Relative Perturbation Bounds for the Unitary Polar Factor

Ren-Cang Li *

July 25, 1994

Abstract

Let B be an $m \times n$ $(m \geq n)$ complex matrix. It is known that there is a unique *polar decomposition* $B = QH$, where $Q^*Q = I$, the $n \times n$ identity matrix, and H is positive definite, provided B has full column rank. This paper addresses the following question: how much may Q change if B is perturbed to $\widetilde{B} = D_1^* B D_2$? Here D_1 and D_2 are two nonsingular matrices and close to the identities of suitable dimensions.

Known perturbation bounds for complex matrices indicate that in the worst case, the change in Q is proportional to the reciprocal of the smallest singular value of B. In this paper, we will prove that for the above mentioned perturbations to B, the change in Q is bounded only by the distances from D_1 and D_2 to identities!

As an application, we will consider perturbations for one-side scaling, i.e., the case when $G = D^* B$ is perturbed to $\widetilde{G} = D^* \widetilde{B}$, where D is usually a nonsingular diagonal scaling matrix but for our purpose we do not have to assume this, and B and $\widetilde{B}$ are nonsingular.

Let B be an $m \times n$ $(m \geq n)$ complex matrix. It is known that there are Q with orthonormal column vectors, i.e., $Q^*Q = I$, and a unique positive semidefinite H such that

$$B = QH. \tag{1}$$

Hereafter I denotes an identity matrix with appropriate dimensions which should be clear from the context or specified. The decomposition (1) is called the *polar decomposition* of B. If, in addition, B has full column rank then Q is uniquely determined also. In fact,

$$H = (B^*B)^{1/2}, \quad Q = B(B^*B)^{-1/2}, \tag{2}$$

*Department of Mathematics, University of California at Berkeley, Berkeley, California 94720, (`li@math.berkeley.edu`).

where superscript "*" denotes conjugate transpose. The decomposition (1) can also be computed from the *singular value decomposition* (SVD) $B = U\Sigma V^*$ by

$$H = V\Sigma_1 V^*, \quad Q = U_1 V^*, \tag{3}$$

where $U = (U_1, U_2)$ and V are unitary, U_1 is $m \times n$, $\Sigma = \begin{pmatrix} \Sigma_1 \\ 0 \end{pmatrix}$ and $\Sigma_1 = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$ is nonnegative.

There are many published bounds upon how much the two factor matrices Q and H may change if entries of B are perturbed in arbitrary manner [1, 2, 3, 4, 6, 5, 7, 8, 9]. In these papers, no assumption was made on how B was perturbed unlike what we are going to do here.

In this paper, we obtain some bounds for the perturbations of Q, assuming B is complex and is perturbed to $\widetilde{B} = D_1^* B D_2$, where D_1 and D_2 are two nonsingular matrices and close to the identities of suitable dimensions. Assume also B has full column rank and so do $\widetilde{B} = D_1^* B D_2$. Let

$$B = QH, \quad \widetilde{B} = \widetilde{Q}\widetilde{H} \tag{4}$$

be the polar decompositions of B and $\widetilde{B}$ respectively, and let

$$B = U\Sigma V^*, \quad \widetilde{B} = \widetilde{U}\widetilde{\Sigma}\widetilde{V}^* \tag{5}$$

be the SVDs of B and $\widetilde{B}$, respectively, where $\widetilde{U} = (\widetilde{U}_1, \widetilde{U}_2)$, $\widetilde{U}_1$ is $m \times n$, and $\widetilde{\Sigma} = \begin{pmatrix} \widetilde{\Sigma}_1 \\ 0 \end{pmatrix}$ and $\widetilde{\Sigma}_1 = \mathrm{diag}(\widetilde{\sigma}_1, \ldots, \widetilde{\sigma}_n)$. Assume as usual that

$$\sigma_1 \geq \cdots \geq \sigma_n > 0, \quad \text{and} \quad \widetilde{\sigma}_1 \geq \cdots \geq \widetilde{\sigma}_n > 0. \tag{6}$$

It follows from (2) and (5) that

$$Q = U_1 V^*, \quad \widetilde{Q} = \widetilde{U}_1 \widetilde{V}^*.$$

In what follows, $\|X\|_{\mathrm{F}}$ denotes the Frobenius norm which is the square root of the trace of X^*X. Then

$$\begin{aligned}
\widetilde{U}^*(\widetilde{B} - B)V &= \widetilde{\Sigma}\widetilde{V}^*V - \widetilde{U}^*U\Sigma, \\
\widetilde{U}^*(\widetilde{B} - B)V &= \widetilde{U}^*(D_1^*BD_2 - D_1^*B + D_1^*B - B)V \\
&= \widetilde{U}^*\left[\widetilde{B}(I - D_2^{-1}) + (D_1^* - I)B\right]V \\
&= \widetilde{\Sigma}\widetilde{V}^*(I - D_2^{-1})V + \widetilde{U}^*(D_1^* - I)U\Sigma,
\end{aligned}$$

and similarly

$$\begin{aligned}
U^*(\widetilde{B} - B)\widetilde{V} &= U^*\widetilde{U}\widetilde{\Sigma} - \Sigma V^*\widetilde{V}, \\
U^*(\widetilde{B} - B)\widetilde{V} &= U^*(D_1^*BD_2 - BD_2 + BD_2 - B)\widetilde{V} \\
&= U^*\left[(I - D_1^{-*})\widetilde{B} + B(D_2 - I)\right]\widetilde{V} \\
&= U^*(I - D_1^{-*})\widetilde{U}\widetilde{\Sigma} + \Sigma V^*(D_2 - I)\widetilde{V}.
\end{aligned}$$

Therefore, we obtained two perturbation equations.

$$\widetilde{\Sigma}\widetilde{V}^*V - \widetilde{U}^*U\Sigma = \widetilde{\Sigma}\widetilde{V}^*(I - D_2^{-1})V + \widetilde{U}^*(D_1^* - I)U\Sigma, \tag{7}$$
$$U^*\widetilde{U}\widetilde{\Sigma} - \Sigma V^*\widetilde{V} = U^*(I - D_1^{-*})\widetilde{U}\widetilde{\Sigma} + \Sigma V^*(D_2 - I)\widetilde{V}. \tag{8}$$

The first n rows of the equation (7) yields

$$\widetilde{\Sigma}_1\widetilde{V}^*V - \widetilde{U}_1^*U_1\Sigma_1 = \widetilde{\Sigma}_1\widetilde{V}^*(I - D_2^{-1})V + \widetilde{U}_1^*(D_1^* - I)U_1\Sigma_1. \tag{9}$$

The first n rows of the equation (8) yields

$$U_1^*\widetilde{U}_1\widetilde{\Sigma}_1 - \Sigma_1V^*\widetilde{V} = U_1^*(I - D_1^{-*})\widetilde{U}_1\widetilde{\Sigma}_1 + \Sigma_1V^*(D_2 - I)\widetilde{V},$$

on taking conjugate transpose of which, one has

$$\widetilde{\Sigma}_1\widetilde{U}_1^*U_1 - \widetilde{V}^*V\Sigma_1 = \widetilde{\Sigma}_1\widetilde{U}_1^*(I - D_1^{-1})U_1 + \widetilde{V}^*(D_2^* - I)V\Sigma_1. \tag{10}$$

Now subtracting (10) from (9) leads to

$$\begin{aligned} &\widetilde{\Sigma}_1(\widetilde{U}_1^*U_1 - \widetilde{V}^*V) + (\widetilde{U}_1^*U_1 - \widetilde{V}^*V)\Sigma_1 \qquad (11)\\ &= \widetilde{\Sigma}_1\left[\widetilde{U}_1^*(I - D_1^{-1})U_1 - \widetilde{V}^*(I - D_2^{-1})V\right] \\ &\quad + \left[\widetilde{V}^*(D_2^* - I)V - \widetilde{U}_1^*(D_1^* - I)U_1\right]\Sigma_1. \end{aligned}$$

Set

$$X = \widetilde{U}_1^*U_1 - \widetilde{V}^*V = (x_{ij}), \tag{12}$$
$$E = \widetilde{U}_1^*(I - D_1^{-1})U_1 - \widetilde{V}^*(I - D_2^{-1})V = (e_{ij}), \tag{13}$$
$$\widetilde{E} = \widetilde{V}^*(D_2^* - I)V - \widetilde{U}_1^*(D_1^* - I)U_1 = (\widetilde{e}_{ij}). \tag{14}$$

Then the equation (11) reads $\widetilde{\Sigma}_1X + X\Sigma_1 = \widetilde{\Sigma}_1E + \widetilde{E}\Sigma_1$, or componentwisely, $\widetilde{\sigma}_ix_{ij} + x_{ij}\sigma_j = \widetilde{\sigma}_ie_{ij} + \widetilde{e}_{ij}\sigma_j$. Thus

$$\begin{aligned} |(\widetilde{\sigma}_i + \sigma_j)x_{ij}| &\leq \sqrt{\widetilde{\sigma}_i^2 + \sigma_j^2}\sqrt{|e_{ij}|^2 + |\widetilde{e}_{ij}|^2} \\ &\Rightarrow \quad |x_{ij}|^2 \leq \frac{\widetilde{\sigma}_i^2 + \sigma_j^2}{(\widetilde{\sigma}_i + \sigma_j)^2}(|e_{ij}|^2 + |\widetilde{e}_{ij}|^2) \leq |e_{ij}|^2 + |\widetilde{e}_{ij}|^2. \end{aligned}$$

Summing on i and j for $i,\, j = 1, 2, \cdots, n$ produces

$$\|X\|_{\mathrm{F}}^2 = \sum_{i,\,j=1}^{n} |x_{ij}|^2 \leq \|E\|_{\mathrm{F}}^2 + \|\widetilde{E}\|_{\mathrm{F}}^2. \tag{15}$$

Notice that

$$\begin{aligned} X &= \widetilde{U}_1^*U_1 - \widetilde{V}^*V = \widetilde{V}^*(\widetilde{V}\widetilde{U}_1^*U_1V^* - I)V = \widetilde{V}^*(\widetilde{Q}^*Q - I)V, \\ &\Rightarrow \quad \|X\|_{\mathrm{F}} = \|\widetilde{Q}^*Q - I\|_{\mathrm{F}}, \end{aligned}$$

and

$$\begin{aligned} \|E\|_{\mathrm{F}} &\le \|I - D_1^{-1}\|_{\mathrm{F}} + \|I - D_2^{-1}\|_{\mathrm{F}}, \\ \|\widetilde{E}\|_{\mathrm{F}} &\le \|D_2^* - I\|_{\mathrm{F}} + \|D_1^* - I\|_{\mathrm{F}}. \end{aligned}$$

Lemma 1

$$\begin{aligned} &\|\widetilde{Q}^* Q - I\|_{\mathrm{F}} \\ &\quad \le \sqrt{\left(\|I - D_1^{-1}\|_{\mathrm{F}} + \|I - D_2^{-1}\|_{\mathrm{F}}\right)^2 + (\|D_2^* - I\|_{\mathrm{F}} + \|D_1^* - I\|_{\mathrm{F}})^2}. \end{aligned}$$

When $m = n$, both Q and $\widetilde{Q}$ are unitary. Thus $\|\widetilde{Q}^* Q - I\|_{\mathrm{F}} = \|Q - \widetilde{Q}\|_{\mathrm{F}}$, and Lemma 1 yields

Theorem 1 *Let B and $\widetilde{B} = D_1^* B D_2$ be two $n \times n$ nonsingular complex matrices whose polar decompositions are given by (4). Then*

$$\begin{aligned} \|Q - \widetilde{Q}\|_{\mathrm{F}} &\le \sqrt{\left(\|I - D_1^{-1}\|_{\mathrm{F}} + \|I - D_2^{-1}\|_{\mathrm{F}}\right)^2 + \left(\|D_2 - I\|_{\mathrm{F}} + \|D_1 - I\|_{\mathrm{F}}\right)^2} \quad (16) \\ &\le \sqrt{2}\sqrt{\|I - D_1^{-1}\|_{\mathrm{F}}^2 + \|I - D_2^{-1}\|_{\mathrm{F}}^2 + \|D_2 - I\|_{\mathrm{F}}^2 + \|D_1 - I\|_{\mathrm{F}}^2}. \end{aligned}$$

If, however, $m > n$, then it follows from the last $m - n$ rows of the equations (7) and (8) that

$$\begin{aligned} \widetilde{U}_2^* U_1 \Sigma_1 &= \widetilde{U}_2^* (D_1^* - I) U_1 \Sigma_1 \qquad \text{and} \\ U_2^* \widetilde{U}_1 \widetilde{\Sigma} &= U_2^* (I - D_1^{-*}) \widetilde{U}_1 \widetilde{\Sigma}_1. \end{aligned}$$

Since we assume that both B and $\widetilde{B}$ have full column rank, both Σ_1 and $\widetilde{\Sigma}_1$ are nonsingular diagonal matrices. So

$$\widetilde{U}_2^* U_1 = \widetilde{U}_2^* (D_1^* - I) U_1 \quad \text{and} \quad U_2^* \widetilde{U}_1 = U_2^* (I - D_1^{-*}) \widetilde{U}_1.$$

Therefore, we have

$$\|\widetilde{U}_2^* U_1\|_{\mathrm{F}} \le \|D_1^* - I\|_{\mathrm{F}} \quad \text{and} \quad \|U_2^* \widetilde{U}_1\|_{\mathrm{F}} = \|I - D_1^{-*}\|_{\mathrm{F}}. \tag{17}$$

Notice that $(U_1 V^*, U_2) = (Q, U_2)$ and $(\widetilde{U}_1 \widetilde{V}^*, \widetilde{U}_2) = (\widetilde{Q}, \widetilde{U}_2)$ are unitary. Hence $U_2^* Q = 0 = \widetilde{U}_2^* \widetilde{Q}$ and

$$\begin{aligned} &\|Q - \widetilde{Q}\|_{\mathrm{F}} = \|(Q, U_2)^* (Q - \widetilde{Q})\|_{\mathrm{F}} = \left\| \begin{pmatrix} I - Q^* \widetilde{Q} \\ -U_2^* \widetilde{Q} \end{pmatrix} \right\|_{\mathrm{F}} \\ &\quad \le \sqrt{\|I - Q^* \widetilde{Q}\|_{\mathrm{F}}^2 + \|-U_2^* \widetilde{U}_1 \widetilde{V}^*\|_{\mathrm{F}}^2} \\ &\quad \le \sqrt{\|I - Q^* \widetilde{Q}\|_{\mathrm{F}}^2 + \|U_2^* \widetilde{U}_1\|_{\mathrm{F}}^2} \\ &\quad \le \sqrt{\left(\|I - D_1^{-1}\|_{\mathrm{F}} + \|I - D_2^{-1}\|_{\mathrm{F}}\right)^2 + (\|D_2^* - I\|_{\mathrm{F}} + \|D_1^* - I\|_{\mathrm{F}})^2 + \|I - D_1^{-*}\|_{\mathrm{F}}^2}. \end{aligned} \tag{18}$$

Similarly, we have

$$\|Q-\widetilde{Q}\|_{\mathrm{F}} = \|(\widetilde{Q},\widetilde{U}_2)^*(Q-\widetilde{Q})\|_{\mathrm{F}} = \left\|\begin{pmatrix} \widetilde{Q}^*Q - I \\ \widetilde{U}_2 Q \end{pmatrix}\right\|_{\mathrm{F}}$$

$$\leq \sqrt{\left(\|I-D_1^{-1}\|_{\mathrm{F}} + \|I-D_2^{-1}\|_{\mathrm{F}}\right)^2 + (\|D_2^*-I\|_{\mathrm{F}} + \|D_1^*-I\|_{\mathrm{F}})^2 + \|D_1^*-I\|_{\mathrm{F}}^2}. \tag{19}$$

Theorem 2 below follows from (18) and (19).

Theorem 2 *Let A and $\widetilde{A}$ be two $m \times n$ $(m > n)$ complex matrices having full column rank and with the polar decompositions (4). Then*

$$\|Q-\widetilde{Q}\|_{\mathrm{F}} \leq \Big[\left(\|I-D_1^{-1}\|_{\mathrm{F}} + \|I-D_2^{-1}\|_{\mathrm{F}}\right)^2$$

$$+ (\|I-D_2\|_{\mathrm{F}} + \|I-D_1\|_{\mathrm{F}})^2 + \min\left\{\|I-D_1^{-1}\|_{\mathrm{F}}^2, \|I-D_1\|_{\mathrm{F}}^2\right\}\Big]^{\frac{1}{2}}$$

$$\leq \sqrt{3}\sqrt{\|I-D_2\|_{\mathrm{F}}^2 + \|I-D_2^{-1}\|_{\mathrm{F}}^2 + \|I-D_1\|_{\mathrm{F}}^2 + \|I-D_1^{-1}\|_{\mathrm{F}}^2}.$$

Now we are in the position to apply Theorem 1 to perturbations for one-side scaling (from the left). Here we consider two $n \times n$ nonsingular matrices $G = D^*B$ and $\widetilde{G} = D^*\widetilde{B}$, where D is a scaling matrix and usually diagonal (but this is not necessary to the theorem that follows). B is nonsingular and usually better conditioned than G itself. Set

$$\Delta B \stackrel{\mathrm{def}}{=} \widetilde{B} - B.$$

$\widetilde{B}$ is also nonsingular by the condition $\|\Delta B\|_2\|B^{-1}\|_2 < 1$ which will be assumed henceforth. Notice that

$$\widetilde{G} = D^*\widetilde{B} = D^*(B+\Delta B) = D^*B(I+B^{-1}(\Delta B)) = G(I+B^{-1}(\Delta B)).$$

So applying Theorem 1 with $D_1 = 0$ and $D_2 = I + B^{-1}(\Delta B)$ leads to

Theorem 3 *Let $G = D^*B$ and $\widetilde{G} = D^*\widetilde{B}$ be two $n \times n$ nonsingular matrices, and let*

$$G = QH \quad \textit{and} \quad \widetilde{G} = \widetilde{Q}\widetilde{H}$$

be their polar decompositions. Set $\Delta B \stackrel{\mathrm{def}}{=} \widetilde{B} - B$. If $\|\Delta B\|_2\|B^{-1}\|_2 < 1$ then

$$\|Q-\widetilde{Q}\|_{\mathrm{F}} \leq \sqrt{\|B^{-1}(\Delta B)\|_{\mathrm{F}}^2 + \left\|I - \left(I + B^{-1}(\Delta B)\right)^{-1}\right\|_{\mathrm{F}}^2}$$

$$\leq \sqrt{1 + \frac{1}{(1-\|B^{-1}\|_2\|\Delta B\|_2)^2}}\,\|B^{-1}\|_2\|\Delta B\|_{\mathrm{F}}.$$

One can deal with one-side scaling from the right in the similar way.

Acknowledgement: I thank Professor W. Kahan for his supervision and Professor J. Demmel for valuable discussions.

References

[1] A. Barrlund. Perturbation bounds on the polar decomposition. *BIT*, 30:101–113, 1990.

[2] C.-H. Chen and J.-G. Sun. Perturbation bounds for the polar factors. *J. Comp. Math.*, 7:397–401, 1989.

[3] N. J. Higham. Computing the polar decomposition–with applications. *SIAM Journal on Scientific and Statistical Computing*, 7:1160–1174, 1986.

[4] C. Kenney and A. J. Laub. Polar decompostion and matrix sign function condition estimates. *SIAM Journal on Scientific and Statistical Computing*, 12:488–504, 1991.

[5] R.-C. Li. New perturbation bounds for the unitary polar factor. Manuscript submitted to SIAM J. Matrix Anal. Appl., 1993.

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

[7] J.-Q. Mao. The perturbation analysis of the product of singular vector matrices uv^h. *J. Comp. Math.*, 4:245–248, 1986.

[8] R. Mathias. Perturbation bounds for the polar decomposition. *SIAM J. Matrix Anal. Appl.*, 14:588–597, 1993.

[9] J.-G. Sun and C.-H. Chen. Generalized polar decomposition. *Math. Numer. Sinica*, 11:262–273, 1989. In Chinese.