2013

On condition numbers for the canonical generalized polar decompostion of real matrices

Ze-Jia Xie
xiezejia2012@gmail.com

Wen Li

Xiao-Qing Jin

ELA

ON CONDITION NUMBERS FOR THE CANONICAL GENERALIZED POLAR DECOMPOSITION OF REAL MATRICES[*]

ZE-JIA XIE[†], WEN LI[‡], AND XIAO-QING JIN[§]

Abstract. Three different kinds of condition numbers: normwise, mixed and componentwise, are discussed for the canonical generalized polar decomposition (CGPD) of real matrices. The technique used herein is different from the ones in previous literatures of the polar decomposition. With some modifications of the definition of the componentwise condition number, its application scope is extended. Explicit expressions and computable upper bounds of these three condition numbers for the CGPD are presented. Besides, some first order normwise and componentwise perturbation bounds for the CGPD are also obtained. At last, some numerical examples are given to demonstrate the theoretical results.

Key words. Condition number, Canonical generalized polar decomposition, Perturbation analysis, Sensitivity.

AMS subject classifications. 15A12, 65F35.

1. Introduction. In this paper, we are interested in the discussion of perturbation bounds and condition numbers for two factors of the canonical generalized polar decomposition (CGPD) of real matrices. Higham et al. [16] proposed the CGPD which was a generalization of the (generalized) polar decomposition, weighted generalized polar decomposition [31], and H-polar decomposition [4, 5]. Therefore, the CGPD has the same applications as those of these polar decompositions (see [6, 14, 19]). Before we introduce the definition of the CGPD, we will introduce several concepts, which can be found in [16].

The symbol I_n stands for the identity matrix of order n. A^T and A^* denote the transpose and the conjugate transpose of matrix A, respectively. A scalar product between two vectors $\mathbf{x} \in \mathbb{K}^m$ and $\mathbf{y} \in \mathbb{K}^m$ ($\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$) in terms of a nonsingular

[†]School of Mathematical Sciences, South China Normal University, Guangzhou 510631, China (xiezejia2012@gmail.com).

[‡]School of Mathematical Sciences, South China Normal University, Guangzhou 510631, China (liwen@scnu.edu.cn). This author was supported in part by National Natural Science Foundations of China (no. 1097107511271144), Guangdong Provincial Natural Science Foundations (no. 09150631000021, s2012010009985) and Research Fund for the Doctoral Program of Higher Education of China (no. 20104407110001).

[§]Department of Mathematics, University of Macau, Macau, China (xqjin@umac.mo). This author was supported by the grant UL020/08-Y5/MAT/JXQ01/FST from University of Macau.

ELA

http://math.technion.ac.il/iic/ela

matrix $M \in \mathbb{K}^{m \times m}$ is defined by

$$\langle \mathbf{x}, \mathbf{y} \rangle_M = \begin{cases} \mathbf{x}^T M \mathbf{y}, & \text{for bilinear forms,} \\ \mathbf{x}^* M \mathbf{y}, & \text{for sesquilinear forms.} \end{cases}$$

Assume that $\mathbb{K}^m$ and $\mathbb{K}^n$ are equipped with scalar products $\langle \cdot, \cdot \rangle_M$ and $\langle \cdot, \cdot \rangle_N$ induced by the nonsingular matrices $M \in \mathbb{K}^{m \times m}$ and $N \in \mathbb{K}^{n \times n}$, respectively. The (M, N)-adjoint of a matrix $A \in \mathbb{K}^{m \times n}$ is defined to be the unique matrix $A^{\star M,N} \in \mathbb{K}^{n \times m}$ satisfying the identity

$$\langle A\mathbf{x}, \mathbf{y} \rangle_M = \langle \mathbf{x}, A^{\star M,N} \mathbf{y} \rangle_N$$

for all $\mathbf{x} \in \mathbb{K}^n$ and all $\mathbf{y} \in \mathbb{K}^m$, and $A^{\star M,N}$ is defined by

$$A^{\star M,N} \equiv \begin{cases} N^{-1} A^T M, & \text{for bilinear forms,} \\ N^{-1} A^* M, & \text{for sesquilinear forms.} \end{cases}$$

The nonsingular matrices $M \in \mathbb{K}^{m \times m}$ and $N \in \mathbb{K}^{n \times n}$ form an orthosymmetric pair if (i) for bilinear forms,

$$M^T = \beta M, \quad N^T = \beta N, \quad \beta = \pm 1,$$

or (ii) for sesquilinear forms,

$$M^* = \alpha M, \quad N^* = \alpha N, \quad \alpha \in \mathbb{C}, \quad |\alpha| = 1.$$

A matrix $W \in \mathbb{K}^{m \times n}$ is a partial (M, N)-isometry if

$$(1.1) \qquad\qquad W W^{\star M,N} W = W.$$

A matrix $S \in \mathbb{K}^{n \times n}$ is said to be N-selfadjoint if

$$(1.2) \qquad\qquad S^{\star N} \equiv N^{-1} S^T N = S.$$

DEFINITION 1.1 ([16]). Let the nonsingular matrices $M \in \mathbb{K}^{m \times m}$ and $N \in \mathbb{K}^{n \times n}$ form an orthosymmetric pair. A canonical generalized polar decomposition of $A \in \mathbb{K}^{m \times n}$ is a decomposition $A = WS$, where $W \in \mathbb{K}^{m \times n}$ is a partial (M, N)-isometry, $S \in \mathbb{K}^{n \times n}$ is an N-selfadjoint matrix whose nonzero eigenvalues are contained in the open right half-plane, and $\text{range}(W^{\star M,N}) = \text{range}(S)$.

Higham et al. [16] also provided the necessary and sufficient condition for the existence and uniqueness of the CGPD, which is given below.

$$\boxed{\textbf{ELA}}$$

THEOREM 1.2. *Let the nonsingular matrices $M \in \mathbb{K}^{m \times m}$ and $N \in \mathbb{K}^{n \times n}$ form an orthosymmetric pair. Then $A \in \mathbb{K}^{m \times n}$ has a unique canonical generalized polar decomposition if and only if*

(i) *$A^{\star M,N} A$ has no negative real eigenvalues;*

(ii) *if zero is an eigenvalue of $A^{\star M,N} A$, then it is semisimple; and*

(iii) *$\ker(A^{\star M,N} A) = \ker(A)$.*

If $\mathbb{K} = \mathbb{R}$, then we have

$$(1.3) \qquad\qquad A^{\star M,N} = N^{-1} A^T M.$$

From Theorem 1.2, it is easy to obtain the following theorem.

THEOREM 1.3. *Let the nonsingular matrices $M \in \mathbb{R}^{m \times m}$ and $N \in \mathbb{R}^{n \times n}$ form an orthosymmetric pair. Then $A \in \mathbb{R}^{m \times n}$ has a unique canonical generalized polar decomposition $A = WS$ with S being nonsingular if and only if $A^{\star M,N} A$ has no nonpositive eigenvalues.*

In the rest of this paper, we always assume that $\mathbb{K} = \mathbb{R}$ and $A^{\star M,N} A$ has no nonpositive eigenvalues when we refer to the CGPD.

We study three different kinds of condition numbers: normwise, mixed and componentwise, for the CGPD. The classical condition number is the normwise one, which has a drawback that it ignores the structure of both input and output data. To be more accurate, Gohberg and Koltracht [12] proposed another two different kinds of condition numbers: mixed and componentwise. More about these two condition numbers can be found in [10, 11, 12, 29]. In this paper, we also modify the definition of the componentwise condition number to extend its scope of application.

The perturbation analysis for the (generalized) polar decomposition has been studied by many authors [1, 2, 3, 8, 9, 20, 21, 22, 23, 24, 25]. They mainly provided the normwise perturbation bounds for the (subunitary) unitary and Hermite positive semidefinite polar factors. Yang and Li [31] explored a weighted generalized polar decomposition, but it is defined only with respect to positive definite scalar products, which is a special case of the CGPD. Yang and Li [26, 30] also gave the normwise perturbation bounds for the weighted polar decomposition. Some authors [7, 15, 18, 27] gave the normwise condition number for the (generalized) polar decomposition. However, so far, no one has presented the perturbation bounds and condition numbers for the CGPD. By this motivation, we discuss three different kinds of condition numbers: normwise, mixed and componentwise, for the CGPD, and present their first order normwise and componentwise perturbation bounds.

The rest of this paper is organized as follows. In Section 2, we introduce some preliminaries. In Section 3, we discuss some first order normwise and componentwise

ELA

http://math.technion.ac.il/iic/ela

perturbation bounds and condition numbers for two factors of the CGPD. Finally, we give some numerical examples in Section 4 to demonstrate our theoretical results.

2. Preliminaries.

2.1. Vector and matrix norms. For a matrix $A \in \mathbb{R}^{m \times n}$, the Frobenius norm and the spectral norm of A are denoted by $\|A\|_F$ and $\|A\|_2$, respectively. $\|\mathbf{x}\|_2$ denotes the Euclidean norm of a vector $\mathbf{x} \in \mathbb{R}^n$. The ∞-norm of a vector $\mathbf{x} \in \mathbb{R}^n$ and a matrix $A \in \mathbb{R}^{m \times n}$ are defined respectively by

$$\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i| \quad \text{and} \quad \|A\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|.$$

The maximum norm of a matrix $A \in \mathbb{R}^{m \times n}$ is defined by

$$\|A\|_{\max} = \max_{i,j} |a_{ij}|.$$

One should note that the inequality

$$\|AB\|_{\max} \leq \|A\|_{\max} \|B\|_{\max}$$

does not necessarily hold for all matrices A, B with appropriate sizes. This is different from the spectral norm, the Frobenius norm, and the ∞-norm.

For the Kronecker product of $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times q}$, we have [17],

(2.1) $$\|A \otimes B\|_\infty = \|A\|_\infty \|B\|_\infty \quad \text{and} \quad \|A \otimes B\|_2 = \|A\|_2 \|B\|_2.$$

2.2. Three different kinds of condition numbers. The normwise condition number measures the size of both input perturbations and output errors by using some norms. However, it ignores the structure of both input and output data with respect to scaling or sparsity. To be more accurate, two different kinds of condition numbers: mixed and componentwise, are proposed. The mixed condition number measures the errors in the output normwise and the input perturbations componentwise, and the componentwise one measures both the errors in the output and the perturbation in the input componentwise.

To define these three different kinds of condition numbers, we first introduce some preliminaries. For a matrix $X \in \mathbb{R}^{m \times n}$, $|X|$ denotes the m-by-n matrix whose (i,j)-entry is just the absolute value of the (i,j)-entry of X, and $\text{vec}(X)$ denotes the vector as follows:

$$\text{vec}(X) = (x_{1,1}, \ldots, x_{m,1}, x_{1,2}, \ldots, x_{m,2}, \ldots, x_{1,n}, \ldots, x_{m,n})^T.$$

ELA

http://math.technion.ac.il/iic/ela

For a number $c \in \mathbb{R}$, we define

$$c^{\ddagger} = \begin{cases} c^{-1}, & \text{if } c \neq 0, \\ 1, & \text{if } c = 0. \end{cases}$$

We use the mark "$\ddagger$" here just for distinguishing from "$\dagger$", whose value at zero is zero. For any two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, we define the vector $\mathbf{a}/\mathbf{b}$ as

$$(2.2) \qquad \frac{\mathbf{a}}{\mathbf{b}} = \text{diag}^{\ddagger}(\mathbf{b})\mathbf{a},$$

where $\text{diag}(\mathbf{b})$ and $\text{diag}^{\ddagger}(\mathbf{b})$ are n-by-n diagonal matrices with diagonal entries being equal to $b_1, b_2, \ldots, b_n$ and $b_1^{\ddagger}, b_2^{\ddagger}, \ldots, b_n^{\ddagger}$, respectively. It is obvious that $\mathbf{a}/\mathbf{b}$ has components

$$\left(\frac{\mathbf{a}}{\mathbf{b}}\right)_i = b_i^{\ddagger} a_i.$$

Moreover, we can define the componentwise distance between $\mathbf{a}$ and $\mathbf{b}$ by

$$(2.3) \qquad d(\mathbf{a}, \mathbf{b}) = \left\| \frac{\mathbf{a} - \mathbf{b}}{\mathbf{b}} \right\|_{\infty} = \max_{1 \leq i \leq n} \left\{ |b_i^{\ddagger}||a_i - b_i| \right\}.$$

That is, we consider the relative distance at nonzero components, while the absolute distance at zero components. It is obvious that $d(\mathbf{a}, \mathbf{b}) = 0$ if and only if $\mathbf{a} = \mathbf{b}$.

Let $A, B \in \mathbb{R}^{m \times n}$. We define A/B as an entrywise division with entries

$$\left(\frac{A}{B}\right)_{ij} = b_{ij}^{\ddagger} a_{ij},$$

and the componentwise distance of A and B as

$$d(A, B) = \left\| \frac{A - B}{B} \right\|_{\max} = \max_{i,j} \left\{ |b_{ij}^{\ddagger}||a_{ij} - b_{ij}| \right\}.$$

For a vector $\mathbf{a} = (a_1, a_2, \ldots, a_p)^T \in \mathbb{R}^p$, we define

$$\Omega(\mathbf{a}) = \{k \mid a_k = 0, \ 1 \leq k \leq p\} \text{ and } |\mathbf{a}| = (|a_1|, |a_2|, \ldots, |a_p|)^T.$$

Given $\varepsilon > 0$, we denote

$$B^o(\mathbf{a}, \varepsilon) = \{\mathbf{x} \in \mathbb{R}^p \mid |x_i - a_i| \leq \varepsilon|a_i|, \ i = 1 : p\}.$$

It is obvious that if $x \in B^o(\mathbf{a}, \varepsilon)$, then $\Omega(\mathbf{a}) \subseteq \Omega(\mathbf{x})$ and $\mathbf{x} = \text{diag}(\mathbf{a})\text{diag}^{\ddagger}(\mathbf{a})\mathbf{x}$. We also denote

$$B(\mathbf{a}, \varepsilon) = \{\mathbf{x} \in \mathbb{R}^p \mid \|\mathbf{x} - \mathbf{a}\|_2 \leq \varepsilon\|\mathbf{a}\|_2\}.$$

ELA

http://math.technion.ac.il/iic/ela

Now we introduce the definitions of three different kinds of condition numbers. The first one is the usual condition number given by Rice [28]. The last two definitions are given by Gohberg and Koltracht [12].

DEFINITION 2.1. Let $F : \mathbb{R}^p \to \mathbb{R}^q$ be a continuous mapping defined on an open set $D_F \subset \mathbb{R}^p$, and $\mathbf{a} \in D_F$, $\mathbf{a} \neq \mathbf{0}$, such that $F(\mathbf{a}) \neq \mathbf{0}$.

(i) The normwise condition number of F at $\mathbf{a}$ is defined by

$$\kappa(F, \mathbf{a}) = \lim_{\varepsilon \to 0} \sup_{\substack{\mathbf{x} \neq \mathbf{a} \\ \mathbf{x} \in B(\mathbf{a}, \varepsilon)}} \frac{\|F(\mathbf{x}) - F(\mathbf{a})\|_2 / \|F(\mathbf{a})\|_2}{\|\mathbf{x} - \mathbf{a}\|_2 / \|\mathbf{a}\|_2}.$$

(ii) The mixed condition number of F at $\mathbf{a}$ is defined by

$$m(F, \mathbf{a}) = \lim_{\varepsilon \to 0} \sup_{\substack{\mathbf{x} \neq \mathbf{a} \\ \mathbf{x} \in B^o(\mathbf{a}, \varepsilon)}} \frac{\|F(\mathbf{x}) - F(\mathbf{a})\|_\infty}{\|F(\mathbf{a})\|_\infty} \frac{1}{d(\mathbf{x}, \mathbf{a})}.$$

(iii) The componentwise condition number of F at $\mathbf{a}$ is defined by

$$c(F, \mathbf{a}) = \lim_{\varepsilon \to 0} \sup_{\substack{\mathbf{x} \neq \mathbf{a} \\ \mathbf{x} \in B^o(\mathbf{a}, \varepsilon)}} \frac{d(F(\mathbf{x}), F(\mathbf{a}))}{d(\mathbf{x}, \mathbf{a})}.$$

REMARK 2.2. (i) It is noted that Definition 2.1 (iii) is the same as the one given in [12] when $F(\mathbf{a}) = (f_1(\mathbf{a}), \ldots, f_q(\mathbf{a}))$ has no zero components. Since the distance d we defined is always finite, which is different from δ defined in [12], the hypothesis that $F(\mathbf{a})$ has no zero components in [12] can be removed.

(ii) From Definition 2.1, we can see that the mixed and componentwise condition numbers demand that the zero components of $\mathbf{a}$ are not perturbed, while the normwise one does not have this requirement. Actually, the demand that zero components should not be perturbed is the case when $x_i = fl(a_i)$ is the representation of a_i in a computer arithmetic, which has the property $|x_i - a_i| \leq u|a_i|$, where u is the unit roundoff (see [13, Theorem 2.2]). Note that if $\mathbf{x} \in B^o(\mathbf{a}, \varepsilon)$, then $\mathbf{x} \in B(\mathbf{a}, \varepsilon)$. Obviously, this is not true in the opposite direction. Therefore, the problem of computing F at $\mathbf{a}$ could be ill conditioned with respect to perturbations in point $\mathbf{a}$ satisfying $\mathbf{x} \in B(\mathbf{a}, \varepsilon)$ while being well conditioned with respect to perturbations satisfying $\mathbf{x} \in B^o(\mathbf{a}, \varepsilon)$.

The following lemma, which gives the explicit expressions for these three condition numbers, can be found in [10, Lemma 2]. Since the definition of the componentwise condition number is modified correspondingly, we give a new proof for it.

ELA

http://math.technion.ac.il/iic/ela

LEMMA 2.3. *With the same assumptions as in Definition* 2.1, *and supposing that* F *is Fréchet differentiable at* $\mathbf{a}$, *we have*

$$(2.4) \qquad \kappa(F, \mathbf{a}) = \frac{\|F'(\mathbf{a})\|_2 \|\mathbf{a}\|_2}{\|F(\mathbf{a})\|_2},$$

$$(2.5) \qquad m(F, \mathbf{a}) = \frac{\|F'(\mathbf{a})\mathrm{diag}(\mathbf{a})\|_\infty}{\|F(\mathbf{a})\|_\infty} = \frac{\||F'(\mathbf{a})||\mathbf{a}|\|_\infty}{\|F(\mathbf{a})\|_\infty},$$

$$(2.6) \qquad c(F, \mathbf{a}) = \|\mathrm{diag}^\ddagger(F(\mathbf{a}))F'(\mathbf{a})\mathrm{diag}(\mathbf{a})\|_\infty = \left\|\frac{|F'(\mathbf{a})||\mathbf{a}|}{|F(\mathbf{a})|}\right\|_\infty,$$

where $F'(\mathbf{a})$ *is the Fréchet derivative of* F *at* $\mathbf{a}$.

Proof. Here we only prove (2.6), (2.4) and (2.5) can be proved similarly. By Definition 2.1, (2.2) and (2.3), we have

$$c(F, \mathbf{a}) = \lim_{\varepsilon \to 0} \sup_{\substack{\mathbf{x} \neq \mathbf{a} \\ \mathbf{x} \in B^o(\mathbf{a}, \varepsilon)}} \frac{d(F(\mathbf{x}), F(\mathbf{a}))}{d(\mathbf{x}, \mathbf{a})}$$

$$= \lim_{\varepsilon \to 0} \sup_{\substack{\mathbf{x} \neq \mathbf{a} \\ \mathbf{x} \in B^o(\mathbf{a}, \varepsilon)}} \frac{\|\mathrm{diag}^\ddagger(F(\mathbf{a}))[F(\mathbf{x}) - F(\mathbf{a})]\|_\infty}{\|\mathrm{diag}^\ddagger(\mathbf{a})(\mathbf{x} - \mathbf{a})\|_\infty}.$$

Denote $D_\mathbf{a} = \mathrm{diag}(\mathbf{a})$, $D_\mathbf{a}^\ddagger = \mathrm{diag}^\ddagger(\mathbf{a})$ and $D_{F\mathbf{a}}^\ddagger = \mathrm{diag}^\ddagger(F(\mathbf{a}))$. Since $\mathbf{x} \in B^o(\mathbf{a}, \varepsilon)$, we know that $\Omega(\mathbf{a}) \subseteq \Omega(\mathbf{x})$ and $\mathbf{x} = D_\mathbf{a} D_\mathbf{a}^\ddagger \mathbf{x}$. Let $\mathbf{y} = D_\mathbf{a}^\ddagger \mathbf{x}$ and $\mathbf{b} = D_\mathbf{a}^\ddagger \mathbf{a}$. Then $\mathbf{x} = D_\mathbf{a}\mathbf{y}$, $\mathbf{a} = D_\mathbf{a}\mathbf{b}$, and $\mathbf{x} \neq \mathbf{a}$ if and only if $\mathbf{y} \neq \mathbf{b}$. By the Chain Rule of the Fréchet derivative, we have

$$c(F, \mathbf{a}) = \lim_{\varepsilon \to 0} \sup_{\substack{\mathbf{x} \neq \mathbf{a} \\ \mathbf{x} \in B^o(\mathbf{a}, \varepsilon)}} \frac{\|D_{F\mathbf{a}}^\ddagger F(\mathbf{x}) - D_{F\mathbf{a}}^\ddagger F(\mathbf{a})\|_\infty}{\|D_\mathbf{a}^\ddagger \mathbf{x} - D_\mathbf{a}^\ddagger \mathbf{a}\|_\infty}$$

$$= \lim_{\varepsilon \to 0} \sup_{\substack{\mathbf{y} \neq \mathbf{b} \\ \mathbf{y} \in B^o(\mathbf{b}, \varepsilon)}} \frac{\|D_{F\mathbf{a}}^\ddagger F(D_\mathbf{a}\mathbf{y}) - D_{F\mathbf{a}}^\ddagger F(D_\mathbf{a}\mathbf{b})\|_\infty}{\|\mathbf{y} - \mathbf{b}\|_\infty}$$

$$= \|D_{F\mathbf{a}}^\ddagger F'(D_\mathbf{a}\mathbf{b})D_\mathbf{a}\|_\infty = \|D_{F\mathbf{a}}^\ddagger F'(\mathbf{a})D_\mathbf{a}\|_\infty.$$

Hence, by considering the proof of Lemma 2 in [10] and (2.2), we obtain

$$\|D_{F\mathbf{a}}^\ddagger F'(\mathbf{a})D_\mathbf{a}\|_\infty = \left\|\frac{|F'(\mathbf{a})||\mathbf{a}|}{|F(\mathbf{a})|}\right\|_\infty,$$

which proves (2.6). □

REMARK 2.4. It is obvious that (2.6) is a generalization of

$$c(x) = \frac{|f'(x)||x|}{|f(x)|},$$

which is the relative condition number for computing the scalar real function $y = f(x) \in \mathbb{R}$ (see [13]). And if $f(x) = 0$, we then usually consider the absolute condition number $|f'(x)|$ as its sensitivity for computing $f(x)$. From (2.6), we see that if $F(\mathbf{a}) = \mathbf{0}$, then $c(F, \mathbf{a}) = \||F'(\mathbf{a})||\mathbf{a}|\|_\infty$, which measures the absolute error in the output for a given relative perturbation in the input. But if we define $c(F, \mathbf{a}) = \|\mathrm{diag}^\dagger(F(\mathbf{a}))F'(\mathbf{a})\mathrm{diag}(\mathbf{a})\|_\infty$, where $\mathrm{diag}^\dagger(F(\mathbf{a}))$ denotes the Moore-Penrose inverse of $\mathrm{diag}(F(\mathbf{a}))$, then it will be less informative when $F(\mathbf{a})$ has zero components.

3. Perturbation analysis for the CGPD. For the CGPD of $A = WS$, we define two mappings as follows:

$$\varphi_S : \mathrm{vec}(A) \mapsto \mathrm{vec}(S),$$
$$\varphi_W : \mathrm{vec}(A) \mapsto \mathrm{vec}(W).$$

It is noted that S and W are unique, hence φ_S and φ_W are well-defined.

3.1. Perturbation bounds for the factor S and W. Let $\widetilde{A} = A + \Delta A \in \mathbb{R}^{m \times n}$ be a perturbed matrix of A. Suppose $A = WS$ and

$$(3.1) \qquad A + \Delta A = (W + \Delta W)(S + \Delta S)$$

are the CGPDs of A and $\widetilde{A}$, respectively. Since $A^{\star M,N}A = S^2$ (see [16, Lemma 3.7]), we have

$$(3.2) \qquad N^{-1}(A + \Delta A)^T M(A + \Delta A) = (S + \Delta S)(S + \Delta S).$$

Omitting the second-order terms, (3.2) can be turned into

$$(3.3) \qquad N^{-1}A^T M \Delta A + N^{-1}\Delta A^T MA \approx S\Delta S + \Delta SS.$$

Using the vec operation, we have

$$(3.4) \qquad \begin{aligned} &[(I_n \otimes N^{-1}A^T M) + (A^T M^T \otimes N^{-1})\Pi]\mathrm{vec}(\Delta A) \\ \approx &(I_n \otimes S + S^T \otimes I_n)\mathrm{vec}(\Delta S), \end{aligned}$$

where

$$(3.5) \qquad \Pi = \sum_{i=1}^{m}\sum_{j=1}^{n} E_{ij} \otimes E_{ij}^T \in \mathbb{R}^{mn \times mn}$$

is a permutation matrix and each $E_{ij} \in \mathbb{R}^{m \times n}$ has "1" in the (i,j)th entry and all other entries are zero.

ELA

http://math.technion.ac.il/iic/ela

Since all the eigenvalues of S lie in the open right half-plane, we know that $0 \notin \lambda(I_n \otimes S + S^T \otimes I_n)$. Then $I_n \otimes S + S^T \otimes I_n$ is nonsingular. Hence, it follows from (3.4) that

$$(3.6) \qquad \mathrm{vec}(\Delta S) \approx K_S^{-1} K_{AMN} \mathrm{vec}(\Delta A),$$

where

$$(3.7) \qquad \begin{aligned} K_S &= I_n \otimes S + S^T \otimes I_n, \\ K_{AMN} &= (I_n \otimes N^{-1} A^T M) + (A^T M^T \otimes N^{-1})\Pi. \end{aligned}$$

Omitting the second-order term of (3.1) gives

$$(3.8) \qquad \Delta A \approx W \Delta S + \Delta W S.$$

Since S is nonsingular, (3.8) can be changed into

$$(3.9) \qquad \Delta W \approx \Delta A S^{-1} - W \Delta S S^{-1}.$$

Taking the vec operation in both sides of (3.9) yields

$$(3.10) \qquad \mathrm{vec}(\Delta W) \approx (S^{-T} \otimes I_m)\mathrm{vec}(\Delta A) - (S^{-T} \otimes W)\mathrm{vec}(\Delta S).$$

Substituting (3.6) into (3.10), we obtain

$$(3.11) \qquad \mathrm{vec}(\Delta W) \approx [(S^{-T} \otimes I_m) - (S^{-T} \otimes W)K_S^{-1} K_{AMN}]\mathrm{vec}(\Delta A).$$

From the definitions of φ_S, φ_W and the Fréchet derivative, and combining (3.6) and (3.11), we can easily obtain the following theorem.

THEOREM 3.1. *Let the nonsingular matrices $M \in \mathbb{R}^{m \times m}$ and $N \in \mathbb{R}^{n \times n}$ form an orthosymmetric pair. Suppose that $A^{\star M,N} A$ has no nonpositive real eigenvalues, then $A = WS$ is a unique CGPD of $A \in \mathbb{R}^{m \times n}$ with S being nonsingular. Furthermore, the Fréchet derivatives of φ_S and φ_W at $\mathbf{a} = \mathrm{vec}(A)$ are given respectively by*

$$(3.12) \qquad \varphi_S'(\mathbf{a}) = K_S^{-1} K_{AMN},$$

and

$$(3.13) \qquad \varphi_W'(\mathbf{a}) = (S^{-T} \otimes I_m) - (S^{-T} \otimes W)K_S^{-1} K_{AMN},$$

where K_S and K_{AMN} are defined by (3.7).

Taking the Euclidean norm and the absolute value on (3.6) and (3.11) respectively, we can easily get the normwise and componentwise perturbation bounds for the factors S and W.

ELA

http://math.technion.ac.il/iic/ela

THEOREM 3.2. *With the same assumptions as in Theorem* 3.1, *suppose that* $(A + \Delta A)^{\star M,N}(A + \Delta A)$ *has no nonpositive real eigenvalues. We have*

$$(3.14) \qquad \|\Delta S\|_F \lesssim \|K_S^{-1} K_{AMN}\|_2 \|\Delta A\|_F,$$

$$(3.15) \qquad \|\Delta W\|_F \lesssim \|(S^{-T} \otimes I_m) - (S^{-T} \otimes W)K_S^{-1} K_{AMN}\|_2 \|\Delta A\|_F,$$

and

$$(3.16) \qquad \mathrm{vec}(|\Delta S|) \lesssim |K_S^{-1} K_{AMN}| \mathrm{vec}(|\Delta A|),$$

$$(3.17) \qquad \mathrm{vec}(|\Delta W|) \lesssim \left|(S^{-T} \otimes I_m) - (S^{-T} \otimes W)K_S^{-1} K_{AMN}\right| \mathrm{vec}(|\Delta A|),$$

where K_S *and* K_{AMN} *are defined by* (3.7).

3.2. Condition numbers for the factor S. By Lemma 2.3 and Theorem 3.1, we get the expressions of the condition numbers for the factor S.

THEOREM 3.3. *With the same assumptions as in Theorem* 3.1, *we have*

$$(3.18) \qquad \kappa_S(A) = \frac{\|K_S^{-1} K_{AMN}\|_2 \|A\|_F}{\|S\|_F},$$

$$(3.19) \qquad m_S(A) = \frac{\||K_S^{-1} K_{AMN}| \mathrm{vec}(|A|)\|_\infty}{\|\mathrm{vec}(S)\|_\infty},$$

$$(3.20) \qquad c_S(A) = \left\| \frac{|K_S^{-1} K_{AMN}| \mathrm{vec}(|A|)}{\mathrm{vec}(|S|)} \right\|_\infty,$$

where K_S *and* K_{AMN} *are defined by* (3.7).

In Theorem 3.3, we see that the expressions of these three condition numbers contain many Kronecker products and the vec-permutation matrix Π, which require large computer storage and high computational complexity when the size of the given matrix is large. It may be expensive to compute them. For reducing the computer storage and the computational complexity, we give some upper bounds for these three condition numbers.

COROLLARY 3.4. *With the same assumptions as in Theorem* 3.1, *we have*

$$(3.21) \qquad \kappa_S(A) \le \frac{\|K_S^{-1}\|_2 \|A\|_F (\|N^{-1} A^T M\|_2 + \|MA\|_2 \|N^{-1}\|_2)}{\|S\|_F},$$

$$(3.22) \qquad m_S(A) \le \frac{\|K_S^{-1}\|_\infty \|2|N^{-1}||A|^T |M||A|\|_{\max}}{\|S\|_{\max}},$$

$$(3.23) \qquad c_S(A) \le \left\| \mathrm{diag}^{\ddagger}(\mathrm{vec}(S))K_S^{-1} \right\|_\infty \|2|N^{-1}||A|^T |M||A|\|_{\max},$$

where K_S *is defined by* (3.7).

$$\boxed{\textbf{ELA}}$$

Ze-Jia Xie, Wen Li, and Xiao-Qing Jin

Proof. By (2.1), we have

$$\begin{aligned}
\|K_S^{-1}K_{AMN}\|_2 &\le \|K_S^{-1}\|_2\|(I_n \otimes N^{-1}A^T M) + (A^T M^T \otimes N^{-1})\Pi\|_2 \\
&\le \|K_S^{-1}\|_2(\|I_n \otimes N^{-1}A^T M\|_2 + \|A^T M^T \otimes N^{-1}\|_2) \\
&= \|K_S^{-1}\|_2(\|N^{-1}A^T M\|_2 + \|MA\|_2\|N^{-1}\|_2),
\end{aligned}$$

which, combining (3.18), leads to (3.21). As for (3.22), it can be obtained from (3.19) and

$$\begin{aligned}
&\||K_{AMN}|\mathrm{vec}(|A|)\|_\infty \\
=&\||(I_n \otimes N^{-1}A^T M) + (A^T M^T \otimes N^{-1})\Pi|\mathrm{vec}(|A|)\|_\infty \\
\le&\|(I_n \otimes |N^{-1}||A|^T|M|)\mathrm{vec}(|A|) + (|A|^T|M|^T \otimes |N^{-1}|)\mathrm{vec}(|A|^T)\|_\infty \\
=&\|\mathrm{vec}(|N^{-1}||A|^T|M||A|) + \mathrm{vec}(|N^{-1}||A|^T|M||A|)\|_\infty \\
\text{(3.24)} \qquad =&\|2|N^{-1}||A|^T|M||A|\|_{\max}.
\end{aligned}$$

Also, according to (2.2) and (3.24), we have

$$\begin{aligned}
c_S(A) &= \left\|\mathrm{diag}^{\ddagger}(\mathrm{vec}(|S|))|K_S^{-1}K_{AMN}|\mathrm{vec}(|A|)\right\|_\infty \\
&\le \left\|\mathrm{diag}^{\ddagger}(\mathrm{vec}(S))K_S^{-1}\right\|_\infty \||K_{AMN}|\mathrm{vec}(|A|)\|_\infty \\
&\le \left\|\mathrm{diag}^{\ddagger}(\mathrm{vec}(S))K_S^{-1}\right\|_\infty \|2|N^{-1}||A|^T|M||A|\|_{\max},
\end{aligned}$$

which implies that (3.23) holds. □

3.3. Condition numbers for the factor W. By Lemma 2.3 and Theorem 3.1, we get the expressions of the condition numbers for the factor W.

THEOREM 3.5. *With the same assumptions as in Theorem 3.1, we have*

$$\text{(3.25)} \qquad \kappa_W(A) = \frac{\|(S^{-T} \otimes I_m) - K_{SW}K_{AMN}\|_2\|A\|_F}{\|W\|_F},$$

$$\text{(3.26)} \qquad m_W(A) = \frac{\||(S^{-T} \otimes I_m) - K_{SW}K_{AMN}|\mathrm{vec}(|A|)\|_\infty}{\|\mathrm{vec}(W)\|_\infty},$$

$$\text{(3.27)} \qquad c_W(A) = \left\|\frac{|(S^{-T} \otimes I_m) - K_{SW}K_{AMN}|\mathrm{vec}(|A|)}{\mathrm{vec}(|W|)}\right\|_\infty,$$

where

$$\text{(3.28)} \qquad K_{SW} = (S^{-T} \otimes W)K_S^{-1},$$

and K_S and K_{AMN} are defined by (3.7).

ELA

http://math.technion.ac.il/iic/ela

The following corollary gives computable upper bounds for these three condition numbers.

COROLLARY 3.6. *With the same assumptions as in Theorem* 3.1, *we have*

(3.29)
$$\kappa_W(A) \leq \frac{\|S^{-1}\|_2 \|A\|_F \left[1 + \|W\|_2 \|K_S^{-1}\|_2 (\|N^{-1}A^T M\|_2 + \|MA\|_2 \|N^{-1}\|_2)\right]}{\|W\|_F},$$

(3.30)
$$m_W(A) \leq \frac{\||A||S^{-1}|\|_{\max} + \|K_{SW}\|_\infty \|2|N^{-1}||A|^T|M||A|\|_{\max}}{\|W\|_{\max}},$$

(3.31)
$$c_W(A) \leq \left\|\frac{|A||S^{-1}|}{|W|}\right\|_{\max} + \left\|\mathrm{diag}^\ddagger(\mathrm{vec}(W))K_{SW}\right\|_\infty \|2|N^{-1}||A|^T|M||A|\|_{\max},$$

where K_S *and* K_{SW} *are defined by* (3.7) *and* (3.28), *respectively.*

The proof of this corollary is similar to the one of Corollary 3.4.

REMARK 3.7. Let $\sigma_1 \geq \cdots \geq \sigma_{n-1} \geq \sigma_n$ be n nonzero singular values of $A \in \mathbb{R}^{m \times n}$ with full column rank, and $\kappa(A) = \sigma_1/\sigma_n$ be the generalized condition number of A. In [7], the authors gave the absolute normwise condition numbers for the polar decomposition (a special CGPD where $M = I_m$ and $N = I_n$) of a full column rank real matrix A. Their results are showed in Table 1.

Table 1. Absolute normwise condition numbers.

Factor W	$m = n$	$2/(\sigma_n + \sigma_{n-1})$
	$m > n$	$1/\sigma_n$
Factor S	$m \geq n$	$\sqrt{2}\frac{(1+\kappa(A)^2)^{1/2}}{1+\kappa(A)}$

While the ones we give in (3.18) and (3.25) are the relative normwise condition numbers, we change them into the absolute ones as follows:

(3.32) $$\kappa_S^{abs}(A) = \|K_S^{-1} K_{AMN}\|_2,$$
(3.33) $$\kappa_W^{abs}(A) = \|(S^{-T} \otimes I_m) - K_{SW} K_{AMN}\|_2,$$

where K_S, K_{AMN} and K_{SW} are defined by (3.7) and (3.28), respectively. We can see that the explicit expressions of normwise condition numbers in Table 1 are obviously different from (3.32) and (3.33) (where $M = I_m$ and $N = I_n$). Actually, no matter

ELA

http://math.technion.ac.il/iic/ela

by the implicit definition of the absolute normwise condition number or numerical computation, we can find that their values are the same as (3.32) and (3.33). For example, for the factor W, the implicit definition of its absolute normwise condition number, given in [7], is

$$\kappa(A, W) = \lim_{\varepsilon \to 0} \sup_{\|\Delta A\|_F \leq \varepsilon} \frac{\|\Delta W\|_F}{\|\Delta A\|_F}.$$

It is obvious that

$$\kappa(A, W) = \lim_{\varepsilon \to 0} \sup_{\|\mathrm{vec}(\Delta A)\|_2 \leq \varepsilon} \frac{\|\mathrm{vec}(\Delta W)\|_2}{\|\mathrm{vec}(\Delta A)\|_2} = \kappa_W^{abs}(A).$$

4. Numerical examples. In this section, we consider the following examples to demonstrate our theoretical results.

EXAMPLE 4.1. We use the example given in [21, Remark 5] to show that sometimes the mixed and componentwise condition numbers may be more accurate than the normwise one. Let $M = I_3$, $N = I_2$ and

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0.000008 \\ 0 & 0 \end{bmatrix}.$$

Then the CGPD (also the polar decomposition) of A is as follows:

$$W = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad S = \begin{bmatrix} 1 & 0 \\ 0 & 0.000008 \end{bmatrix},$$

and the values of their condition numbers are showed in Table 2.

Table 2. Values of three condition numbers.

$\kappa_W(A)$	$m_W(A)$	$c_W(A)$	$\kappa_S(A)$	$m_S(A)$	$c_S(A)$
$8.8388e + 004$	0	0	1.4142	1	1

From Table 2, we see that the normwise condition number of W is large, while the mixed and componentwise ones of W are both zero. Now let A suffer from a small perturbation (the zero entries are not perturbed), and we obtain

$$\widetilde{A} = \begin{bmatrix} 1.000002 & 0 \\ 0 & 0.000007 \\ 0 & 0 \end{bmatrix}.$$

ELA

The CGPD of $\widetilde{A}$ is as follows:

$$\widetilde{W} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \widetilde{S} = \begin{bmatrix} 1.000002 & 0 \\ 0 & 0.000007 \end{bmatrix}.$$

Next let A suffer from another small perturbation (some zero entries are perturbed), and we obtain

$$\widehat{A} = \begin{bmatrix} 1 & 0 \\ 0 & 0.000008 \\ 0 & 0.000006 \end{bmatrix}.$$

The CGPD of $\widehat{A}$ is as follows:

$$\widehat{W} = \begin{bmatrix} 1 & 0 \\ 0 & 0.8 \\ 0 & 0.6 \end{bmatrix}, \quad \widehat{S} = \begin{bmatrix} 1 & 0 \\ 0 & 0.00001 \end{bmatrix}.$$

The former perturbation example confirms that the mixed and componentwise condition numbers of W are zero, and the later one verifies that the normwise condition number of W is large. Actually, from Remark 2.2, we see that if the zero entries of input data are not perturbed, then the mixed and componentwise condition numbers would be more accurate than the normwise one.
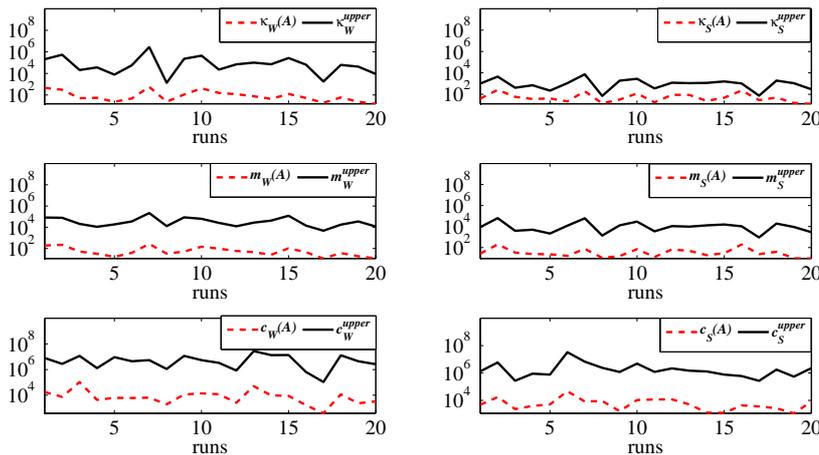

FIG. 4.1. *Condition numbers and their upper bounds.*

EXAMPLE 4.2. Assume $M \in \mathbb{R}^{18 \times 18}$ and $N \in \mathbb{R}^{12 \times 12}$, which are given by MATLAB function *randn* (20 runs), both are nonsingular symmetric matrices. It is

ELA

http://math.technion.ac.il/iic/ela

obvious that M and N form an orthosymmetric pair. Suppose $A \in \mathbb{R}^{18 \times 12}$ is given by function *randn* (20 runs) such that $A^{\star M,N} A$ has no nonpositive real eigenvalues. Denote the upper bounds for condition numbers $\kappa_W(A)$, $m_W(A)$, $c_W(A)$, $\kappa_S(A)$, $m_S(A)$ and $c_S(A)$ by κ_W^{upper}, m_W^{upper}, c_W^{upper}, κ_S^{upper}, m_S^{upper} and c_S^{upper}, respectively. Figure 4.1 shows the condition numbers and their upper bounds. Unfortunately, these upper bounds seem to be not sharp enough.

REFERENCES

[1] A. Barrlund. Perturbation bounds on the polar decomposition. *BIT*, 30:101–113, 1989.

[2] R. Bhatia. *Matrix Analysis*. Springer, New York, 1997.

[3] R. Bhatia and K. Mukherjea. Variation of the unitary part of a matrix. *SIAM J. Matrix Anal. Appl.*, 15:1007–1014, 1994.

[4] Y. Bolshakov, C.V.M. van der Mee, A.C.M. Ran, B. Reichstein, and L. Rodman. Polar decompositions in finite dimensional indefinite scalar product spaces: General Theory. *Linear Algebra Appl.*, 261:91–141, 1997.

[5] Y. Bolshakov, C.V.M. van der Mee, A.C.M. Ran, B. Reichstein, and L. Rodman. Extension of isometries in finite-dimensional indefinite scalar product spaces and polar decompositions. *SIAM J. Matrix Anal. Appl.*, 18:752–774, 1997.

[6] Y. Bolshakov, C.V.M. van der Mee, A.C.M. Ran, B. Reichstein, and L. Rodman. Polar decompositions in finite-dimensional indefinite scalar product spaces: special cases and applications. In: I. Gohberg, P. Lancaster, and P.N. Shivakumar (editors), *Recent Developments in Operator Theory and its Applications*, Birkhäuser, Basel, 1996.

[7] F. Chaitin-Chatelin and S. Gratton. On the condition numbers associated with the polar factorization of a matrix. *Numer. Linear Algebra Appl.*, 7:337–354, 2000.

[8] X.-S. Chen and W. Li. Variations for Q- and H-fators in the polar decomposition. *Calcolo*, 45:99–109, 2008.

[9] X.-S. Chen, W. Li, and W.-W. Sun. Some new perturbation bounds for generalized polar decompositions. *BIT*, 44:237–244, 2004.

[10] F. Cucker, H.-A. Diao, and Y.-M. Wei. On mixed and componentwise condition numbers for Moore-Penrose inverse and linear least squares problems. *Math. Comput.*, 76:947–963, 2007.

[11] H.-A. Diao, W.-G. Wang, Y.-M. Wei, and S.-Z. Qiao. On condition numbers for Moore-Penrose inverse and linear least squares problem involving Kronecker products. *Numer. Linear Algebra Appl.*, 20:44–59, 2013.

[12] I. Gohberg and I. Koltracht. Mixed, componentwise, and structured condition numbers. *SIAM J. Matrix Anal. Appl.*, 14:688–704, 1993.

[13] N.J. Higham. *Accuracy and Stability of Numerical Algorithms*. SIAM, Philadelphia, PA, 2002.

[14] N.J. Higham. Computing the polar decomposition with applications. *SIAM J. Sci. Statist. Comput.*, 7:1160–1174, 1986.

[15] N.J. Higham. The matrix sign decomposition and its relation to the polar decomposition. *Linear Algebra Appl.*, 212/213:3–20, 1994.

[16] N.J. Higham, C. Mehl, and F. Tisseur. The canonical generalized polar decomposition. *SIAM J. Matrix Anal. Appl.*, 31:2163–2180, 2010.

[17] R.A. Horn and C.R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1991.

ELA

http://math.technion.ac.il/iic/ela

[18] C. Kenney and A.J. Laub. Polar decomposition and matrix sign function condition estimates. *SIAM J. Sci. Stat. Comput.*, 12:488–504, 1991.

[19] U. Kintzel. Procrustes problems in finite dimensional indefinite scalar product spaces. *Linear Algebra Appl.*, 402:1–28, 2005.

[20] R.-C. Li. A perturbation bounds for the generalized polar decomposition. *BIT*, 33:304–308, 1993.

[21] R.-C. Li. New perturbation bounds for the unitary polar factor. *SIAM J. Matrix Anal. Appl.*, 16:327–332, 1995.

[22] R.-C. Li. Relative perturbation bounds for positive polar factors of graded matrices. *SIAM J. Matrix Anal. Appl.*, 27:424–433, 2005.

[23] W. Li and W.-W. Sun. Perturbation bounds for unitary and subunitary polar factors. *SIAM J. Matrix Anal. Appl.*, 23:1183–1193, 2002.

[24] W. Li and W.-W. Sun. New perturbation bounds for unitary polar factors. *SIAM J. Matrix Anal. Appl.*, 25:362–372, 2003.

[25] W. Li and W.-W. Sun. Some remarks on perturbation of polar decomposition for rectangular matrices. *Numer. Linear Algebra Appl.*, 13:327–328, 2006.

[26] H.-Y. Li and H. Yang. Relative perturbation bounds for weighted polar decomposition. *Comput. Math. Appl.*, 59:853–860, 2010.

[27] R. Mathias. Perturbation bounds for the polar decomposition. *SIAM J. Matrix Anal. Appl.*, 14:588–593, 1993.

[28] J.R. Rice. A theory of condition. *SIAM J. Numer. Anal.*, 3:287–310, 1966.

[29] H. Xiang and Y.-M. Wei. Structured mixed and componentwise condition numbers of some structured matrices. *J. Comput. Appl. Math.*, 202:217–229, 2007.

[30] H. Yang and H.-Y. Li. Perturbation bounds for weighted polar decomposition in the weighted unitarily invariant norm. *Numer. Linear Algebra Appl.*, 15:685–700, 2008.

[31] H. Yang and H.-Y. Li. Weighted polar decomposition and WGL partial ordering of rectangular complex matrices. *SIAM J. Matrix Anal. Appl.*, 30:898–924, 2008.