Linear Algebra and its Applications 313 (2000) 41–51

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

A new relative perturbation theorem for singular subspaces

Ren-Cang Li [a,*,1], G.W. Stewart [b,2]

[a] *Department of Mathematics, 715 Patterson Office Tower, University of Kentucky, Lexington, KY 40506, USA*

[b] *Department of Computer Science, University of Maryland, College Park, MD 20742-3255, USA*

Received 25 January 1999; accepted 7 February 2000

Submitted by H. Schneider

Abstract

This note addresses the sensitivity of singular subspaces of a matrix under relative perturbations. It employs a new technique of separating a multiplicative perturbation D into two components: one is the distance of a scalar multiple of D to the nearest unitary matrix Q and the other is the distance of Q to the identity. Consequently, the new bounds reflect the intrinsic differences in how left and right multiplicative perturbations affect left and right singular subspaces. © 2000 Elsevier Science Inc. All rights reserved.

AMS classification: 15A18; 15A42; 65F15; 65F35; 65G99

Keywords: Multiplicative perturbation; Relative perturbation theory; Relative gap; Singular subspace

1. Introduction

The purpose of this note is to extend a result on the sensitivity of singular subspaces of a matrix to relative perturbations in the matrix (Theorem 2.2). We begin by setting the background and notation.

* Corresponding author.

E-mail addresses: rcli@ms.uky.edu (R.-C. Li), stewart@cs.umd.edu (G.W. Stewart).

[1] This work was supported in part by the National Science Foundation under Grant No. ACI-9721388 and by the National Science Foundation CAREER award under Grant No. CCR-9875201.

[2] This work was supported in part by the National Science Foundation under Grant No. 970909-8426.

PII: S 0 0 2 4 - 3 7 9 5 (0 0) 0 0 0 7 4 - 4

Let A be an $m \times n$ matrix with $m \geqslant n$, and let A have the singular value decomposition

$$A = (U_1 \quad U_2 \quad U_3) \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1^* \\ V_2^* \end{pmatrix},$$

where

$$\Sigma_1 = \mathrm{diag}(\sigma_1, \ldots, \sigma_k) \geqslant 0 \quad \text{and} \quad \Sigma_2 = \mathrm{diag}(\sigma_{k+1}, \ldots, \sigma_n) \geqslant 0,$$

$(U_1 \quad U_2\ U_3)$ and $(V_1 \quad V_2)$ are unitary and the asterisk denotes the conjugate transpose. Note that we do not impose any particular ordering on the singular values σ_i of A. For later reference set

$$U_\perp = (U_2 \quad U_3),$$

and define

$$\mathscr{U}_1 = \mathscr{R}(U_1), \quad \mathscr{U}_\perp = \mathscr{R}(U_\perp), \quad \text{and} \quad \mathscr{V}_1 = \mathscr{R}(V_1), \quad \mathscr{V}_\perp = \mathscr{R}(V_2),$$

where $\mathscr{R}(\cdot)$ represents the subspace spanned by the columns of a matrix.

The spaces $\mathscr{U}_1$ and $\mathscr{U}_\perp$ are complementary left singular subspaces of A. Likewise, $\mathscr{V}_1$ and $\mathscr{V}_\perp$ are complementary right singular subspaces of A. The purpose of this note is to derive bounds on the changes in these subspaces when A is replaced by a perturbation $\tilde{A}$. We will assume that $\tilde{A}$ has the singular value decomposition

$$\tilde{A} = (\tilde{U}_1 \ \ \tilde{U}_2 \ \ \tilde{U}_3) \begin{pmatrix} \tilde{\Sigma}_1 & 0 \\ 0 & \tilde{\Sigma}_2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{V}_1^* \\ \tilde{V}_2^* \end{pmatrix},$$

and will focus on the relation of the spaces $\tilde{\mathscr{U}}_1 = \mathscr{R}(\tilde{U}_1)$ and $\tilde{\mathscr{V}}_1 = \mathscr{R}(\tilde{V}_1)$ of $\tilde{A}$ to the spaces $\mathscr{U}_1$ and $\mathscr{V}_1$ of A. Since the spaces $\mathscr{U}_\perp$ and $\mathscr{V}_\perp$ are unitary complements of $\mathscr{U}_1$ and $\mathscr{V}_1$, perturbation bounds for the latter will apply without change to the former.

We will measure the difference between, say, $\mathscr{V}_1$ and $\tilde{\mathscr{V}}_1$ by the quantity

$$\| \sin \Theta(\mathscr{V}_1, \tilde{\mathscr{V}}_1) \|_{\mathrm{F}} = \| \tilde{V}_2^* V_1 \|_{\mathrm{F}} = \| V_2^* \tilde{V}_1 \|_{\mathrm{F}}, \tag{1.1}$$

where $\| \cdot \|_{\mathrm{F}}$ denotes the Frobenius matrix norm. As the left-hand side of (1.1) suggests, this measure is related to the canonical angles between the subspaces $\mathscr{V}_1$ and $\tilde{\mathscr{V}}_1$. Specifically, the measure is the square root of the sum of squares of the sines of the canonical angles. This measure is also a metric and hence satisfies the triangle inequality. (For more on canonical angles see [5].)

If Σ_1 and Σ_2 have a common singular value, the subspaces $\mathscr{U}_1$ and $\mathscr{V}_1$ are not uniquely defined. Hence we must posit a gap between the two sets of singular values. Moreover, if $m > n$, then A has additional left null vectors that do not correspond to Σ_1 and Σ_2. These additional null vectors can be regarded as corresponding to additional "honorary" zero singular values. This leads to the following definitions:

$$\mathscr{S}_1 = \{\sigma_1, \ldots, \sigma_k\},$$

and

$$\mathscr{S}_\perp = \begin{cases} \{\sigma_{k+1}, \ldots, \sigma_n\} & \text{if } m = n, \\ \{\sigma_{k+1}, \ldots, \sigma_n, 0\} & \text{if } m > n, \end{cases}$$

and analogously for $\tilde{\mathscr{S}}_1$, $\tilde{\mathscr{S}}_\perp$. Our separation hypothesis will be stated in terms of the sets $\mathscr{S}_1$ and $\tilde{\mathscr{S}}_\perp$ (or of the sets $\tilde{\mathscr{S}}_1$ and $\mathscr{S}_\perp$). When A and $\tilde{A}$ are close, those gaps should be comparable to the gap between the two sets of singular values in Σ_1 and Σ_2.

The paper is organized as follows. In Section 2, we review the relevant literature and motivate our new bounds. In Section 3, we establish our main results. These results depend on two free parameters, and in Section 4, we show how to choose them to optimize the bound. The paper concludes with an example and a brief summary.

2. Motivation

The classic theorem on the perturbation of singular subspaces is due to Wedin [6]. In stating it we assume the above notation and definitions.

Theorem 2.1. *Let*

$$\tilde{A} = A + E,$$

and let

$$\delta = \min|\mathscr{S}_1 - \tilde{\mathscr{S}}_\perp| \equiv \min\{|\mu_1 - \tilde{\mu}_\perp| : \mu_1 \in \mathscr{S}_1, \tilde{\mu}_\perp \in \tilde{\mathscr{S}}_\perp\}.$$

Set

$$R = \tilde{A}V_1 - U_1\Sigma_1 \equiv EV_1 \quad \textit{and} \quad S = \tilde{A}^*U_1 - V_1\Sigma_1 \equiv E^*U_1.$$

Then

$$\sqrt{\|\sin\Theta(\mathscr{U}_1, \tilde{\mathscr{U}}_1)\|_F^2 + \|\sin\Theta(\mathscr{V}_1, \tilde{\mathscr{V}}_1)\|_F^2} \leqslant \frac{\sqrt{\|R\|_F^2 + \|S\|_F^2}}{\delta}.$$

From the statement of Wedin's theorem we see that the perturbation in A is additive or absolute; that is, it is of the form $\tilde{A} = A + E$, where E is presumed small. Moreover, the gap δ is simply the distance between the sets $\mathscr{S}_1$ and $\tilde{\mathscr{S}}_\perp$. In this paper, our concern is with multiplicative or relative perturbations of the form $\tilde{A} = D_L^* A D_R$, where D_L and D_R are nonsingular. Clearly certain restrictions must be placed on D_L and D_R in order for singular subspaces to have small changes. The following theorem [4] shows how these relative perturbations in A affect singular subspaces when D_L and D_R are assumed to be near the identity matrices. Note that it uses a relative version of Wedin's gap δ.

Theorem 2.2. *Let*

$$\tilde{A} = D_L^* A D_R,$$

where $D_{\rm L}$ *and* $D_{\rm R}$ *are nonsingular. Let*

$$\rho = \min_{\mu_1\in\mathscr{S}_1,\tilde{\mu}_\perp\in\tilde{\mathscr{S}}_\perp} \frac{|\mu_1-\tilde{\mu}_\perp|}{\sqrt{\mu_1^2+\tilde{\mu}_\perp^2}}. \tag{2.1}$$

If $\rho > 0$, *then*

$$\sqrt{\|\sin\Theta(\mathscr{U}_1,\tilde{\mathscr{U}}_1)\|_{\rm F}^2+\|\sin\Theta(\mathscr{V}_1,\tilde{\mathscr{V}}_1)\|_{\rm F}^2} \leqslant \frac{\sqrt{\|(I-D_{\rm L}^*)U_1\|_{\rm F}^2+\|(I-D_{\rm L}^{-1})U_1\|_{\rm F}^2+\|(I-D_{\rm R}^*)V_1\|_{\rm F}^2+\|(I-D_{\rm R}^{-1})V_1\|_{\rm F}^2}}{\rho}. \tag{2.2}$$

The bound in the theorem, like that in Wedin's theorem, consists of an error expression (in this case involving the deviation of $D_{\rm L}$ and $D_{\rm R}$ from the identity) divided by a measure of the gap between the singular values. The bound produces reasonably sharp results when $D_{\rm L}$ and $D_{\rm R}$ are nearly identity matrices. In other cases, however, the bound can be pessimistic. There are two problems.

The first problem can be illustrated by a simple example. Let $D_{\rm L}, D_{\rm R} = 2I$. Then the numerator of the bound (2.2) becomes

$$\frac{\sqrt{5}}{2}\sqrt{\|U_1\|_{\rm F}^2+\|V_1\|_{\rm F}^2},$$

which is not small. Since the denominator cannot be larger than $\sqrt{2}$, the entire bound is large. But the singular subspaces of A and $D_{\rm L}AD_{\rm R}$ are the same.

More generally, we can write

$$\alpha\beta\tilde{A} = (\alpha D_{\rm L})^*A(\beta D_{\rm R}).$$

Since the singular subspaces of $\alpha\beta\tilde{A}$ are the same as those of $\tilde{A}$, we may replace $D_{\rm L}$ and $D_{\rm R}$ in (2.2) by $\alpha D_{\rm L}$ and $\beta D_{\rm R}$. In effect, the bound (2.2) contains two free parameters, and unless we optimize the bound with respect to them, we cannot expect the best results.

The second problem can also be illustrated by an example. Let $D_{\rm L}$ be unitary but not near the identity matrix and let $D_{\rm R} = I$. Then the right singular subspaces of $D_{\rm L}^*AD_{\rm R}$ are the same as those of A; but the bound (2.2), which contains the sum $\|(I-D_{\rm L}^*)U_1\|_{\rm F}^2+\|(I-D_{\rm L}^{-1})U_1\|_{\rm F}^2$ cannot be small. When $D_{\rm L}$ is near a unitary matrix and $D_{\rm R}$ is near an identity matrix, the right singular subspaces change, but only slightly. However, we know of no published result that gives direct bounds on this error. [3]

[3] There are some complicated bounds in the technical report version of [4]. It is available as LAPACK Working Note 85 at `http://www.netlib.org/lapack/lawns/lawn85.ps`.

3. The main result

We propose to solve these two problems raised in Section 2 simultaneously. Let α and β be two positive numbers, which will be determined later to optimize our bounds. Let Q_{L} be the unitary matrix nearest αD_{L}. It turns out that Q_{L} is the unitary factor in the polar decompositions of D_{L} and is independent of α (see, e.g. [1, p. 276]). Similarly let Q_{R} be the unitary matrix nearest βD_{R}. Set

$$E_{\mathrm{L}} = \alpha D_{\mathrm{L}} - Q_{\mathrm{L}} \quad \text{and} \quad E_{\mathrm{R}} = \beta D_{\mathrm{R}} - Q_{\mathrm{R}}.$$

Then

$$\alpha\beta\tilde{A} = (\alpha D_{\mathrm{L}})^* A(\beta D_{\mathrm{R}}) = (I + E_{\mathrm{L}}^* Q_{\mathrm{L}}) Q_{\mathrm{L}}^* A Q_{\mathrm{R}} (I + Q_{\mathrm{R}}^* E_{\mathrm{R}}). \tag{3.1}$$

Now the singular subspaces of $Q_{\mathrm{L}}^* A Q_{\mathrm{R}}$ are $Q_{\mathrm{L}}^*\mathscr{U}_1$ and $Q_{\mathrm{R}}^*\mathscr{V}_1$, and in a moment we will show how to bound their deviations from $\mathscr{U}_1$ and $\mathscr{V}_1$. Note that this can be done separately for $\mathscr{U}_1$ and $\mathscr{V}_1$. We can then apply Theorem 2.2 with $D_{\mathrm{L}} = I + Q_{\mathrm{L}}^* E_{\mathrm{L}}$ and $D_{\mathrm{R}} = I + Q_{\mathrm{R}}^* E_{\mathrm{R}}$ to bound the deviation of $\tilde{\mathscr{U}}_1$ and $\tilde{\mathscr{V}}_1$ from $Q_{\mathrm{L}}^*\mathscr{U}_1$ and $Q_{\mathrm{R}}^*\mathscr{V}_1$. In this way we arrive at bound on, say, $\tilde{\mathscr{V}}_1$, that is the sum of a joint bound involving the distance of βD_{R} from Q_{R} and a separate bound that depends only on Q_{R}. It is easily seen that with proper choice of α and β this approach disposes of the two examples given above.

We will now show how to bound, say, $\|\sin\Theta(\mathscr{U}_1, Q_{\mathrm{L}}^*\mathscr{U}_1)\|_{\mathrm{F}}$.

Lemma 3.1. *Let the columns of U form an orthonormal basis for the subspace $\mathscr{U}$ and let the columns of $U_\perp$ form an orthonormal basis for the unitary complement of $\mathscr{U}$. If Q is unitary, then*

$$\|\sin\Theta(\mathscr{U}, Q^*\mathscr{U})\|_{\mathrm{F}} = \begin{cases} \|U_\perp^*(I - Q^*)U\|_{\mathrm{F}} \leqslant \|(I - Q^*)U\|_{\mathrm{F}}, \\ \|U_\perp^*(I - Q)U\|_{\mathrm{F}} \leqslant \|(I - Q)U\|_{\mathrm{F}}. \end{cases}$$

Proof. By (1.1),

$$\begin{aligned} &\|\sin\Theta(\mathscr{U}, Q^*\mathscr{U})\|_{\mathrm{F}} \\ &\quad = \begin{cases} \|U_\perp^*(Q^*U)\|_{\mathrm{F}} = \|U_\perp^*(I - Q^*)U\|_{\mathrm{F}} \leqslant \|(I - Q^*)U\|_{\mathrm{F}}, \\ \|(Q^*U_\perp)^*U\|_{\mathrm{F}} = \|U_\perp^*(I - Q)U\|_{\mathrm{F}} \leqslant \|(I - Q)U\|_{\mathrm{F}}. \end{cases} \end{aligned}$$ □

We are now in a position to state our main result.

Theorem 3.2. *Let*

$$\tilde{A} = D_{\mathrm{L}}^* A D_{\mathrm{R}},$$

where D_{L} and D_{R} are nonsingular, and for $\alpha, \beta > 0$. Let

$$\rho_{\alpha,\beta} = \min_{\mu_1 \in \mathscr{S}_1, \tilde{\mu}_\perp \in \tilde{\mathscr{S}}_\perp} \frac{|\mu_1 - \alpha\beta\tilde{\mu}_\perp|}{\sqrt{\mu_1^2 + (\alpha\beta)^2\tilde{\mu}_\perp^2}}.$$

Define $\epsilon_{\alpha,\beta} \geqslant 0$ *by*

$$\epsilon_{\alpha,\beta}^2 = \left\|(\alpha D_{\mathrm{L}} - Q_{\mathrm{L}})^* U_1\right\|_{\mathrm{F}}^2 + \left\|\left[(\alpha D_{\mathrm{L}})^{-1} - Q_{\mathrm{L}}^*\right] U_1\right\|_{\mathrm{F}}^2 + \left\|(\beta D_{\mathrm{R}} - Q_{\mathrm{R}})^* V_1\right\|_{\mathrm{F}}^2 + \left\|\left[(\beta D_{\mathrm{R}})^{-1} - Q_{\mathrm{R}}^*\right] V_1\right\|_{\mathrm{F}}^2. \tag{3.2}$$

If the number $\rho_{\alpha,\beta}$ *is positive, then*

$$\|\sin\Theta(\mathscr{U}_1, \tilde{\mathscr{U}}_1)\|_{\mathrm{F}} \leqslant \frac{\epsilon_{\alpha,\beta}}{\rho_{\alpha,\beta}} + \|(I - Q_{\mathrm{L}}^*)U_1\|_{\mathrm{F}}, \tag{3.3}$$

and

$$\|\sin\Theta(\mathscr{V}_1, \tilde{\mathscr{V}}_1)\|_{\mathrm{F}} \leqslant \frac{\epsilon_{\alpha,\beta}}{\rho_{\alpha,\beta}} + \|(I - Q_{\mathrm{R}}^*)V_1\|_{\mathrm{F}}. \tag{3.4}$$

Proof. From (3.1) and Theorem 2.2 applied to $Q_{\mathrm{L}}^* A Q_{\mathrm{R}}$ and $(I + E_{\mathrm{L}}^* Q_{\mathrm{L}}) Q_{\mathrm{L}}^* A Q_{\mathrm{R}} (I + Q_{\mathrm{R}}^* E_{\mathrm{R}})$, it follows that $\|\sin\theta(Q_{\mathrm{L}}\mathscr{U}_1, \tilde{\mathscr{U}}_1)\|_{\mathrm{F}} \leqslant \epsilon/\rho_{\alpha,\beta}$ and $\|\sin\theta(Q_{\mathrm{R}}\mathscr{V}_1, \tilde{\mathscr{V}}_1)\|_{\mathrm{F}} \leqslant \epsilon/\rho_{\alpha,\beta}$, where

$$\epsilon^2 = \|E_{\mathrm{L}}^* U_1\|_{\mathrm{F}}^2 + \|(I + Q_{\mathrm{L}}^* E_{\mathrm{L}})^{-1} Q_{\mathrm{L}}^* E_{\mathrm{L}} Q_{\mathrm{L}}^* U_1\|_{\mathrm{F}}^2 + \|E_{\mathrm{R}}^* V_1\|_{\mathrm{F}}^2 + \|(I + Q_{\mathrm{R}}^* E_{\mathrm{R}})^{-1} Q_{\mathrm{R}}^* E_{\mathrm{R}} Q_{\mathrm{R}}^* V_1\|_{\mathrm{F}}^2.$$

This ϵ is just $\epsilon_{\alpha,\beta}$, since

$$\begin{aligned}(I + Q_{\mathrm{L}}^* E_{\mathrm{L}})^{-1} Q_{\mathrm{L}}^* E_{\mathrm{L}} Q_{\mathrm{L}}^* &= (Q_{\mathrm{L}} + E_{\mathrm{L}})^{-1} Q_{\mathrm{L}} Q_{\mathrm{L}}^* E_{\mathrm{L}} Q_{\mathrm{L}}^* \\ &= (\alpha D_{\mathrm{L}})^{-1} (\alpha D_{\mathrm{L}} - Q_{\mathrm{L}}) Q_{\mathrm{L}}^* \\ &= Q_{\mathrm{L}}^* - (\alpha D_{\mathrm{L}})^{-1},\end{aligned}$$

and similarly $(I + Q_{\mathrm{R}}^* E_{\mathrm{R}})^{-1} Q_{\mathrm{R}}^* E_{\mathrm{R}} Q_{\mathrm{R}}^* = Q_{\mathrm{R}}^* - (\alpha D_{\mathrm{R}})^{-1}$. But from Lemma 3.1 it follows that $\|\sin\Theta(\mathscr{U}_1, Q_{\mathrm{L}}^*\mathscr{U}_1)\|_{\mathrm{F}} \leqslant \|(I - Q_{\mathrm{L}}^*)U_1\|_{\mathrm{F}}$ and $\|\sin\Theta(\mathscr{V}_1, Q_{\mathrm{R}}^*\mathscr{V}_1)\|_{\mathrm{F}} \leqslant \|(I - Q_{\mathrm{R}}^*)V_1\|_{\mathrm{F}}$. The theorem now follows from the fact that our measure of distance between subspaces satisfies the triangle inequality. □

Here are some comments on this theorem.

Note that by Lemma 3.1, $\|(I - Q_{\mathrm{L}}^*)U_1\|_{\mathrm{F}}$ may be replaced by $\|(I - Q_{\mathrm{L}})U_1\|_{\mathrm{F}}$ in (3.3) and $\|(I - Q_{\mathrm{R}}^*)V_1\|_{\mathrm{F}}$ by $\|(I - Q_{\mathrm{R}})V_1\|_{\mathrm{F}}$ in (3.4).

The final result of the theorem is the two bounds (3.3) and (3.4), the first bounding the perturbation in the subspace $\mathscr{U}_1$ and the second in the subspace $\mathscr{V}_1$. This is in contrast to Theorems 2.1 and 2.2, which give a single bound for both subspaces.

The quantity $\epsilon_{\alpha,\beta}$ is common to both bounds. It can be small only when D_{L} and D_{R} are near a multiple of a unitary matrix. Thus this term accounts for the effects of deviation from orthogonality. We will show later how to choose α and β to make this term small (if possible).

The second term is different for each bound. In the first, it is small when $I - Q_{\mathrm{L}}$ is small, i.e., when the optimal unitary approximation Q_{L} to D_{L} is near the identity matrix. From the bound [3, Theorem 1]

$$\|I - Q_{\mathrm{L}}\|_{\mathrm{F}} \leqslant 2\|I - \alpha D_{\mathrm{L}}\|_{\mathrm{F}},$$

we see that if D_{L} is near a multiple of an identity matrix, the term is small. Thus this term accounts for the effects of deviation from the identity. [4]

Putting the two terms together, we see that we can get a small bound for, say, $\mathscr{V}_1$ if D_{L} is near a multiple α of a unitary matrix and D_{R} is near a multiple β of the identity, provided we can obtain optimal (or nearly optimal) α and β. Note that if the nearly unitary matrix, in this case D_{L}, is not near an identity matrix, then conventional bounds (e.g., those in [2,4] and references therein) will be large. Thus our theorem represents an improvement on existing bounds in the literature.

4. Optimization of the bounds

We turn now to the approximation of α and β. The quantity $\epsilon_{\alpha,\beta}$ in Theorem 3.2 contains two free parameters, α and β, which may be used to optimize the bound. Note that α and β play independent roles in the definition of $\epsilon_{\alpha,\beta}$, and we may optimize with respect to each separately. To emphasize this we will drop the subscripts L and R and let γ stand for either α or β and W for U_1 or V_1.

The quantity $\rho_{\alpha,\beta}$ also contains the parameters α and β, and one could object that what we gain in optimizing $\epsilon_{\alpha,\beta}$ we could lose by $\rho_{\alpha,\beta}$ becoming smaller. Fortunately, the definition (3.2) implies that $\epsilon_{\alpha,\beta}$ can be small only if αD_{L} and βD_{R} are near unitary matrices, in which case the singular values of the matrix $\alpha\beta\tilde{A}$ are near those of A. More specifically,

$$\lim_{\epsilon_{\alpha,\beta}\to 0} \rho_{\alpha,\beta} = \min_{\mu_1\in\mathscr{S}_1,\mu_\perp\in\mathscr{S}_\perp} \frac{|\mu_1 - \mu_\perp|}{\sqrt{\mu_1^2 + \mu_\perp^2}}. \tag{4.1}$$

Thus optimizing $\epsilon_{\alpha,\beta}$ drives $\rho_{\alpha,\beta}$ toward the relative gap between $\mathscr{S}_1$ and $\mathscr{S}_2$, which is what we should hope for.

Turning now to the optimization of $\epsilon_{\alpha,\beta}$, we note that its definition contains terms of the form

$$\begin{aligned}&\|(\gamma D - Q)^* W\|_{\mathrm{F}}^2 + \|[(\gamma D)^{-1} - Q^*]W\|_{\mathrm{F}}^2\\ &\quad\leqslant \|\gamma D - Q\|_{\mathrm{F}}^2 + \|(\gamma D)^{-1} - Q^*\|_{\mathrm{F}}^2.\end{aligned} \tag{4.2}$$

[4] It is possible for $I - Q_{\mathrm{L}}$ to be small when D_{L} is not near a multiple of the identity, but then $\alpha D_{\mathrm{L}} - Q_{\mathrm{L}}$ must be large, and hence $\epsilon_{\alpha,\beta}$ is also large.

Optimizing the left-hand side of this relation appears to be an intractable problem. The right-hand side, however, can be optimized. Specifically, the optimal Q is the unitary factor of the polar decomposition of D. Unfortunately, the optimal value of γ is the root of a quartic equation with no convenient closed form solution.

Fortunately, the two terms in the bound are such that one cannot be small unless the other is approximately the same size. This is made precise in (4.6). The common sense of the matter is that if a matrix is near a unitary matrix, its inverse must be approximately as near to the transpose of the unitary matrix. This means that optimizing one of the terms essentially optimizes the other.

We will therefore choose γ to solve the problem

$$\begin{array}{ll} \text{minimize} & \|\gamma D - Q\|_{\mathrm{F}} \\ \text{subject to} & Q \text{ unitary and } \gamma \geqslant 0. \end{array} \tag{4.3}$$

We have already noted that the value of Q must be the unitary factor of the polar decomposition of D. That given, the value of γ is easily determined.

Theorem 4.1. *Let $\tau_1, \ldots, \tau_n$ be the singular values of D. The matrix Q in the solution of* (4.3) *is the unitary factor of the polar decomposition of D. The value of γ is given by*

$$\gamma = \frac{\sum \tau_i}{\sum \tau_i^2}. \tag{4.4}$$

At the solution

$$\|\gamma D - Q\|_{\mathrm{F}} = \sqrt{n - \frac{(\sum \tau_i)^2}{\sum \tau_i^2}} \equiv \nu(D). \tag{4.5}$$

Proof. Let $Y \operatorname{diag}(\tau_1, \ldots, \tau_n) Z^*$ be the singular value decomposition of D. Then the unitary factor of the polar decomposition of D is YZ^*, which we have noted above is Q. It follows that

$$\|\gamma D - Q\|_{\mathrm{F}}^2 = \sum_i (\gamma \tau_i - 1)^2.$$

Minimizing this sum with respect to γ gives (4.4) and (4.5). □

This theorem takes care of the term $\|\gamma D - Q\|_{\mathrm{F}}^2 = \nu(D)$ in the right-hand side of (4.2). To handle the second term, $\|(\gamma D)^{-1} - Q^*\|_{\mathrm{F}}^2$, we will assume that $\nu(D)$ is less than 1. (This is not a very strong assumption, since we are bounding sines of angles.) It then follows from the perturbation theory of matrix inverses (see, e.g., [5, Chapter III]) that

$$\|(\gamma D)^{-1} - Q^*\|_{\mathrm{F}} \leqslant \frac{\|\gamma D - Q\|_{\mathrm{F}}}{1 - \|\gamma D - Q\|_2} \leqslant \frac{\nu(D)}{1 - \nu(D)}, \tag{4.6}$$

where $\|\cdot\|_2$ denotes the spectral matrix norm—the largest singular value. This gives us our final bound.

Corollary 4.2. *In Theorem* 3.2, *assume that* $\nu(D_{\mathrm{L}}), \nu(D_{\mathrm{R}}) < 1$, *where* $\nu(D)$ *is defined by* (4.5). *Then we may replace* $\epsilon_{\alpha,\beta}$ *by*

$$\epsilon = \sqrt{\nu(D_{\mathrm{L}})^2\left[1+\frac{1}{[1-\nu(D_{\mathrm{L}})]^2}\right]+\nu(D_{\mathrm{R}})^2\left[1+\frac{1}{[1-\nu(D_{\mathrm{R}})]^2}\right]}, \tag{4.7}$$

and α *and* β *in* $\rho_{\alpha\beta}$ *by the values determined as in Theorem* 4.1.

In (4.6), the norm $\|\gamma D - Q\|_2$ is majorized by $\nu(D)$, and this leads to the above corollary. We can obtain another bound by observing that the γ^{-1} is the weighted average of the τ_i. Hence $\min \tau_i \leqslant \gamma^{-1} \leqslant \max \tau_i$, and $\|\gamma D - Q\|_2 \leqslant \kappa(D) - 1$, where $\kappa(D) = \max \tau_i / \min \tau_i$. Thus in place of (4.7) we can use

$$\epsilon = \sqrt{\nu(D_{\mathrm{L}})^2\left[1+\frac{1}{[2-\kappa(D_{\mathrm{L}})]^2}\right]+\nu(D_{\mathrm{R}})^2\left[1+\frac{1}{[2-\kappa(D_{\mathrm{R}})]^2}\right]},$$

provided $\kappa(D_{\mathrm{L}}), \kappa(D_{\mathrm{R}}) < 2$.

Corollary 4.2 can be improved when $k = 1$. Specifically, we can replace the definition of $\epsilon_{\alpha,\beta}$ in Theorem 3.2 by

$$\begin{aligned}\epsilon_{\alpha,\beta}^2 = &\|\alpha D_{\mathrm{L}} - Q_{\mathrm{L}}\|_2^2 + \|(\alpha D_{\mathrm{L}})^{-1} - Q_{\mathrm{L}}^*\|_2^2\\ &+\|\beta D_{\mathrm{R}} - Q_{\mathrm{R}}\|_2^2 + \|(\beta D_{\mathrm{R}})^{-1} - Q_{\mathrm{R}}^*\|_2^2.\end{aligned} \tag{4.8}$$

Once again we face a minimization problem over the two parameters α and β. As above, a complete solution also appears to be intractable, and instead we consider the following problem:

$$\begin{array}{ll}\text{minimize} & \|\gamma D - Q\|_2\\ \text{subject to} & Q \text{ unitary and } \gamma \geqslant 0.\end{array} \tag{4.9}$$

Theorem 4.3. *Let* $\tau_1 \leqslant \tau_2 \leqslant \cdots \leqslant \tau_n$ *be the singular values of D. The matrix Q in the solution of* (4.9) *is the unitary factor of the polar decomposition of D. The value of* γ *is given by*

$$\gamma = \frac{2}{\tau_1 + \tau_n}. \tag{4.10}$$

At the solution

$$\|\gamma D - Q\|_2 = \frac{\kappa(D) - 1}{\kappa(D) + 1}, \tag{4.11}$$

where $\kappa(D) = \tau_n/\tau_1$.

Proof. Again Q is the unitary factor of the polar decomposition of D, and then

$$\|\gamma D - Q\|_2 = \max_i |\gamma\tau_i - 1| = \max\{|\gamma\tau_1 - 1|,\ |\gamma\tau_n - 1|\}.$$

Minimizing this with respect to γ gives (4.10) and (4.11). □

We now have an improved bound for the case $k = 1$.

Corollary 4.4. *In Theorem* 3.2, *if* $k = 1$, *we may replace* $\epsilon_{\alpha,\beta}$ *by*

$$\epsilon = \sqrt{(\kappa(D_\mathrm{L}) - 1)^2 \left[\frac{1}{(\kappa(D_\mathrm{L}) + 1)^2} + \frac{1}{4}\right] + (\kappa(D_\mathrm{R}) - 1)^2 \left[\frac{1}{(\kappa(D_\mathrm{R}) + 1)^2} + \frac{1}{4}\right]}, \tag{4.12}$$

and α *and* β *in* $\rho_{\alpha\beta}$ *by the values determined as in Theorem* 4.3.

5. An example and summary

Since the bounds are rather complicated, we illustrate them with a simple example, which was generated by the following matlab code:

```
n = 10;
err = 1e-4;
A = randn(n);
DR = 3*(eye(n) + err*randn(n));
for\-mat short e
[Q,R] = qr(randn(n));
DL = .5*(Q + err*randn(n));
AT = DL*A*DR;
```

Thus the original matrix A is a random matrix of order 10, the matrix D_R is three times a perturbation of order 10^{-4} of the identity matrix, and D_L is 0.5 times a perturbation of order 10^{-4} of a random unitary matrix. We bound the perturbation of the dominant right singular subspace of order 2.

The optimal values of α and β, determined from Theorem 4.1, are

$$\alpha_{\mathrm{opt}} = 1.9999064 \quad \text{and} \quad \beta_{\mathrm{opt}} = 0.3333337.$$

Note that the reciprocals of these values reproduce the values used to generate the example to about four places. Using these values, we get

$$\epsilon_{\alpha,\beta} = 1.4 \times 10^{-3},$$

and

$$\rho_{\alpha,\beta} = 0.14494. \tag{5.1}$$

The final bound is on the sines of the angles between the original and the perturbed subspace is 1.0×10^{-2}, whereas the true value is 1.1×10^{-3}. Thus the bound gives away about an order of magnitude.

In Section 3, we argued that optimizing $\epsilon_{\alpha,\beta}$ would produce a value of $\rho_{\alpha,\beta}$ that approximates the value that would be calculated from the original matrix (see (4.1)). In fact that value is

$$\rho = 0.14585,$$

which agrees very well with the value in (5.1).

In deciding what to optimize, we rejected the left-hand side of (4.2) as intractable. However, having determined values of α and β that optimize the right-hand side, we can use them in the left-hand side to get a sharper bound. In our example, this procedure causes the bound to decrease from 1.0×10^{-2} to 5.9×10^{-3}.

In conclusion, the bounds we have developed above separate a multiplicative perturbation D into two components. One is the distance of a scalar multiple of D to the nearest unitary matrix Q and the other is the distance of Q to the identity. By making this decomposition, we were able to take the bound in Theorem 2.2 and make it say more about the effects of multiplicative perturbations on the right and left singular subspaces. For simplicity, we have confined our exposition to this bound and to the Frobenius norm. But our technique applies to other, independent bounds in arbitrary unitarily invariant norms, such as those in [4]. This technique of decomposing multiplicative perturbation D into two components also extends to relative eigenspace variations as well.

References

[1] R. Bhatia, Matrix analysis, Graduate Texts in Mathematics, vol. 169, Springer, New York, 1996.

[2] S.C. Eisenstat, I.C.F. Ipsen, Relative perturbation bounds for eigenspaces and singular vector subspaces, in: J.G. Lewis (Ed.), Proceedings of the Fifth SIAM Conference on Applied Linear Algebra, SIAM, Philadelphia, PA, 1994, pp. 62–66.

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

[4] R.-C. Li, Relative perturbation theory: II. eigenspace and singular subspace variations, SIAM J. Matrix Anal. Appl. 20 (1999) 471–492.

[5] G.W. Stewart, J.-G. Sun, Matrix Perturbation Theory, Academic Press, Boston, 1990.

[6] P.-Å. Wedin, Perturbation bounds in connection with singular value decomposition, BIT 12 (1972) 99–111.