,, \quad k \in \mathbb{N}\,.$$

c) *In addition to* (3.3), *let* $[h]$ *fulfill*

$$\| d([h]([x],\tilde{x})) \|_\infty \;\leq\; 2\,\delta\, \| |[x]-\tilde{x}| \|_\infty \, \| d([x]) \|_\infty \tag{3.5}$$

for all interval vectors $[x] \subseteq D$ *and for some positive number* δ *which is independent of* $[x]$ *but which may depend on* $\tilde{x}$. *Define* $\hat{\Delta}, \hat{r}^-, \hat{r}^+$ *as* Δ, r^-, r^+, *with* γ *being replaced by* $\hat{\gamma} := \max\{\gamma, \delta\}$. *If* $\hat{\Delta} \geq 0$ *and if* $r \in [\hat{r}^-, (\hat{r}^- + \hat{r}^+)/2)$ *then the function* t *has exactly one fixed point* $x^* \in [x]^0$; $[g]([x]^0, \tilde{x}) \subseteq [x]^0$ *holds, and the iteration*

$$[x]^{k+1} \;:=\; [g]([x]^k, \tilde{x})\,, \quad k \;=\; 0\,, 1\,, \ldots\,,$$

converges to x^* *with*

$$x^* \in [x]^k \subseteq [x]^{k-1} \subseteq \cdots \subseteq [x]^0\,, \quad k \in \mathbb{N}\,.$$

Although variants of Theorem 3.1 have already been proved in [5], [15] and [19] we shortly repeat the major steps of its proof.

Proof. From $\beta < 1$ and $\Delta \geq 0$ we have $0 \leq r^- \leq r^+$.
a) is proved by b) applied to some vector $[\hat{x}]^0 = \tilde{x} + [-\hat{r}, \hat{r}]$ with $r \geq \hat{r} \in [r^-, r^+]$.
b) The inclusion $[g]([x]^0, \tilde{x}) \subseteq [x]^0$ is equivalent to

$$[g]([x]^0, \tilde{x}) - \tilde{x} \subseteq [x]^0 - \tilde{x}\,. \tag{3.6}$$

Therefore, this inclusion certainly holds for $[x]^0 := \tilde{x} + [-r, r]e$ if

$$|t(\tilde{x}) - \tilde{x}| + |t'(\tilde{x})|re + \left|[h]([x]^0, \tilde{x})\right| \;\leq\; re\,.$$

This, in turn, is true if

$$\alpha + \beta r + \gamma r^2 \;\leq\; r \quad\Longleftrightarrow\quad \gamma\,(r - r^-)\,(r - r^+) \;\leq\; 0\,. \tag{3.7}$$

Hence (3.7) is fulfilled for each $r \in [r^-, r^+]$. Since

$$t(x) \in [g]([x]^0, \tilde{x}) \subseteq [x]^0 \quad \text{for} \quad x \in [x]^0\,,$$

Brouwer's fixed point theorem guarantees that t has at least one fixed point $x^* \in [x]^0$. The relation

$$x^* \;=\; t(x^*) \in [g]([x]^k, \tilde{x}) \;=\; [x]^{k+1} \subseteq [g]([x]^{k-1}, \tilde{x}) \;=\; [x]^k\,, \quad k \;=\; 1\,, 2\,, \ldots\,,$$

holds by induction; in particular, $[x]^* := \lim_{k\to\infty} [x]^k$ exists.
c) Since $\hat{\gamma} \geq \gamma$ we have $\hat{\Delta} \leq \Delta$ and

$$r^- \;=\; \frac{2\alpha}{1 - \beta + \sqrt{\Delta}} \;\leq\; \hat{r}^- \;\leq\; \hat{r}^+ \;\leq\; r^+ \;=\; \frac{2\alpha}{1 - \beta - \sqrt{\Delta}}\,.$$

Therefore, r is contained in $[r^-, r^+]$, and b) is used to prove most of the assertions in c). In particular, x^* and $[x]^* = [g]([x]^*, \tilde{x})$ exist, and

$$d([x]^*) \;=\; d([g]([x]^*, \tilde{x})) \;\leq\; |t'(\tilde{x})|\, d([x]^*) + d([h]([x]^*, \tilde{x}))$$

by the elementary rules for the diameter mentioned in Section 2. Apply $\|\cdot\|_\infty$ to this inequality and let $d^* := \|d([x]^*)\|_\infty$. Then

$$d^* \;\leq\; \beta d^* + 2\delta \,\big\|\,|[x]^* - \tilde{x}|\,\big\|_\infty d^* \;\leq\; \beta d^* + 2\delta \,\big\|\,|[x]^0 - \tilde{x}|\,\big\|_\infty d^* \;\leq\; \beta d^* + 2r\hat{\gamma} d^* .$$

If $d^* > 0$, we obtain $1 \leq \beta + 2r\hat{\gamma}$ which yields to the contradiction

$$r \;\geq\; \frac{1-\beta}{2\hat{\gamma}} \;=\; \frac{\hat{r}^- + \hat{r}^+}{2} .$$

Therefore, $d^* = 0$, and $x^* \in [x]^*$ implies $[x]^* = [x^*, x^*]$. In particular, this proves uniqueness. □

One often chooses t as in (3.4) and

$$[h]([x], \tilde{x}) \;:=\; -\frac{1}{2}(Pf)''([x] \sqcup \tilde{x})\,([x] - \tilde{x})\,([x] - \tilde{x}) \tag{3.8}$$

with $f''(x)yz = \left(y^T \left(\frac{\partial^2 f_i(x)}{\partial x_l\, \partial x_k}\right) z\right) \in \mathbb{R}^n$ for $f(x) = (f_i(x)) \in \mathbb{R}^n$, $x \in D$ and $y, z \in \mathbb{R}^n$. For $\tilde{x} \in [x]$ the function $[g]$ is identical with the function k_2 in [5], p. 239, and k_1 in [19], p. 29. As was indicated in [15], there are cases, in which $[h]$ differs from the choice in (3.8).

If one chooses t according to (3.4) without knowing about the regularity of the matrix P then one can verify its nonsingularity by checking the assumption $\beta < 1$ of Theorem 3.1. If it is true then the spectral radius of $t'(\tilde{x}) = I - Pf'(\tilde{x})$ is less than one; hence P and $f'(\tilde{x})$ cannot be singular.

4. Enclosures for generalized singular values

We now will present our method for verifying and enclosing generalized singular values and the corresponding vectors from the two decompositions (1.10), (1.11). Let u^i, v^i, x^i be the i-th columns of U, V and X, respectively. We want to express these vectors and the corresponding generalized singular values (c_i, s_i) as a zero of some function f. To this end we multiply (1.10) by U and (1.11) by V in order to obtain

$$AX \;=\; U\Sigma_A , \tag{4.1}$$

$$BX \;=\; V\Sigma_B . \tag{4.2}$$

By transposing (1.10), (1.11) we get

$$X^T A^T U \;=\; \Sigma_A^T , \tag{4.3}$$

$$X^T B^T V \;=\; \Sigma_B^T . \tag{4.4}$$

Let $\hat{\Sigma}_A := \begin{pmatrix} C \\ O \end{pmatrix} \in \mathbb{R}^{q \times n}$, $\hat{\Sigma}_B := \begin{pmatrix} S \\ O \end{pmatrix} \in \mathbb{R}^{p \times n}$. Note that $\hat{\Sigma}_A$, $\hat{\Sigma}_B$ differ from Σ_A, Σ_B by the number of rows. Multiply (4.3) by $\hat{\Sigma}_B$ and (4.4) by $\hat{\Sigma}_A$. Then

$$X^T A^T U \hat{\Sigma}_B = \Sigma_A^T \hat{\Sigma}_B = CS, \tag{4.5}$$

$$X^T B^T V \hat{\Sigma}_A = \Sigma_B^T \hat{\Sigma}_A = SC = CS, \tag{4.6}$$

where the last equality holds since C, S are $n \times n$ diagonal matrices. Multiplying (4.5) and (4.6) by $\left(X^{-1}\right)^T$ and equating the left-hand sides results in

$$A^T U \hat{\Sigma}_B = B^T V \hat{\Sigma}_A . \tag{4.7}$$

From (4.1), (4.2) and (4.7) we obtain the following set of equations:

$$Ax^i = c_i u^i , \tag{4.8}$$

$$Bx^i = s_i v^i , \tag{4.9}$$

$$s_i A^T u^i = c_i B^T v^i , \tag{4.10}$$

which we complete by

$$\left(u^i\right)^T u^i = 1 , \tag{4.11}$$

$$c_i^2 + s_i^2 = 1 \quad (c_i, s_i \geq 0) . \tag{4.12}$$

Equation (4.11) follows from the orthogonality $U^T U = I$, and (4.12) is part of the normalization (1.12).

From (4.10) we get

$$s_i (x^i)^T A^T u^i = c_i \left(x^i\right)^T B^T v^i \iff s_i \left(Ax^i\right)^T u^i = c_i \left(Bx^i\right)^T v^i$$

which yields to

$$c_i s_i \left(u^i\right)^T u^i = c_i s_i \left(v^i\right)^T v^i$$

by (4.8) and (4.9). Using (4.11) we obtain

$$c_i s_i \left(1 - \left(v^i\right)^T v^i\right) = 0$$

which implies the corresponding normalization

$$\left(v^i\right)^T v^i = 1 \tag{4.13}$$

for v, provided that $c_i s_i \neq 0$. Note that in the case $c_i s_i = 0$ the equations (4.8) – (4.12) do not necessarily imply (4.13). If $c_i = 0$ then $s_i = 1$, $Ax^i = 0$ and $Bx^i = v^i$. Hence x^i, v^i can be replaced by any pair τx^i, τv^i ($\tau \in \mathbb{R}$) in the solution vector of (4.8) – (4.12). Similarly, in the case $s_i = 0$, $c_i = 1$ any vector from the null space of B^T can be chosen for v^i.

Therefore we will assume $c_i s_i \neq 0$ in the sequel.

Let us now start conversely with (4.8) – (4.12). Then (4.8) and (4.11) guarantee $x^i \neq 0$, and (4.8) – (4.10) yield to

$$\begin{aligned}(s_i^2 A^T A - c_i^2 B^T B)x^i &= s_i^2 c_i A^T u^i - c_i^2 s_i B^T v^i \\ &= s_i c_i (s_i A^T u^i - c_i B^T v^i) \\ &= 0\end{aligned} \tag{4.14}$$

whence $\det(s_i^2 A^T A - c_i^2 B^T B) = 0$. This proves (c_i, s_i) to be a generalized singular value of (A, B). Assume now that

$$z^i = \left((u^i)^T, (v^i)^T, (x^i)^T, c_i, s_i\right)^T \quad \text{and} \quad z^j = \left((u^j)^T, (v^j)^T, (x^j)^T, c_j, s_j\right)^T$$

are two solutions of the system

$$Ax = cu, \quad Bx = sv, \quad sA^T u = cB^T v, \quad u^T u = 1, \quad c^2 + s^2 = 1 \quad (c, s \geq 0)$$

with $(c_i, s_i) \neq (c_j, s_j)$. We want to show that $(u^i)^T u^j = (v^i)^T v^j = 0$. To this end multiply (4.14) by $(x^j)^T$ from the left. Using (4.8), (4.9) yields to

$$0 = s_i^2 (x^j)^T A^T A x^i - c_i^2 (x^j)^T B^T B x^i = s_i^2 c_i c_j (u^j)^T u^i - c_i^2 s_i s_j (v^j)^T v^i$$

whence

$$s_i c_j (u^j)^T u^i - s_j c_i (v^j)^T v^i = 0. \tag{4.15}$$

Exchanging the roles of i and j results in

$$s_j c_i (u^j)^T u^i - s_i c_j (v^j)^T v^i = 0. \tag{4.16}$$

Multiply (4.15) by $s_i c_j$ and (4.16) by $s_j c_i$ and subtract both equations in order to get

$$(s_i^2 c_j^2 - s_j^2 c_i^2)(u^j)^T u^i = 0$$

or, by (4.12),

$$0 = (c_j^2 - c_i^2)(u^j)^T u^i = (s_i^2 - s_j^2)(u^j)^T u^i.$$

Since $(c_i, s_i) \neq (c_j, s_j)$ the orthogonality $(u^i)^T u^j = 0$ follows, and (4.15) or (4.16) proves $(v^i)^T v^j = 0$. Therefore, each zero z^* of the function

$$f : \begin{cases} \mathbb{R}^{p+q+n+2} & \longrightarrow & \mathbb{R}^{p+q+n+2} \\ z = (u^T, v^T, x^T, c, s)^T & \longmapsto & f(z) := \begin{pmatrix} Ax - cu \\ Bx - sv \\ sA^T u - cB^T v \\ 1 - u^T u \\ 1 - c^2 - s^2 \end{pmatrix} \end{cases} \tag{4.17}$$

with $c^*s^* \neq 0$ and $c^* \geq 0$, $s^* \geq 0$ has the following property:

1. (c^*, s^*) is a generalized singular value of (A, B). This property still holds if our general assumptions (1.8), p, $q \geq n$ and $c^*s^* \neq 0$ are not fulfilled, as the deduction of (4.14) shows.

2. If (c^*, s^*) is a simple generalized singular value, then u^*, v^*, x^* are columns of U, V, and X, respectively, in the generalized singular value decomposition (1.10), (1.11).

3. If (c^*, s^*) is a multiple generalized singular value, then u^*, v^* belong to the subspace which is spanned by the corresponding columns of U and V respectively. This can be seen as follows:

Since x^* solves (4.14) it can be represented as a linear combination $x^* = \sum_j \alpha_{i_j} x^{i_j}$ of those vectors x^{i_j} which are part of the solutions of (4.8) – (4.12) arising from the generalized singular value decomposition of (A, B) and belonging to the multiple generalized singular value (c^*, s^*). By (4.8), (4.9) the vectors u^*, v^* can then be represented as a linear combination of the corresponding columns of U and V, respectively, with the same coefficients α_{i_j} as in x^*.

We remark that the vectors u^*, v^* need not coincide with the columns themselves as can be seen from the example $A = B = I \in \mathbb{R}^{2\times 2}$ for which $(c^*, s^*) = (\alpha, \alpha)$ with $\alpha := \frac{1}{\sqrt{2}}$ is the unique generalized singular value which is a double one. The zeros $z^1 = (1,0,1,0,\alpha,0,\alpha,\alpha)^T$ and $z^2 = \alpha\,(1,1,1,1,\alpha,\alpha,1,1)^T$ of f contain the two linearly independent vectors $u^1 = (1,0)^T$, $u^2 = (\alpha,\alpha)^T$ which are certainly not orthogonal and which therefore cannot both be a column of the orthogonal matrix U.

We note that a zero $\hat{z}$ of f with $\hat{c} < 0$ yields at once to a zero z^* of f with $c^* > 0$ by replacing $\hat{u}$, $\hat{c}$ in $\hat{z}$ by $u^* := -\hat{u}$ and $c^* := -\hat{c}$. One can proceed similarly if $\hat{s} < 0$.

Let now $\tilde{z} := \left(\tilde{u}^T, \tilde{v}^T, \tilde{x}^T, \tilde{s}, \tilde{c}\right)^T$ be an approximation of a zero z^* of f and let $t(z) := z - Pf(z)$ with a nonsingular $(p+q+n+2)\times(p+q+n+2)$ matrix P. Expand the function t in a Taylor series at $z = \tilde{z}$. Then

$$
\begin{aligned}
t(z) &= t(\tilde{z} + (z - \tilde{z})) \\
&= \tilde{z} + (z - \tilde{z}) - Pf(\tilde{z} + (z - \tilde{z})) \\
&= \tilde{z} + (z - \tilde{z}) - P \begin{pmatrix} A(\tilde{x} + (x - \tilde{x})) - (\tilde{c} + (c - \tilde{c}))(\tilde{u} + (u - \tilde{u}) \\ B(\tilde{x} + (x - \tilde{x})) - (\tilde{s} + (s - \tilde{s}))(\tilde{v} + (v - \tilde{v})) \\ (\tilde{s} + (s - \tilde{s}))A^T(\tilde{u} + (u - \tilde{u})) - (\tilde{c} + (c - \tilde{c}))B^T(\tilde{v} + (v - \tilde{v})) \\ 1 - (\tilde{u} + (u - \tilde{u}))^T(\tilde{u} + (u - \tilde{u})) \\ 1 - (\tilde{c} + (c - \tilde{c}))^2 - (\tilde{s} + (s - \tilde{s}))^2 \end{pmatrix} \\
&= \tilde{z} - Pf(\tilde{z}) + (I - PR)(z - \tilde{z}) + h(z, \tilde{z})
\end{aligned}
$$

with the $(p+q+n+2)\times(p+q+n+2)$ matrix

$$
R := \begin{pmatrix} -\tilde{c}I_p & O & A & -\tilde{u} & 0 \\ O & -\tilde{s}I_q & B & 0 & -\tilde{v} \\ \tilde{s}A^T & -\tilde{c}B^T & O & -B^T\tilde{v} & A^T\tilde{u} \\ -2\tilde{u}^T & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2\tilde{c} & -2\tilde{s} \end{pmatrix} \tag{4.18}
$$

and the vector

$$(4.19)\quad h(z,\tilde{z}) := P\begin{pmatrix} (c-\tilde{c})(u-\tilde{u}) \\ (s-\tilde{s})(v-\tilde{v}) \\ (c-\tilde{c})B^T(v-\tilde{v}) - (s-\tilde{s})A^T(u-\tilde{u}) \\ (u-\tilde{u})^T(u-\tilde{u}) \\ (c-\tilde{c})^2 + (s-\tilde{s})^2 \end{pmatrix} \in \mathbb{R}^{p+q+n+2}.$$

Define the interval function $[g]$ by

$$(4.20)\quad [g]([z],\tilde{z}) := \tilde{z} - Pf(\tilde{z}) + (I-PR)([z]-\tilde{z}) + h([z],\tilde{z}) \in \boldsymbol{IR}^{p+q+n+2}$$

where $h([z],\tilde{z})$ is the interval arithmetic evaluation of $h(z,\tilde{z})$. It is clear that the analogue of (3.2) holds. We want to show that $[h] = h([z],\tilde{z})$ fulfills the properties (3.3) and (3.5). With the rules in Section 2 for the absolute values and the diameter we get

$$\begin{aligned} |h([z],\tilde{z})| &\le |P|\,\hat{e}\,\big\||[z]-\tilde{z}|\big\|_\infty^2, \\ d(h([z],\tilde{z})) &\le 2|P|\,\hat{e}\,\big\||[z]-\tilde{z}|\big\|_\infty \big\|d([z])\big\|_\infty, \end{aligned}$$

with

$$\hat{e} := \begin{pmatrix} e_p \\ e_q \\ |A^T|\,e_p + |B^T|\,e_q \\ p \\ 2 \end{pmatrix} \in \mathbb{R}^{p+q+n+2}.$$

Hence (3.3), (3.5) are valid with

$$(4.21)\quad \gamma := \delta := \big\|\,|P|\hat{e}\big\|_\infty \le \left\| |P| \begin{pmatrix} e_p \\ e_q \\ \left(\|A^T\|_\infty + \|B^T\|_\infty\right) e_n \\ p \\ 2 \end{pmatrix} \right\|_\infty$$

or with

$$(4.22)\quad \gamma := \delta := \big\|\,|P|\,\big\|_\infty \max\left\{2, p, \|A^T\|_\infty + \|B^T\|_\infty\right\}.$$

We are now ready to apply Theorem 3.1 to our situation. The results are collected in the following Theorem 4.1. Before formulating it we remark that up to now we generally had assumed

$$(4.23)\quad (1.8),\quad p,q \ge n,\quad \text{and}\quad c^*s^* \ne 0.$$

The first two assumptions were necessary in order to guarantee the representation (1.10), (1.11) in Theorem 1.3 which led us to the crucial function f in (4.17). By means of the third one we showed the normalization (4.13). None of these assumptions will be needed to prove Theorem 4.1 since it only deals with zeros

$$z^* = \left((u^*)^T, (v^*)^T, (x^*)^T, c^*, s^*\right)^T$$

of f from (4.17). Therefore, we will formulate it without requiring (4.23). But we emphasize that even if (4.23) is false, Property 1 below (4.17) guarantees that the last two components c^*, s^* of z^* form a generalized singular value of (A, B). This remark is important since normally we do not know in advance whether (1.8) holds. If (4.23) is not true, a generalized singular value decomposition analogously to (1.10), (1.11) need not exist as we showed by a simple example preceding Corollary 1.4. If it exists, however, and if (c^*, s^*) forms a simple generalized singular value then u^* is a column of U and v^*, x^* coincide with columns of V, X up to a multiplicative factor which is equal to one provided that $c^*s^* \neq 0$. If $c^*s^* = 0$ then $\|v^*\|_2 = 1$ may not be true. In this case one can normalize v^* by $v^*/\|v^*\|_2$ provided that $v^* \neq 0$. If $c^* = 0$ one must, in addition, replace x^* by $x^*/\|v^*\|_2$. This does not influence the representation (1.10) since then $Ax^* = 0$.

Theorem 4.1. *Let $A \in \mathbb{R}^{p\times n}$, $B \in \mathbb{R}^{q\times n}$. Let P be some nonsingular real $(p+q+n+2)\times(p+q+n+2)$ matrix, let $\tilde{z} = \left(\tilde{u}^T, \tilde{v}^T, \tilde{x}^T, \tilde{c}, \tilde{s}\right)^T \in \mathbb{R}^{p+q+n+2}$ and let f be given as in (4.17). Let the assumptions of Theorem 3.1 be fulfilled for $t(z) := z - Pf(z)$, $[g]$, h from (4.20), (4.19), and γ from (4.21) or (4.22), and define $r^{\pm}$ as in that theorem. Then the following assertions hold for $[z]^0 := \tilde{z} + [-r, r]e \in \boldsymbol{IR}^{p+q+n+2}$:*

a) *If $r \geq r^-$ then f has at least one zero $z^* \in [z]^0$. The iteration*

$$[z]^{k+1} := [g]\left([z]^k, \tilde{z}\right) \cap [z]^k, \quad k = 0, 1, \ldots,$$

converges to some interval vector $[z]^*$ with

$$z^* \in [z]^* \subseteq [z]^k \subseteq [z]^{k-1} \subseteq \cdots \subseteq [z]^0, \quad k \in \mathbb{N}.$$

b) *If $r \in [r^-, r^+]$ then f has at least one zero $z^* \in [z]^0$. In addition, $[g]\left([z]^0, \tilde{z}\right) \subseteq [z]^0$ holds and the iteration*

$$[z]^{k+1} := [g]\left([z]^k, \tilde{z}\right), \quad k = 0, 1, \ldots,$$

converges to some interval vector $[z]^*$ with

$$z^* \in [z]^* \subseteq [z]^k \subseteq [z]^{k-1} \subseteq \cdots \subseteq [z]^0, \quad k \in \mathbb{N}.$$

c) *If $r \in [r^-, (r^- + r^+)/2)$ then the function f has exactly one zero $z^* \in [z]^0$; $[g]\left([z]^0, \tilde{z}\right) \subseteq [z]^0$ holds, and the iteration*

$$[z]^{k+1} := [g]\left([z]^k, \tilde{z}\right), \quad k = 0, 1, \ldots,$$

converges to z^* with

$$z^* \in [z]^k \subseteq [z]^{k-1} \subseteq \cdots \subseteq [z]^0, \quad k \in \mathbb{N}.$$

Note that we could use $r^{\pm}$ in Theorem 4.1 c) since $\gamma = \delta$ (cf. (4.21), (4.22)) implies $r^{\pm} = \hat{r}^{\pm}$ in Theorem 3.1.

In order to fulfill the assumption in Theorem 4.1 for β, one normally chooses the matrix P as $P \approx R^{-1}$. Then $\beta \approx 0$; in particular, $\beta < 1$. In this case P is nonsingular according to a remark after the proof of Theorem 3.1.

Due to the choice of $P \approx R^{-1}$, it is interesting to know whether all the matrices near R are invertible. Together with the continuity of the matrix inversion the following Theorem 4.2 guarantees this, provided that (c^*, s^*) is a simple generalized singular value and provided that $\tilde{z}$ comes sufficiently close to z^*.

But Theorem 4.2 can also be used to guarantee the simplicity of (c^*, s^*) itself, provided that $\beta < 1$ is known and provided that $\tilde{z}$ approximates z^* very well. In this case R^{-1} exists and the continuity of the matrix inversion implies the nonsingularity of the matrix R^* in the subsequent theorem.

Note that a good approximation $\tilde{z} \approx z^*$ also guarantees $\Delta \geq 0$ in Theorem 3.1.

Theorem 4.2. *Let $A \in \mathbb{R}^{p\times n}$, $B \in \mathbb{R}^{q\times n}$ and f as in (4.17). Let $p, q \geq n$ and let (1.8) hold. If $z^* = \left((u^*)^T, (v^*)^T, (x^*)^T, c^*, s^*\right)^T$ is a zero of f with $c^*s^* \neq 0$ then the following statements are equivalent:*

a) *The pair (c^*, s^*) is a simple generalized singular value of (A, B).*

b) *The real $(p+q+n+2)\times(p+q+n+2)$ matrix*

$$R^* := \begin{pmatrix} -c^*I_p & O & A & -u^* & 0 \\ O & -s^*I_q & B & 0 & -v^* \\ s^*A^T & -c^*B^T & O & -B^Tv^* & A^Tu^* \\ -2(u^*)^T & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2c^* & -2s^* \end{pmatrix} \tag{4.24}$$

is nonsingular.

Proof. a) $\Rightarrow$ b): Let (c^*, s^*) be a simple generalized singular value of (A, B) and assume that R^* is singular. Then there exists a vector $w = \left(w_1^T, w_2^T, w_3^T, w_4, w_5\right)^T \in \mathbb{R}^{p+q+n+2}\backslash\{0\}$ with block vectors w_i such that $R^*w = 0$. Thus the following system holds:

$$0 = -c^*w_1 + Aw_3 - u^*w_4\,, \tag{4.25}$$

$$0 = -s^*w_2 + Bw_3 - v^*w_5\,, \tag{4.26}$$

$$0 = s^*A^Tw_1 - c^*B^Tw_2 - B^Tv^*w_4 + A^Tu^*w_5\,, \tag{4.27}$$

$$0 = (u^*)^Tw_1\,, \tag{4.28}$$

$$0 = c^*w_4 + s^*w_5\,. \tag{4.29}$$

Multiply the equations (4.25), (4.26) and (4.27) by $(u^*)^T$, $c^*(v^*)^T$ and $(x^*)^T$, respectively, and use (4.11), (4.28), and (4.10), (4.13), and (4.8), (4.9), (4.11), (4.13), (4.28) respectively, in order to get

$$0 = (u^*)^TAw_3 - w_4\,, \tag{4.30}$$

$$(4.31)\qquad \begin{cases} 0 &= -c^*s^*(v^*)^T w_2 + c^*(v^*)^T B w_3 - c^* w_5 \\ &= -c^*s^*(v^*)^T w_2 + s^*(u^*)^T A w_3 - c^* w_5\,, \end{cases}$$

$$(4.32)\qquad \begin{cases} 0 &= s^*(x^*)^T A^T w_1 - c^*(x^*)^T B^T w_2 - (x^*)^T B^T v^* w_4 + (x^*)^T A^T u^* w_5 \\ &= c^*s^*(u^*)^T w_1 - c^*s^*(v^*)^T w_2 - s^* w_4 + c^* w_5 \\ &= -c^*s^*(v^*)^T w_2 - s^* w_4 + c^* w_5\,. \end{cases}$$

From (4.31), (4.32) and (4.30) we deduce

$$(4.33)\qquad \begin{aligned} 0 &= s^*(u^*)^T A w_3 - c^* w_5 + s^* w_4 - c^* w_5 \\ &= s^* w_4 - c^* w_5 + s^* w_4 - c^* w_5 \\ &= 2(s^* w_4 - c^* w_5)\,. \end{aligned}$$

Together with (4.12) and (4.29) this implies $w_4 = w_5 = 0$, whence from (4.30) – (4.32) we get

$$(4.34)\qquad 0 = (v^*)^T w_2 = (u^*)^T A w_3 = (v^*)^T B w_3\,.$$

In addition, (4.25) – (4.27) reduce to

$$(4.35)\qquad A w_3 = c^* w_1\,,$$

$$(4.36)\qquad B w_3 = s^* w_2\,,$$

$$(4.37)\qquad s^* A^T w_1 = c^* B^T w_2\,,$$

which implies

$$(4.38)\qquad \left((s^*)^2 A^T A - (c^*)^2 B^T B\right) w_3 = (s^*)^2 A^T c^* w_1 - (c^*)^2 B^T s^* w_2 = 0\,.$$

Since we assumed (c^*, s^*) to be a simple generalized singular value, we must have $w_3 = \tau x^*$ with some real number $\tau \neq 0$. Using (4.9) we get from (4.36)

$$(4.39)\qquad s^* w_2 = B w_3 = \tau B x^* = \tau s^* v^*\,.$$

Multiplying (4.39) by $(v^*)^T$ and taking into account (4.13) and (4.34) yields to $\tau s^* = 0$, whence the contradiction $\tau = 0$ follows. Therefore, $w = 0$ in contrast to our assumption.

b) $\Rightarrow$ a): Let R^* be nonsingular and assume that (c^*, s^*) is a generalized singular value which is not simple. Taking into account Theorem 1.3, there are two different zeros z^*, $\hat{z}$ of f such that $(c^*, s^*) = (\hat{c}, \hat{s})$ and $(u^*)^T \hat{u} = 0$. (If $(u^*)^T \hat{u} \neq 0$ one can choose an appropriate normalized linear combination $\check{u} := \zeta u^* + \eta \hat{u}$ of u^*, $\hat{u}$ such that $(u^*)^T \check{u} = 0$ holds.) This implies $R^* \left(\hat{u}^T, \hat{v}^T, \hat{x}^T, 0, 0\right)^T = 0$ whence $\hat{u} = 0$ in contrast to $\hat{u}^T \hat{u} = 1$. □

We now compare the method of this paper with two methods discussed in [11]. There

it was started using the equations

$$
(4.40) \qquad \begin{cases} Ax &= c_1 u, \\ Bx &= sv, \\ sA^T u &= c_2 B^T v, \\ u^T u &= 1, \\ v^T v &= 1, \\ c_1 c_2 + s^2 &= 1 \end{cases}
$$

which differ from (4.8) – (4.12) by the two unknowns c_1, c_2 and by the normalization $v^T v = 1$. However, the normalization can also be deduced from the set (4.8) – (4.12); see (4.13). Adding the equation $v^T v = 1$ to (4.8) – (4.12) implies that the number of scalar equations is increased by one. Therefore in order to have still the same number of unknowns and (scalar) equations one has to introduce an additional unknown. This was done in [11] by replacing c by c_1, c_2. Note, however, that each solution of (4.40) satisfies

$$
(4.41) \qquad sc_1 = sc_1 u^T u = su^T Ax = c_2 v^T Bx = c_2 sv^T v = c_2 s,
$$

whence $c_1 = c_2$ provided that $s \neq 0$. Note that $c_1 = c_2 = 0$ is possible here while we assumed $cs \neq 0$ in the method of this paper thus decreasing the dimension of the vectors and matrices. The generalized singular value $(c, s) = (1, 0)$ can also be handled by (4.40) when interchanging the role of A and B. Assuming $s \neq 0$ the solutions of (4.40) are precisely the zeros of the function

$$
(4.42) \quad f : \begin{cases} \mathbb{R}^{p+q+n+3} & \longrightarrow & \mathbb{R}^{p+q+n+3} \\ z = \left(u^T, v^T, x^T, c_1, c_2, s\right)^T & \longmapsto & f(z) := \begin{pmatrix} Ax - c_1 u \\ Bx - sv \\ sA^T u - c_2 B^T v \\ \frac{1}{p}(1 - u^T u) \\ \frac{1}{q}(1 - v^T v) \\ \frac{1}{2}(1 - c_1 c_2 - s^2) \end{pmatrix} \end{cases}
$$

which was used in [11]. It is obvious that the same steps can be done as for the method of this paper ending up with similar results. The factors $\frac{1}{p}$, $\frac{1}{q}$ and $\frac{1}{2}$ are scaling factors. They have been introduced in order to decrease the value of the factors γ and δ in (4.21). It can easily be checked that these factors now read

$$
(4.43) \qquad \gamma := \delta := \left\| |P| \begin{pmatrix} e_p \\ e_q \\ |A^T| e_p + |B^T| e_q \\ 1 \\ 1 \\ 1 \end{pmatrix} \right\|_\infty .
$$

But note that the scaling influences the matrices P, R and the vector $h(z,\tilde{z})$. We remark that similar scaling factors could also be introduced in (4.17).

Clearly, Theorem 3.1 applies again. The results are then identical with those in [11]. We leave it to the reader to formulate them. The matrix R^* from Theorem 4.2 is now a $(p+q+n+3)\times(p+q+n+3)$ matrix and reads

$$R^* := \begin{pmatrix} -c_1^* I_p & O & A & -u^* & 0 & 0 \\ O & -s^* I_q & B & 0 & 0 & -v^* \\ s^* A^T & -c_2^* B^T & O & 0 & -B^T v^* & A^T u^* \\ -\frac{2}{p}(u^*)^T & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{2}{q}(v^*)^T & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2}c_2^* & -\frac{1}{2}c_1 & -s^* \end{pmatrix}, \tag{4.44}$$

where we assume $s^* \neq 0$. Theorem 4.2 holds analogously, as was shown for the implication a) $\Rightarrow$ b) in [11]. The proof can be performed analogously to that in this paper.

A second variant in [11] uses the transposed inverse $Y := \left(X^{-1}\right)^T$. Instead of combining (4.3) and (4.4) into (4.7) one rewrites (4.3), (4.4) as

$$A^T U = Y\Sigma_A^T, \tag{4.45}$$

$$B^T U = Y\Sigma_B^T, \tag{4.46}$$

and starts with the equations

$$\begin{cases} Ax = c_1 u, \\ Bx = sv, \\ A^T u = c_2 y, \\ B^T v = sy, \\ u^T u = 1, \\ v^T v = 1, \\ c_1 c_2 + s^2 = 1, \end{cases} \tag{4.47}$$

in which the numbers of equations and unknowns are now increased to $p+q+2n+3$. The third and the fourth equation imply $sA^T u = sc_2 y = c_2 B^T v$ which is the third equation of (4.40). Therefore, as in (4.41) one obtains $c_1 = c_2 =: c$ for each solution of (4.47), provided that $s \neq 0$. In addition $x^T y = 1$ holds in this case because of

$$\begin{aligned} x^T y &= \left(c^2 + s^2\right) x^T y \\ &= x^T \left(c(cy) + s(sy)\right) \\ &= x^T \left(cA^T u + sB^T v\right) \\ &= c^2 u^T u + s^2 v^T v \\ &= c^2 + s^2 \\ &= 1. \end{aligned}$$

Assuming $s \neq 0$ the solutions of (4.47) are the zeros of the function

$$(4.48)\quad f : \begin{cases} \mathbb{R}^{p+q+2n+3} & \longrightarrow & \mathbb{R}^{p+q+2n+3} \\ z = (u^T, v^T, x^T, c_1, c_2, s, y^T) & \longmapsto & f(z) := \begin{pmatrix} Ax - c_1 u \\ Bx - sv \\ A^T u - c_2 y \\ B^T v - sy \\ \frac{1}{p}(1 - u^T u) \\ \frac{1}{q}(1 - v^T v) \\ \frac{1}{2}(1 - c_1 c_2 - s^2) \end{pmatrix} \end{cases}$$

with which one can again construct the functions $t(x) := x - Pf(x)$, h and $[g]$ as above. Theorem 3.1 yields at once to the results in [11]. The factors γ and δ in (4.21) now read

$$(4.49)\qquad \gamma := \delta := \| |P| \|_\infty .$$

The matrix R^* from Theorem 4.2 is here a $(p+q+2n+3) \times (p+q+2n+3)$ matrix which is given by

$$(4.50)\quad R^* := \begin{pmatrix} -c_1^* I_p & O & A & -u^* & 0 & 0 & O \\ O & -s^* I_q & B & 0 & 0 & -v^* & O \\ A^T & O & O & 0 & -y^* & 0 & -c_2^* I_n \\ O & B^T & O & 0 & 0 & -y^* & -s^* I_n \\ -\frac{2}{p}(u^*)^T & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{2}{q}(v^*)^T & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2}c_2^* & -\frac{1}{2}c_1^* & -s^* & 0 \end{pmatrix}.$$

To avoid confusions we rename the matrix R^* in (4.44) by $\hat{R}^*$. The determinant of R^* from (4.50) is connected to $\det \hat{R}^*$ by

$$(4.51)\qquad |\det R^*| = \left|(-s^*)^n \det \hat{R}^*\right|$$

as was shown in [11]. It can be seen by adding the $(-c^*)$ multiple of the fourth block row of R^* to the s^* multiple of the third block row. (Here we used again $c^* := c_1^* = c_2^*$.) This cancels $-c_2^* I$ at the end of the third block row. Evaluating now the determinant along the last n columns and taking into account $c^* y^* = A^T u^*$, $s^* y^* = B^T v^*$ yields to (4.51). Therefore, $\hat{R}^*$ is nonsingular if and only if R^* has this property. Hence Theorem 4.2 holds again.

5. A numerical example

We consider here an example from [13], p. 16, which was also used in [11]:

$$A = \begin{pmatrix} 0.1376 & 0.4087 & 0.1593 & 0.4308 & 0.4163 \\ 0.9667 & 0.6246 & 0.3384 & 0.8397 & 0.8029 \\ 0.8286 & 0.0661 & 0.9111 & 0.7495 & 0.1577 \\ 0.0728 & 0.3485 & 0.8560 & 0.8068 & 0.2644 \\ 0.5501 & 0.9198 & 0.0080 & 0.7910 & 1.0323 \\ 0.6498 & 0.2725 & 0.3599 & 0.5350 & 0.3801 \end{pmatrix},$$

$$B = \begin{pmatrix} 0.4087 & 0.1593 & 0.6594 & 0.4302 & 0.3516 \\ 0.6246 & 0.3383 & 0.6591 & 0.9342 & 0.9038 \\ 0.0661 & 0.9112 & 0.6898 & 0.1931 & 0.1498 \\ 0.2112 & 0.8150 & 0.7983 & 0.3406 & 0.2803 \end{pmatrix}.$$

Since B has less rows than columns the assumption $q \geq n$ of Theorem 1.3 is certainly not fulfilled. Furthermore, we do not know whether (1.8) is true. Therefore, Theorem 1.3 is not applicable, even if we supplement B with a fifth row which consist only of zeros. Nevertheless a generalized singular value decomposition (1.10), (1.11) can exist. We assume this for the moment.

Since the matrices A, B are not representable exactly by machine numbers we multiply them by 10 000 in order to get rid of this problem. Apparently, this scaling does not change the generalized singular values, but it influences the matrix X in (1.10), (1.11) which has to be multiplied by $\frac{1}{10\,000}$. In order to get a good approximation $\tilde{u}^i$, $\tilde{v}^i$, $\tilde{x}^i$, $\tilde{c}_i$, $\tilde{s}_i$ for the $i-$th column of U, V, and X, respectively, and for the $i-$th generalized singular value (c_i, s_i) one can use the LAPACK driver routine SGGSVD which is based on an algorithm described in [8], [9] and [17]; cf. [7], pp. 22 – 24 and pp. 204 – 206. Here, we adopt the approximations given in [11] which result from a maximum of three cycles of the method in [17] for each index i. For the generalized singular values they read

i	$\tilde{c}_i$		$\tilde{s}_i$	
1	9.570 592 041 847	$E-1$	2.898 925 312 698	$E-1$
2	1.465 216 233 701	$E-4$	9.999 999 892 650	$E-1$
3	1.793 178 754 891	$E-5$	9.999 999 998 390	$E-1$
4	9.999 999 993 410	$E-1$	3.628 796 180 383	$E-5$
5	1.000 000 000 000	$E+0$	1.719 618 144 959	$E-27$

For $i = 1$ the approximations of the corresponding vectors are

$$\tilde{u}^1 = \begin{pmatrix} -4.011\,110\,453\,219\,E-1 \\ -1.177\,826\,718\,754\,E-1 \\ 3.394\,476\,862\,897\,E-1 \\ -4.853\,318\,178\,425\,E-1 \\ -6.853\,583\,334\,942\,E-1 \\ 6.891\,604\,385\,773\,E-2 \end{pmatrix}, \quad \tilde{v}^1 = \begin{pmatrix} 7.661\,493\,109\,139\,E-1 \\ -5.975\,833\,109\,085\,E-1 \\ -2.363\,304\,068\,684\,E-1 \\ -7.573\,552\,568\,414\,E-3 \end{pmatrix},$$

$$\tilde{x}^1 = \frac{1}{10\,000} \begin{pmatrix} 7.948\,534\,941\,564\,E-1 \\ -3.765\,662\,673\,654\,E-1 \\ 6.165\,784\,424\,092\,E-1 \\ -1.202\,198\,346\,284\,E+0 \\ 1.929\,618\,810\,567\,E-1 \end{pmatrix}.$$

For the quantities in Theorem 3.1 and 4.1 we obtained

$$\begin{aligned} \alpha &\in [1.614\,862\,135\,784\,14\overset{8}{}_{6}\,E-7], & \beta &\in [1.79\,E-12, 1.81\,E-12], \\ \gamma &\in [5.350\,007\,052\,147\,07\overset{5}{}_{3}\,E+1], & \Delta &\in [9.999\,654\,419\,011\overset{5}{}_{3}\,E-1], \\ r^- &\leq 1.614\,876\,088\,063\,078\,E-7, & r^+ &\geq 1.869\,140\,265\,896\,325\,E-2. \end{aligned}$$

Here, $[1.614\,862\,135\,784\,14\overset{8}{}_{6}\,E-7]$ denotes

$$[1.614\,862\,135\,784\,146\,E-7\ ,\ 1.614\,862\,135\,784\,148\,E-7]\ .$$

We used $r := 1.614\,876\,088\,063\,078\,E-7 \in [r^-, r^+]$ in order to compute $[z]^0 = \tilde{z} + [-r, r]e$ and PASCAL-XSC as programming language running on a workstation HP 715/100. Iterating twice according to Theorem 4.1 we got the following enclosures for c_1, s_1, u^1, v^1, x^1:

$$\begin{aligned} [c_1] &= [9.570\,592\,041\,841\,63\overset{6}{}_{4}\,E-1], \\ [s_1] &= [2.898\,925\,312\,704\,27\overset{1}{}_{0}\,E-1], \end{aligned}$$

$$[u^1] = \begin{pmatrix} [\,-4.011\,108\,838\,356\,7\overset{29}{}_{31}\,E-1\,] \\ [\,-1.177\,828\,179\,159\,8\overset{89}{}_{90}\,E-1\,] \\ [\,\ 3.394\,477\,797\,779\,47\overset{4}{}_{3}\,E-1\,] \\ [\,-4.853\,319\,058\,148\,42\overset{4}{}_{6}\,E-1\,] \\ [\,-6.853\,582\,905\,535\,60\overset{0}{}_{2}\,E-1\,] \\ [\,\ 6.891\,608\,116\,971\,8\overset{51}{}_{49}\,E-2\,] \end{pmatrix},$$

$$[v^1] = \begin{pmatrix} [\ 7.661\,493\,109\,070\ \frac{800}{798}\ E-1\ ] \\ [\ -5.975\,833\,109\,169\ 76\frac{3}{5}\ E-1\ ] \\ [\ -2.363\,304\,068\,704\ \frac{099}{100}\ E-1\ ] \\ [\ -7.573\,552\,573\,095\ 4\frac{58}{61}\ E-3\ ] \end{pmatrix},$$

$$[x^1] = \begin{pmatrix} [\ 7.948\,535\,396\,408\ 3\frac{30}{27}\ E-5\ ] \\ [\ -3.765\,661\,076\,084\ 87\frac{6}{7}\ E-5\ ] \\ [\ 6.165\,781\,424\,226\ 27\frac{4}{2}\ E-5\ ] \\ [\ -1.202\,197\,947\,052\ 54\frac{8}{9}\ E-4\ ] \\ [\ 1.929\,614\,108\,910\ 59\frac{8}{7}\ E-5\ ] \end{pmatrix}.$$

We remark that the two variants in [11] produce essentially the same inclusions. Similarly, we get

i	$[c_i]$	$[s_i]$
1	$[9.570\,592\,041\,841\,63\frac{6}{4}\ E-1]$	$[2.898\,925\,312\,704\,27\frac{1}{0}\ E-1]$
2	$[1.465\,216\,233\,516\,22\frac{6}{4}\ E-4]$	$[9.999\,999\,892\,657\,0\frac{70}{68}\ E-1]$
3	$[1.793\,178\,750\,459\ \frac{800}{700}\ E-5]$	$[9.999\,999\,998\,392\,25\frac{7}{4}\ E-1]$
4	$[9.999\,999\,993\,415\,9\frac{21}{18}\ E-1]$	$[3.628\,796\,176\,862\,38\frac{3}{1}\ E-5]$

If we know that (1.8) holds, then the generalized singular value $(c_5, s_5) = (1, 0)$ can be deduced directly from the dimensions of A, B and from the enclosures $[s_i]$, $i = 1, \ldots, 4$, as can be seen in the following way: We apply Theorem 1.3 and Corollary 1.4 to A and $\hat{B} := \begin{pmatrix} B \\ 0 \end{pmatrix} \in \mathbb{R}^{5\times 5}$, where we have to enlarge B to $\hat{B}$ since we assumed $q \geq n$ in Theorem 1.3. The generalized singular values remain the same as for (A, B), therefore

$$4 \geq \operatorname{rank}(B) = \operatorname{rank}\left(\hat{B}\right) = \operatorname{rank}(S) \geq 4 \tag{5.1}$$

which means $\operatorname{rank}(B) = 4$. For the last inequality in (5.1) we used the fact that the enclosures for s_i, $i = 1, \ldots, 4$ are pairwise disjoint and do not contain zero. Thus, $\operatorname{rank}(S) = 4$, hence s_5 must be zero and Theorem 1.3 implies $c_5 = 1$.

Unfortunately, we do not yet know whether (1.8) is true. We are now going to prove this using the iteration method of this paper: The approximation $(\tilde{c}_5, \tilde{s}_5)$ indicates $(c_5, s_5) = (1, 0)$. Therefore, we start our program with $\tilde{c}_5 = 1$, $\tilde{s}_5 = 0$, $\tilde{v}^5 = 0$ and

$$\tilde{u}^5 = \begin{pmatrix} 6.373\,109\,613\,136\ E-1 \\ 5.561\,243\,807\,402\ E-1 \\ -1.129\,007\,623\,404\ E-1 \\ 2.801\,674\,929\,222\ E-1 \\ 4.393\,577\,391\,708\ E-1 \end{pmatrix},\ \tilde{x}^5 = \frac{1}{10\,000} \begin{pmatrix} 9.619\,099\,638\,573\ E-1 \\ 3.564\,487\,242\,938\ E-1 \\ -4.914\,070\,317\,944\ E-1 \\ 4.307\,023\,711\,361\ E-1 \\ -8.850\,092\,357\,967\ E-1 \end{pmatrix}.$$

The choice $\tilde{v}^5 = 0$ can be motivated as follows: The enclosures for s_i, $i = 1, \ldots, 4$, are pairwise disjoint and do not contain zero. In addition, $0 \notin [c_i]$ whence $s_i c_i \neq 0$ for $i = 1, \ldots, 4$. Therefore, the corresponding vectors $v^1, \ldots, v^4$ are pairwise orthonormal according to the arguments preceding (4.17) for which no assumptions (4.23) are needed. Hence (4.9) implies $\operatorname{rank}(B^T) = \operatorname{rank}(B) = 4$ since $\operatorname{rank}(B)$ is the dimension of the range of the linear mapping $x \mapsto Bx$. From (4.10) and from the assumption $(c_5, s_5) = (1, 0)$ we get $B^T v^5 = 0$ which implies $v^5 = 0$ by the rank of B. This is the reason, why we put $\tilde{v}^5 = 0$.

For the quantities in Theorem 3.1 and 4.1 we obtained

$$\begin{aligned}
\alpha &\in [3.755\,282\,052\,424\,44^{2}_{0}\ E-11]\,,\\
\beta &\in [1.975\,959\,3^{9}_{7}\ E-8]\,,\\
\gamma &\in [3.519\,783\,275\,163\,5^{80}_{78}\ E+5]\,,\\
\Delta &\in [9.999\,470\,893\,649\,6^{65}_{57}\ E-1]\,,\\
r^- &\leq 3.755\,331\,764\,470\,081\ E-11\,,\\
r^+ &\geq 2.841\,046\,405\,733\,589\ E-6\,.
\end{aligned}$$

We used $r := 3.755\,331\,764\,470\,081\ E-11 \in [r^-, r^+]$. After two iterations we got the following enclosures for c_5, s_5, u^5 and x^5:

$$\begin{aligned}
[c_5] &= [9.999\,999\,999\,999\,998\ E-1\,,\ 1.000\,000\,000\,000\,001\ E+0]\,,\\
[s_5] &= [-3.2\ E-30\,,\ 3.2\ E-30]\,,
\end{aligned}$$

$$[u^5] = \begin{pmatrix}
[\ \ 1.687\,550\,110\,072\,0^{80}_{79}\ E-2\ ]\\
[\ \ 6.373\,109\,613\,222\,01^{8}_{6}\ E-1\ ]\\
[\ \ 5.561\,243\,807\,173\,0^{40}_{38}\ E-1\ ]\\
[-1.129\,007\,623\,779\,52^{8}_{9}\ E-1\ ]\\
[\ \ 2.801\,674\,929\,369\,78^{2}_{1}\ E-1\ ]\\
[\ \ 4.393\,577\,391\,703\,15^{8}_{6}\ E-1\ ]
\end{pmatrix},$$

$$[x^5] = \begin{pmatrix}
[\ \ 9.619\,099\,638\,742\,1^{12}_{09}\ E-5\ ]\\
[\ \ 3.564\,487\,242\,945\,36^{8}_{6}\ E-5\ ]\\
[-4.914\,070\,317\,892\,76^{3}_{5}\ E-5\ ]\\
[\ \ 4.307\,023\,710\,673\,22^{4}_{3}\ E-5\ ]\\
[-8.850\,092\,357\,418\,17^{1}_{3}\ E-5\ ]
\end{pmatrix}.$$

The enclosures $[c_i]$, $i = 1, \ldots 5$, are pairwise disjoint and do not contain zero. Therefore, the corresponding vectors $u^1, \ldots, u^5$ are pairwise orthonormal, where we again use the arguments preceding (4.17). Hence (4.8) implies $\operatorname{rank}(A) = 5$ which guarantees (1.8). From this, $(c_5, s_5) = (1, 0)$ follows, as we already saw above. Note that in [13], p. 16, it was stated that the rank of both matrices A and B is three.

Acknowledgements

The authors are indebted to an anonymous referee whose remarks and comments helped to improve the paper.

References

[1] ALEFELD, G.: Berechenbare Fehlerschranken für ein Eigenpaar unter Einschluß von Rundungsfehlern bei Verwendung des genauen Skalarprodukts, Z. angew. Math. Mech. **67** (1987), 145–152

[2] ALEFELD, G.: Rigorous Error Bounds for Singular Values of a Matrix Using the Precise Scalar Product. In: KAUCHER, E., KULISCH, U., and ULLRICH, CH. (eds.): Computerarithmetic, Teubner, Stuttgart, 1987, 9–30

[3] ALEFELD, G.: Errorbounds for Quadratic Systems of Nonlinear Equations Using the Precise Scalar Product. In: KULISCH, U., and STETTER, H. J. (eds.): Scientific Computation with Automatic Result Verification, Computing, Suppl. **6** (1988), 59–68

[4] ALEFELD, G.: Berechenbare Fehlerschranken für ein Eigenpaar beim verallgemeinerten Eigenwertproblem, Z. angew. Math. Mech. **68** (1988), 181–184

[5] ALEFELD, G., and HERZBERGER, J.: Introduction to Interval Computations, Academic Press, New York, 1983

[6] ALEFELD, G., and SPREUER, H.: Iterative Improvement of Componentwise Errorbounds for Invariant Subspaces Belonging to a Double or Nearly Double Eigenvalue, Computing **36** (1986), 321–334

[7] ANDERSON, E., BAI, Z., BISCHOF, J., DEMMEL, J., DONGARRA, J., DU CROZ, J., GREENBAUM, A., HAMMARLING, S., MCKENNEY, A., OSTROUCHOV, S., and SORENSEN, D.: LAPACK, User's Guide, 2nd ed., SIAM, Philadelphia, 1995

[8] BAI, Z., and DEMMEL, J. W.: Computing the Generalized Singular Value Decomposition, SIAM J. Sci. Stat. Comput. **14** (1993), 1464–1486

[9] BAI, Z., and ZHA, H.: A New Preprocessing Algorithm for the Computation of the Generalized Singular Value Decomposition, SIAM J. Sci. Stat. Comput. **14** (1993), 1007–1012

[10] GOLUB, G. H., and VAN LOAN, C. F.: Matrix Computations, The Johns Hopkins University Press, Baltimore, Maryland, 1983

[11] HOFFMANN, R.: Konstruktion von Fehlerschranken bei der verallgemeinerten Singulärwertzerlegung und ihre iterative Verbesserung, Thesis, Universität Karlsruhe, 1993

[12] KLATTE, R., KULISCH, U., NEAGA, M., RATZ, D., and ULLRICH, CH.: PASCAL–XSC. Language Reference with Examples, Springer–Verlag, Berlin, 1992

[13] LAWSON, C. L., and HANSON, R. J.: Solving Least Squares Problems, Classics in Applied Mathematics **15**, SIAM, Philadelphia, 1995

[14] MAYER, G.: Result Verification for Eigenvectors and Eigenvalues. In: HERZBERGER, J. (ed.): Topics in Validated Computations, Studies in Computational Mathematics **5**, Elsevier (North–Holland), Amsterdam, 1994, 209–276

[15] MAYER, G.: On a Unified Representation of Some Interval Analytic Algorithms, Rostock. Math. Kolloq. **49** (1995), 75–88

[16] NEUMAIER, A.: Interval Methods for Systems of Equations, Cambridge University Press, Cambridge, 1990

[17] PAIGE, C. C.: Computing the Generalized Singular Value Decomposition, SIAM J. Sci. Stat. Comput. **7** (1986), 1126–1146

[18] PAIGE, C. C., and SAUNDERS, M. A.: Towards a Generalized Singular Value Decomposition, SIAM J. Numer. Anal. **18** (1981), 398–405

[19] PLATZÖDER, L.: Einige Beiträge über die Existenz von Lösungen nichtlinearer Gleichungssysteme und Verfahren zu ihrer Berechnung, Thesis, Berlin, 1981

[20] VAN LOAN, C. F.: Generalizing the Singular Value Decomposition, SIAM J. Numer. Anal. **13** (1976), 76–83

[21] VAN LOAN, C. F.: Computing the CS and the Generalized Singular Value Decompositions, Numer. Math. **46** (1985), 479–491

Institut für Angewandte Mathematik
Universität Karlsruhe
D–76128 Karlsruhe
Germany
e–mail:
goetz.alefeld@mathematik.uni–karlsruhe.de

Neutharder Strasse 33
D–76689 Karlsdorf–Neuthard
Germany

Fachbereich Mathematik
Universität Rostock
D–18051 Rostock
Germany
e–mail:
guenter.mayer@mathematik.uni–rostock.de