NUMERICAL ALGEBRA,
CONTROL AND OPTIMIZATION
Volume 1, Number 2, June 2011

doi:10.3934/naco.2011.1.301

pp. 301–316

MULTIPLICATIVE PERTURBATION ANALYSIS FOR QR FACTORIZATIONS

Xiao-Wen Chang

School of Computer Science, McGill University
Montreal, Quebec, Canada H3A 2A7

Ren-Cang Li

Department of Mathematics, University of Texas at Arlington
Arlington, TX 76019-0408, USA

(Communicated by Wenyu Sun)

Abstract. This paper is concerned with how the QR factors change when a real matrix A suffers from a left or right multiplicative perturbation, where A is assumed to have full column rank. It is proved that for a left multiplicative perturbation the relative changes in the QR factors in norm are no bigger than a small constant multiple of the norm of the difference between the perturbation and the identity matrix. One of common cases for a left multiplicative perturbation case naturally arises from computing the QR factorization of A. The newly established bounds can be used to explain the accuracy in the computed QR factors. For a right multiplicative perturbation, the bounds on the relative changes in the QR factors are still dependent upon the condition number of the scaled R-factor, however. Some "optimized" bounds are also obtained by taking into account certain invariant properties in the factors.

1. **Introduction.**

Given a matrix $A \in \mathbb{R}^{m\times n}$ with full column rank, there exists a unique QR factorization

$$A = QR,$$

where $Q \in \mathbb{R}^{m\times n}$ has orthonormal columns and $R \in \mathbb{R}^{n\times n}$ is upper triangular with positive diagonal elements. The QR factorization is a very important computational tool in numerical linear algebra, e.g., it is used to solve least squares problems [4], compute singular value and eigenvalue decompositions [10]. If A is perturbed, we would like to know how its QR factors Q and R are perturbed. There are extensive studies in this regard, e.g., [1, 3, 5, 6, 7, 8, 9, 17, 18, 20, 21, 22, 23], for the so-called *additive perturbations*, namely A is perturbed to $\widetilde{A} \equiv A + \Delta A$ with an assumption on the smallness of ΔA usually in norm. In this paper we consider the case when A is multiplicatively perturbed, namely A is perturbed to $\widetilde{A} \equiv D_{\rm L}A$ or $AD_{\rm R}$ with both $D_{\rm L}$ and $D_{\rm R}$ near the identities, multiples of the identities, or sometimes orthogonal matrices. $D_{\rm L}$ and $D_{\rm R}$ are, respectively, called the *left* and *right multiplicative perturbations*.

2000 *Mathematics Subject Classification.* 15A23; 65F35.

Key words and phrases. QR factorization, perturbation analysis, multiplicative perturbation.

The first author was supported in part by NSERC of Canada grant RGPIN217191-07, and the second author was supported in part by the NSF grant DMS-0810506.

Multiplicative perturbations can easily be turned into additive perturbations: $\widetilde{A} = D_{\rm L} A D_{\rm R} = A + \Delta A$ with

$$\Delta A = D_{\rm L} A D_{\rm R} - A = (D_{\rm L} - I)AD_{\rm R} + A(D_{\rm R} - I). \tag{1.1}$$

Any bounds on $D_{\rm L} - I$ and $D_{\rm R} - I$ lead to a bound on ΔA, usually in norm. Then the existing additive perturbation analysis can be applied directly to give perturbation bounds. But in general this approach may produce, not surprisingly, unnecessarily conservative perturbation bounds because it ignores the nature of the perturbations. For more realistic bounds, we may have to perform new analysis that takes advantage of any structures in the perturbations. This point of view is not new. In the past, various multiplicative perturbation analyses have been done to certain problems in numerical linear algebra, including the polar decomposition [12, 15], and the eigendecomposition of a Hermitian matrix, and the singular value decomposition [13, 14, 16].

Multiplicative perturbations naturally arise from matrix scaling, a commonly used technique to improve the conditioning of a matrix. For example the matrix A itself may be very ill-conditioned, but there exists a scaling matrix S such that $B = AS^{-1}$ is much better conditioned. This is the so-called *right scaling*. Often S is diagonal but that is not necessary. Subsequently if $A = BS$ is perturbed to $\widetilde{A}$ scaled also by S into $\widetilde{A} = (B + \Delta B)S$, where ΔB is tiny relative to B, and if[1] the row space of ΔB is contained in that of B, then

$$(B + \Delta B)S = [I + (\Delta B)B^{\dagger}]BS \equiv D_{\rm L} A \tag{1.2}$$

which is in the form of having a left multiplicative perturbation[2] $D_{\rm L} = I + (\Delta B)B^{\dagger}$ that is close to the identity matrix, where $B^{\dagger}$ is B's Moore-Penrose pseudo-inverse. Similar arguments can be made for A scaled from the left, too, to give an example of a right multiplicative perturbation.

In this paper, we will establish perturbation bounds for the QR factors when A suffers from a left or right multiplicative perturbation. The bounds indicate that both the Q- and R-factors are well-conditioned with respect to a left multiplicative perturbation. This is utterly different from the case where A is subject to a general additive perturbation: the Q- and R-factors can be very ill-conditioned with respect to the additive perturbation, see e.g., [7]. Also the two factors behave very differently with respect to a right multiplicative perturbation: the first-order perturbation bound for the Q-factor can be arbitrarily small while the bound for the R-factor is about the same as what we may obtain if we apply existing perturbation bounds for the R-factor with respect to an additive perturbation upon using the conversion (1.1).

The rest of this paper is organized as follows. Section 2 presents some preliminaries and existing additive perturbation bounds for the QR factors. Our main results are detailed in Section 3. We then give a couple of numerical examples to

[1] For the interest of this article, B is assumed to have full column rank. Then this assumption is automatically true. In fact $B^{\dagger} = (B^{\rm T}B)^{-1}B^{\rm T}$ and thus $B^{\dagger}B = I_n$. In general, this assumption implies that $\Delta B = MB$ for some $M \in \mathbb{R}^{m\times m}$. Then $(\Delta B)B^{\dagger}B = MBB^{\dagger}B = MB = \Delta B$ which gives the first equality in (1.2).

[2] Alternatively one may use $\widetilde{A} = A + \Delta A = \left[I + (\Delta A)A^{\dagger}\right]A$. But since $(\Delta A)A^{\dagger}$ is usually measured by its upper bound, e.g., $\|\Delta A\|_2\|A^{\dagger}\|_2 = \kappa_2(A)\frac{\|\Delta A\|_2}{\|A\|_2}$ that is much larger than $\|\Delta B\|_2\|B^{\dagger}\|_2 = \kappa_2(B)\frac{\|\Delta B\|_2}{\|B\|_2}$ when A is ill-conditioned, it is usually a good idea to introduce the scaling matrix S mentioned here.

illustrate our new multiplicative perturbation bounds in comparison to the additive perturbation bounds in Section 4. Finally Section 5 gives a few concluding remarks.

Notation. Throughout this paper, $\mathbb{R}^{m\times n}$ is the set of all $m\times n$ matrices with entries in $\mathbb{R}$ (the set of real numbers), $\mathbb{R}^n=\mathbb{R}^{n\times 1}$, and $\mathbb{R}=\mathbb{R}^1$. $\mathbb{D}_n\in\mathbb{R}^{n\times n}$ is the set of real $n\times n$ diagonal matrices with positive diagonal entries. I_n (or simply I if its dimension is clear from the context) is the $n\times n$ identity matrix. The superscript "$\cdot^{\mathrm{T}}$" takes transpose. $\|X\|_2$ and $\|X\|_{\mathrm{F}}$ are the spectral norm and Frobenius norm of X, respectively, and $\kappa_2(X)$ is X's spectral condition number defined as

$$\kappa_2(X)=\|X\|_2\,\|X^{\dagger}\|_2,$$

where $X^{\dagger}$ is X's Moore-Penrose pseudo-inverse. MATLAB-like notation $X_{(:,j)}$ refers to the jth column of X. Symbols A, $\widetilde{A}$, and these for their QR factors are reserved: $A\in\mathbb{R}^{m\times n}$ has full column rank and it is additively/multiplicatively perturbed to $\widetilde{A}\in\mathbb{R}^{m\times n}$. Sufficient conditions will be stated to make sure $\widetilde{A}$ has full column rank, too. Their unique QR factorizations are

$$A=QR,\quad \widetilde{A}=\widetilde{Q}\widetilde{R}, \tag{1.3}$$

where $Q,\widetilde{Q}\in\mathbb{R}^{m\times n}$ have orthonormal columns, i.e., $Q^{\mathrm{T}}Q=\widetilde{Q}^{\mathrm{T}}\widetilde{Q}=I_n$, and $R,\widetilde{R}\in\mathbb{R}^{n\times n}$ are upper triangular with positive diagonal entries.

2. **Preliminaries.**

We write

$$\varDelta A=\widetilde{A}-A,\quad \varDelta Q=\widetilde{Q}-Q,\quad \varDelta R=\widetilde{R}-R \tag{2.1}$$

for the corresponding additive perturbations. Let P_A be the orthogonal projector onto the column space of A, i.e., $P_A=QQ^{\mathrm{T}}$. For any $X\in\mathbb{R}^{n\times n}$, it can be verified that

$$\|P_AX\|_p=\|Q^{\mathrm{T}}X\|_p,\quad \|XP_A\|_p=\|XQ\|_p\quad\text{for}\quad p=2,\mathrm{F}.$$

The following quantities will be used in perturbation bounds:

$$\eta_p=\kappa_2(A)\frac{\|\varDelta A\|_p}{\|A\|_2}\quad\text{for}\quad p=2,\mathrm{F}. \tag{2.2}$$

For any $X=(x_{ij})\in\mathbb{R}^{n\times n}$, as in [6] we define

$$\mathrm{up}(X)=\begin{bmatrix}\frac12 x_{11} & x_{12} & \cdots & x_{1n}\\ & \frac12 x_{22} & \cdots & x_{2n}\\ & & \ddots & \vdots\\ & & & \frac12 x_{nn}\end{bmatrix}, \tag{2.3}$$

$$\mathrm{low}(X)=\begin{bmatrix}\frac12 x_{11} & & & \\ x_{21} & \frac12 x_{22} & & \\ \vdots & \vdots & \ddots & \\ x_{n1} & x_{n2} & \cdots & \frac12 x_{nn}\end{bmatrix}=\left[\mathrm{up}(X^{\mathrm{T}})\right]^{\mathrm{T}}. \tag{2.4}$$

Lemma 2.1. *For any* $D = \operatorname{diag}(\delta_1, \ldots, \delta_n) \in \mathbb{D}_n$, $X_{n-1} \in \mathbb{R}^{n\times(n-1)}$, *and* $X = [X_{n-1}, x_n] \in \mathbb{R}^{n\times n}$,

$$\|\mathrm{up}(X)\|_{\mathrm{F}} \le \|X\|_{\mathrm{F}}, \tag{2.5}$$

$$\|\mathrm{up}(X^{\mathrm{T}} + X)\|_{\mathrm{F}} \le \sqrt{2}\,\|X\|_{\mathrm{F}}, \tag{2.6}$$

$$\|\mathrm{up}(X)\|_{\mathrm{F}} \le \frac{1}{\sqrt{2}}\,\|X\|_{\mathrm{F}} \quad \textit{if } X^{\mathrm{T}} = X, \tag{2.7}$$

$$\|\mathrm{up}(X) + D^{-1}\mathrm{up}(X^{\mathrm{T}})D\|_{\mathrm{F}} \le \rho_D\|X\|_{\mathrm{F}}, \tag{2.8}$$

$$\|D\,\mathrm{low}(X)D^{-1} - D^{-1}[\mathrm{low}(X)]^{\mathrm{T}}D\|_{\mathrm{F}} \le \sqrt{2}\zeta_D\|X_{n-1}\|_{\mathrm{F}} \le \sqrt{2}\zeta_D\|X\|_{\mathrm{F}}, \tag{2.9}$$

where

$$\zeta_D = \max_{1\le i<j\le n} \delta_j/\delta_i, \quad \rho_D = \left(1+\zeta_D^2\right)^{1/2}. \tag{2.10}$$

Proof. Inequalities (2.5), (2.6) and (2.7) can be easily verified. Inequality (2.8) was proved in [7, Lemma 5.1] and (2.6) is a special case of (2.8). Finally

$$\|D\,\mathrm{low}(X)D^{-1} - D^{-1}[\mathrm{low}(X)]^{\mathrm{T}}D\|_{\mathrm{F}}^2 = 2\sum_{1\le i<j\le n} x_{ji}^2\delta_j^2/\delta_i^2 \le 2\zeta_D^2\|X_{n-1}\|_{\mathrm{F}}^2,$$

leading to (2.9). □

Lemma 2.2. *If* $\eta_2 < 1$, *then* $\widetilde{A}$ *has full column rank and its QR factorization satisfies*

$$\|\Delta Q\|_{\mathrm{F}} \le \frac{\sqrt{2}\,\eta_{\mathrm{F}}}{1-\eta_2}, \quad \frac{\|\Delta R\|_{\mathrm{F}}}{\|R\|_2} \le \frac{\sqrt{2}\,\eta_{\mathrm{F}}}{1-\eta_2}. \tag{2.11}$$

If further $\eta_{\mathrm{F}} < \sqrt{3/2} - 1$, *then*

$$\frac{\|\Delta R\|_{\mathrm{F}}}{\|R\|_2} \le \frac{\sqrt{2}\left(\inf_{D\in\mathbb{D}_n}\rho_D\kappa_2(D^{-1}R)\right)\left(\frac{\|Q^T\Delta A\|_{\mathrm{F}}}{\|A\|_2} + \kappa_2(A)\frac{\|\Delta A\|_{\mathrm{F}}^2}{\|A\|_2^2}\right)}{\sqrt{2}-1+\sqrt{1-4\kappa_2(A)\frac{\|\Delta A\|_{\mathrm{F}}}{\|A\|_2} - 2\kappa_2^2(A)\frac{\|\Delta A\|_{\mathrm{F}}^2}{\|A\|_2^2}}} \tag{2.12}$$

$$\le \left(\sqrt{6}+\sqrt{3}\right)\inf_{D\in\mathbb{D}_n}\rho_D\kappa_2(D^{-1}R)\,\frac{\|\Delta A\|_{\mathrm{F}}}{\|A\|_2}, \tag{2.13}$$

where ρ_D *is defined in* (2.10).

The two inequalities in (2.11) were presented in [22, Theorem 5.1] (there was a typo in the bound given there) and [20], respectively, and the inequalities in (2.12) and (2.13) were given in [8].

3. **Main results.**

We adopt all the notations and assumptions specified at the beginning of Section 2.

3.1. **Left multiplicative perturbation.**

We consider in this subsection the so-called *left multiplicative perturbation* to A:

$$\widetilde{A} = D_{\mathrm{L}}\,A, \tag{3.1}$$

where $D_{\mathrm{L}} \in \mathbb{R}^{m\times m}$ is near I_m, or a scalar multiple of I_m, or sometimes an orthogonal matrix.

Theorem 3.1. *Assume (3.1) and write $D_{\rm L} = I_m + E$. If $\|EP_{A}\|_2 < 1$, then $\widetilde{A}$ has full column rank and its unique QR factorization satisfies*

$$\|\Delta Q\|_{\rm F} \leq \frac{\sqrt{2}\|EP_{A}\|_{\rm F}}{1-\|EP_{A}\|_2}, \quad \frac{\|\Delta R\|_{\rm F}}{\|R\|_2} \leq \|(\Delta R)R^{-1}\|_{\rm F} \leq \frac{\sqrt{2}\|EP_{A}\|_{\rm F}}{1-\|EP_{A}\|_2}. \tag{3.2}$$

In particular, if $\|E\|_2 < 1$, they imply

$$\|\Delta Q\|_{\rm F} \leq \frac{\sqrt{2}\|E\|_{\rm F}}{1-\|E\|_2}, \quad \frac{\|\Delta R\|_{\rm F}}{\|R\|_2} \leq \|(\Delta R)R^{-1}\|_{\rm F} \leq \frac{\sqrt{2}\|E\|_{\rm F}}{1-\|E\|_2}. \tag{3.3}$$

Proof. Since $\|EP_{A}\|_p \leq \|E\|_p$ for $p = 2$, F, (3.3) is a result of (3.2) which we will prove now.

Note that the QR factorization of Q is $Q = Q \cdot I_n$ and $(I + E)Q = Q + EQ$ can be regarded as an additively perturbed Q by EQ. Since

$$\kappa_2(Q)\|EQ\|_2/\|Q\|_2 = \|EP_{A}\|_2 < 1,$$

by Lemma 2.2 $(I + E)Q$ has a unique QR factorization

$$(I+E)Q = \widetilde{Q}\widehat{R}, \tag{3.4}$$

satisfying

$$\|\widetilde{Q} - Q\|_{\rm F} \leq \frac{\sqrt{2}\kappa_2(Q)\frac{\|EQ\|_{\rm F}}{\|Q\|_2}}{1-\kappa_2(Q)\frac{\|EQ\|_2}{\|Q\|_2}} = \frac{\sqrt{2}\|EP_{A}\|_{\rm F}}{1-\|EP_{A}\|_2}, \tag{3.5}$$

$$\|\widehat{R} - I\|_{\rm F} \leq \frac{\sqrt{2}\kappa_2(Q)\frac{\|EQ\|_{\rm F}}{\|Q\|_2}}{1-\kappa_2(Q)\frac{\|EQ\|_2}{\|Q\|_2}} = \frac{\sqrt{2}\|EP_{A}\|_{\rm F}}{1-\|EP_{A}\|_2}. \tag{3.6}$$

Therefore

$$D_{\rm L}A = (I+E)QR = \widetilde{Q}\widehat{R}R = \widetilde{Q}\widetilde{R}, \tag{3.7}$$

which is the unique QR factorization of $\widetilde{A}$ because, by construction, $\widetilde{Q}$ has orthonormal columns and $\widetilde{R}$ is upper triangular with positive diagonal entries. At the same time, this implies that $\widetilde{A}$ has full column rank. Inequalities (3.5) and (3.6), together with $\|(\Delta R)R^{-1}\|_{\rm F} = \|(\widetilde{R} - R)R^{-1}\|_{\rm F} = \|\widehat{R} - I\|_{\rm F}$, give (3.2). ☐

Remark **3.1.** Theorem 3.1 indicates that both the Q-factor and R-factor are very well conditioned with respect to the left multiplicative perturbation in A.

Remark **3.2.** Inequalities in (3.3) are obviously less sharper than the ones in (3.2), but they are usually more convenient to use because in practice it is more likely that one knows bounds on the norms of E than bounds on the norms of EP_{A}.

Remark **3.3.** From (3.2), we obtain the following first-order bounds:

$$\|\Delta Q\|_{\rm F} \leq \sqrt{2}\|EP_{A}\|_{\rm F} + O(\|EP_{A}\|_{\rm F}^2), \quad \frac{\|\Delta R\|_{\rm F}}{\|R\|_2} \leq \sqrt{2}\|EP_{A}\|_{\rm F} + O(\|EP_{A}\|_{\rm F}^2).$$

The first-order bound for the R-factor can be improved. In fact, in the proof of Theorem 3.1, if we apply the inequality (2.12) with $D = I$, it can be shown that we have

$$\|\widehat{R} - I\|_{\rm F} \leq \sqrt{2}\|Q^TEQ\|_{\rm F} + O(\|EQ\|_{\rm F}^2).$$

Then it follows that

$$\frac{\|\Delta R\|_{\rm F}}{\|R\|_2} \leq \sqrt{2}\|P_{A}EP_{A}\|_{\rm F} + O(\|EP_{A}\|_{\rm F}^2).$$

Next we observe the following invariant properties of the QR factors:

1. multiplying A by any positive scalar does not change the Q-factor;
2. multiplying A from the left by any orthogonal matrix of apt size does not change the R-factor.

Following an idea in [16], we now show how the inequalities in (3.2) can be refined after taking these two observations into considerations.

For any scalar $\alpha > 0$, we have the QR factorization of $\alpha\, D_{\rm L} A$:

$$\alpha\, D_{\rm L} A = [I + (\alpha D_{\rm L} - I)]A = \widetilde{Q}(\alpha \widetilde{R}).$$

If $\|(\alpha D_{\rm L} - I)Q\|_2 < 1$, then applying the first inequality in (3.2) gives

$$\|\varDelta Q\|_{\rm F} \le \frac{\sqrt{2}\|(\alpha D_{\rm L} - I)P_{\scriptscriptstyle A}\|_{\rm F}}{1 - \|(\alpha D_{\rm L} - I)P_{\scriptscriptstyle A}\|_2}. \tag{3.8}$$

The above bound holds for any $\alpha > 0$ such that $\|(\alpha D_{\rm L} - I)Q\|_2 < 1$. In order to tighten this bound, we choose α such that $\|(\alpha D_{\rm L} - I)Q\|_{\rm F}$ is minimal. Notice that

$$\begin{aligned}\|(\alpha D_{\rm L} - I)Q\|_{\rm F}^2 &= \operatorname{trace}\left([\alpha\, D_{\rm L} Q - Q]^{\rm T}[\alpha\, D_{\rm L} Q - Q]\right)\\ &= \|D_{\rm L} Q\|_{\rm F}^2 \alpha^2 - 2\operatorname{trace}\left(Q^{\rm T} D_{\rm L} Q\right)\alpha + n.\end{aligned}$$

It is minimized at $\alpha = \hat{\alpha}$:

$$\hat{\alpha} = \frac{\operatorname{trace}(Q^{\rm T} D_{\rm L} Q)}{\|D_{\rm L} Q\|_{\rm F}^2} = \frac{\operatorname{trace}(D_{\rm L} P_{\scriptscriptstyle A})}{\|D_{\rm L} P_{\scriptscriptstyle A}\|_{\rm F}^2}. \tag{3.9}$$

It can be shown that $\hat{\alpha} > 0$ if $\|E P_{\scriptscriptstyle A}\|_2 < 1$. Similarly we can obtain

$$\tilde{\alpha} \equiv \underset{\alpha}{\operatorname{argmin}}\, \|\alpha D_{\rm L} - I\|_{\rm F}^2 = \frac{\operatorname{trace}(D_{\rm L})}{\|D_{\rm L}\|_{\rm F}^2}. \tag{3.10}$$

Now apply Theorem 3.1 to get

Theorem 3.2. *Assume* (3.1) *and write* $D_{\rm L} = I_m + E$. *Let* $\hat{\alpha}$ *and* $\tilde{\alpha}$ *be defined by* (3.9) *and* (3.10), *respectively. If* $\|E P_{\scriptscriptstyle A}\|_2 < 1$, *then* $\widetilde{A}$ *has full column rank and its unique QR factorization satisfies*

$$\|\varDelta Q\|_{\rm F} \le \frac{\sqrt{2}\|(\hat{\alpha} D_{\rm L} - I)P_A\|_{\rm F}}{1 - \|(\hat{\alpha} D_{\rm L} - I)P_A\|_2}. \tag{3.11}$$

If the stronger condition $\|E\|_2 < 1$ *holds, then* (3.11) *implies*

$$\|\varDelta Q\|_{\rm F} \le \frac{\sqrt{2}\|\tilde{\alpha} D_{\rm L} - I\|_{\rm F}}{1 - \|\tilde{\alpha} D_{\rm L} - I\|_2}. \tag{3.12}$$

REMARK **3.4.** If E is a scalar multiple of the identity matrix, then

$$\hat{\alpha} D_{\rm L} - I = \tilde{\alpha} D_{\rm L} - I = 0,$$

and therefore the upper bounds in (3.11) and (3.12) are 0. But the upper bounds in the first inequalities in (3.2) and (3.3) are not 0 and can be large.

For any orthogonal matrix $U \in \mathbb{R}^{m\times m}$,

$$U D_{\rm L} A = (U\widetilde{Q})\widetilde{R}.$$

If $\|(U D_{\rm L} - I)P_{\scriptscriptstyle A}\|_2 = \|(U D_{\rm L} - I)Q\|_2 < 1$, then (3.6) in the proof of Theorem 3.1 yields

$$\|(\varDelta R)R^{-1}\|_{\rm F} \le \frac{\sqrt{2}\|U D_{\rm L} Q - Q\|_{\rm F}}{1 - \|U D_{\rm L} Q - Q\|_2}. \tag{3.13}$$

We would like to minimize the right-hand side of this inequality over all orthogonal $U \in \mathbb{R}^{m\times m}$.

Lemma 3.1. *We have*

$$\min_{\text{orthogonal } U} \|UD_{\mathrm{L}}Q - Q\| = \|\mathrm{diag}(\mu_1 - 1, \dots, \mu_n - 1)\| \tag{3.14}$$

for all unitarily invariant norms $\|\cdot\|$, *where* $\mu_1, \dots, \mu_n$ *are the singular values of* $D_{\mathrm{L}}Q$.

Proof. By Mirsky's theorem [19, p.204], we have for any orthogonal matrix U

$$\|UD_{\mathrm{L}}Q - Q\| \geq \|\mathrm{diag}(\mu_1 - 1, \dots, \mu_n - 1)\|. \tag{3.15}$$

Let the singular value decomposition of $D_{\mathrm{L}}Q$ be $D_{\mathrm{L}}Q = V\Sigma W^T$, where $V \in \mathbb{R}^{m\times n}$ has orthonormal columns, $W \in \mathbb{R}^{n\times n}$ is orthogonal, and $\Sigma = \mathrm{diag}(\mu_1, \dots, \mu_n)$. Let $P = VW^T \in \mathbb{R}^{m\times n}$ and $H = W\Sigma W^T \in \mathbb{R}^{n\times n}$. Then $D_{\mathrm{L}}Q = PH$, which is the polar decomposition of $D_{\mathrm{L}}Q$. Suppose that $[P, P_\perp] \in \mathbb{R}^{m\times m}$ and $[Q, Q_\perp] \in \mathbb{R}^{m\times m}$ are orthogonal. Define $U = [Q, Q_\perp][P, P_\perp]^T$, which is orthogonal. Then $UP = Q$. Therefore, for this U,

$$\|UD_{\mathrm{L}}Q - Q\| = \|QH - Q\| \leq \|Q\|_2\|H - I_n\| = \|\mathrm{diag}(\mu_1 - 1, \dots, \mu_n - 1)\|. \tag{3.16}$$

Equality (3.14) is the consequence of (3.15) and (3.16). ☐

Theorem 3.3. *Assume* (3.1), *and let* $\mu_1, \dots, \mu_n$ *be the singular values of* $D_{\mathrm{L}}Q$ *and* $\nu_1, \dots, \nu_m$ *be the singular values of* D_{L}. *If* $\max_i |\mu_i - 1| < 1$, *then* $\widetilde{A}$ *has full column rank and its unique QR factorization satisfies*

$$\frac{\|\Delta R\|_{\mathrm{F}}}{\|R\|_2} \leq \|(\Delta R)R^{-1}\|_{\mathrm{F}} \leq \frac{\sqrt{2}\sqrt{\sum_{i=1}^{n} |\mu_i - 1|^2}}{1 - \max_{1\leq i\leq n} |\mu_i - 1|}. \tag{3.17}$$

If $\max_i |\nu_i - 1| < 1$, *then*

$$\frac{\|\Delta R\|_{\mathrm{F}}}{\|R\|_2} \leq \|(\Delta R)R^{-1}\|_{\mathrm{F}} \leq \frac{\sqrt{2}\sqrt{\sum_{i=1}^{m} |\nu_i - 1|^2}}{1 - \max_{1\leq i\leq m} |\nu_i - 1|}. \tag{3.18}$$

Proof. Inequality (3.17) is the result of (3.13) and Lemma 3.1. Inequality (3.18) can be proved similarly. ☐

These two inequalities in Theorem 3.3 are sharper than their corresponding ones in Theorem 3.1. But their applicability depends on the availability of information on $|\mu_i - 1|$ and $|\nu_i - 1|$. Often such information is hard to come by. Therefore, their most important value is perhaps the revelation of what in the left multiplication perturbation really moves the R-factor on the theoretical side rather than on the practical side of usefulness.

3.2. **Right multiplicative perturbation.** We consider the so-called *right multiplicative perturbation* to A:

$$\widetilde{A} = AD_{\mathrm{R}}, \tag{3.19}$$

where $D_{\mathrm{R}} \in \mathbb{R}^{n\times n}$ is near I_n, or a scalar multiple of I_n.

Theorem 3.4. *Assume* (3.19), *write* $D_{\mathrm{R}} = I_n + F$, *where* $F = [F_{n-1}, f_n]$ *and* $F_{n-1} \in \mathbb{R}^{n\times(n-1)}$, *and write* $R = \begin{bmatrix} R_{n-1} & r \\ 0 & r_{nn} \end{bmatrix}$, *where* $R_{n-1} \in \mathbb{R}^{(n-1)\times(n-1)}$. *If*

$$\kappa_2(A)\|F\|_{\mathrm{F}} < \sqrt{3/2} - 1, \tag{3.20}$$

then $\widetilde{A}$ has full column rank and its unique QR factorization satisfies

$$\begin{aligned}\|\Delta Q\|_{\mathrm{F}} &\le \frac{\sqrt{2}}{1-\sqrt{2}\,\kappa_2(R)\|F\|_{\mathrm{F}}}\Big[\big(\inf_{D\in\mathbb{D}_n}\zeta_D\|D^{-1}R\|_2\|R_{n-1}^{-1}D_{n-1}\|_2\big)\|F_{n-1}\|_{\mathrm{F}} \\ &\quad + (3\sqrt{2}+2\sqrt{3}\,)\kappa_2^2(R)\|F\|_{\mathrm{F}}^2\Big] \qquad (3.21)\\ &\le \frac{1}{2}(1+\sqrt{2}+\sqrt{3})\Big(\inf_{D\in\mathbb{D}_n}\zeta_D\|D^{-1}R\|_2\|R_{n-1}^{-1}D_{n-1}\|_2\Big)\|F_{n-1}\|_{\mathrm{F}} \\ &\quad + \frac{1}{2}\left(3+6\sqrt{2}+5\sqrt{3}+\sqrt{6}\right)\kappa_2^2(R)\|F\|_{\mathrm{F}}^2, \qquad (3.22)\end{aligned}$$

$$\begin{aligned}\frac{\|\Delta R\|_{\mathrm{F}}}{\|R\|_2} &\le \frac{\sqrt{2}\big(\inf_{D\in\mathbb{D}_n}\rho_D\kappa_2(D^{-1}R)\big)\left(\|F\|_{\mathrm{F}}+\kappa_2(A)\|F\|_{\mathrm{F}}^2\right)}{\sqrt{2}-1+\sqrt{1-4\kappa_2(A)\|F\|_{\mathrm{F}}-2\kappa_2^2(A)\|F\|_{\mathrm{F}}^2}} \qquad (3.23)\\ &\le (\sqrt{6}+\sqrt{3})\left(\inf_{D\in\mathbb{D}_n}\rho_D\kappa_2(D^{-1}R)\right)\|F\|_{\mathrm{F}}, \qquad (3.24)\end{aligned}$$

where D_{n-1} is the leading $(n-1)\times(n-1)$ principle matrix of D, and ζ_D and ρ_D are defined in (2.10).

Proof. We could apply (2.13) in Lemma 2.2 to obtain (3.24) easily, but we will provide a detailed proof anyway because much of it is needed to prove (3.22).

Since (3.20) holds, for any $t\in[0,1]$, $I+tF$ is nonsingular. Thus $A(I+tF)$ has full column rank and has the unique QR factorization

$$A(I+tF)=[Q+\Delta Q(t)][R+\Delta R(t)]. \qquad (3.25)$$

At $t=0,1$, it recovers the unique QR factorizations in (1.3) for A and $\widetilde{A}$. From (3.25), we have

$$(I+tF^{\mathrm{T}})A^{\mathrm{T}}A(I+tF)=[R+\Delta R(t)]^{\mathrm{T}}[R+\Delta R(t)].$$

By simple algebraic manipulations and using $A^{\mathrm{T}}A=R^{\mathrm{T}}R$, we obtain

$$\begin{aligned}&R^{-\mathrm{T}}\Delta R(t)^{\mathrm{T}}+\Delta R(t)R^{-1}\\ &=tR^{-\mathrm{T}}F^{\mathrm{T}}R^{\mathrm{T}}+tRFR^{-1}+t^2R^{-\mathrm{T}}F^{\mathrm{T}}R^{\mathrm{T}}RFR^{-1}-R^{-\mathrm{T}}\Delta R(t)^{\mathrm{T}}\Delta R(t)R^{-1}.\end{aligned}$$

Since $\Delta R(t)R^{-1}$ is upper triangular, it follows by using the "up" notation in (2.3) that

$$\begin{aligned}\Delta R(t)R^{-1}=\mathrm{up}\big(&tR^{-\mathrm{T}}F^{\mathrm{T}}R^{\mathrm{T}}+tRFR^{-1}\\ &+t^2R^{-\mathrm{T}}F^{\mathrm{T}}R^{\mathrm{T}}RFR^{-1}-R^{-\mathrm{T}}\Delta R(t)^{\mathrm{T}}\Delta R(t)R^{-1}\big).\end{aligned} \qquad (3.26)$$

Then, by (2.7),

$$\|\Delta R(t)R^{-1}\|_{\mathrm{F}}\le\frac{1}{\sqrt{2}}\left(2t\,\kappa_2(R)\|F\|_{\mathrm{F}}+t^2\kappa_2^2(R)\|F\|_{\mathrm{F}}^2+\|\Delta R(t)R^{-1}\|_{\mathrm{F}}^2\right). \qquad (3.27)$$

Define $\alpha(t)=\|\Delta R(t)R^{-1}\|_{\mathrm{F}}$, $\beta(t)=2t\,\kappa_2(R)\|F\|_{\mathrm{F}}+t^2\kappa_2^2(R)\|F\|_{\mathrm{F}}^2$, and rewrite (3.27) as

$$\alpha(t)^2-\sqrt{2}\,\alpha(t)+\beta(t)\ge 0.$$

Inequality (3.20) ensures $1-2\beta(t)>0$. Thus we have either $\alpha(t)\le\alpha_1(t)$ or $\alpha(t)\ge\alpha_2(t)$, where

$$\alpha_1(t)=\frac{1}{\sqrt{2}}\left[1-\sqrt{1-2\beta(t)}\right]<\alpha_2(t)=\frac{1}{\sqrt{2}}\left[1+\sqrt{1-2\beta(t)}\right].$$

But $\alpha(t)$ is continuous and $\alpha(0) = \alpha_1(0) = 0 < \alpha_2(0)$. Therefore we must have $\alpha(t) \leq \alpha_1(t)$ for any $t \in [0, 1]$, and in particular, $\alpha(1) \leq \alpha_1(1)$, i.e.,

$$\|\Delta R R^{-1}\|_{\mathrm{F}} \leq \frac{1}{\sqrt{2}} \left(1 - \sqrt{1 - 4\kappa_2(R)\|F\|_{\mathrm{F}} - 2\kappa_2^2(R)\|F\|_{\mathrm{F}}^2}\right) < \frac{1}{\sqrt{2}}. \tag{3.28}$$

Also the first inequality in (3.28) gives

$$\begin{aligned}\|\Delta R R^{-1}\|_{\mathrm{F}} &\leq \frac{1}{\sqrt{2}} \frac{4\kappa_2(R)\|F\|_{\mathrm{F}} + 2\kappa_2^2(R)\|F\|_{\mathrm{F}}^2}{1 + \sqrt{1 - 4\kappa_2(R)\|F\|_{\mathrm{F}} - 2\kappa_2^2(R)\|F\|_{\mathrm{F}}^2}} \\ &\leq \frac{1}{\sqrt{2}} \left[4\kappa_2(R)\|F\|_{\mathrm{F}} + 2\kappa_2^2(R)\|F\|_{\mathrm{F}}^2\right] \\ &\leq (\sqrt{2} + \sqrt{3})\kappa_2(R)\|F\|_{\mathrm{F}}, \end{aligned} \tag{3.29}$$

where we have used the assumption $\kappa_2(R)\|F\|_{\mathrm{F}} < \sqrt{3/2} - 1$. For any $D \in \mathbb{D}_n$, we have from (3.26) with $t = 1$ that

$$\begin{aligned}\Delta R R^{-1} D = {} & \mathrm{up}\left(D^{-1}(DR^{-\mathrm{T}}F^{\mathrm{T}}R^{\mathrm{T}})D + RFR^{-1}D\right) \\ & + \mathrm{up}\left(R^{-\mathrm{T}}F^{\mathrm{T}}R^{\mathrm{T}}RFR^{-1}D\right) - \mathrm{up}\left(R^{-\mathrm{T}}\Delta R^{\mathrm{T}}\Delta R R^{-1}D\right).\end{aligned}$$

Then, from (2.8) and (2.5), it follows that

$$\begin{aligned}\|\Delta R R^{-1} D\|_{\mathrm{F}} \leq {} & \rho_D \|R\|_2 \|R^{-1}D\|_2 \|F\|_{\mathrm{F}} + \kappa_2(R)\|R\|_2\|R^{-1}D\|_2\|F\|_{\mathrm{F}}^2 \\ & + \|\Delta R R^{-1}\|_{\mathrm{F}} \|\Delta R R^{-1} D\|_{\mathrm{F}}.\end{aligned}$$

Therefore, using (3.28) and the fact that $\rho_D \geq 1$ by definition and (3.20), we obtain

$$\begin{aligned}\|\Delta R R^{-1} D\|_{\mathrm{F}} &\leq \frac{\sqrt{2}\rho_D \|R^{-1}D\|_2 \|R\|_2 \|F\|_{\mathrm{F}} \left[1 + \kappa_2(R)\|F\|_{\mathrm{F}}\right]}{\sqrt{2} - 1 + \sqrt{1 - 4\kappa_2(A)\|F\|_{\mathrm{F}} - 2\kappa_2^2(A)\|F\|_{\mathrm{F}}^2}} \\ &\leq \left(\sqrt{6} + \sqrt{3}\right) \rho_D \|R^{-1}D\|_2 \|R\|_2 \|F\|_{\mathrm{F}}.\end{aligned}$$

Combining the inequality $\|\Delta R\|_{\mathrm{F}} \leq \|\Delta R R^{-1} D\|_{\mathrm{F}} \|D^{-1}R\|_2$ and the above inequalities and noticing that $D \in \mathbb{D}_n$ is arbitrary, we obtain (3.23) and (3.24).

Now we prove (3.22). From (3.25) with $t = 1$ it follows that

$$\Delta Q = QRFR^{-1} - (Q + \Delta Q)\Delta R R^{-1}.$$

Then, using (3.26) with $t = 1$, we obtain

$$\begin{aligned}\Delta Q &= QRFR^{-1} - Q\,\mathrm{up}\big(R^{-\mathrm{T}}F^{\mathrm{T}}R^{\mathrm{T}} + RFR^{-1}\big) - \Delta Q\,\mathrm{up}\big(R^{-\mathrm{T}}F^{\mathrm{T}}R^{\mathrm{T}} + RFR^{-1}\big) \\ &\quad - (Q + \Delta Q)\mathrm{up}\big(R^{-\mathrm{T}}F^{\mathrm{T}}R^{\mathrm{T}}RFR^{-1}\big) + (Q + \Delta Q)\mathrm{up}\big(R^{-\mathrm{T}}\Delta R^{\mathrm{T}}\Delta R R^{-1}\big) \\ &= Q\left\{\mathrm{low}(RFR^{-1}) - [\mathrm{low}(RFR^{-1})]^{\mathrm{T}}\right\} - \Delta Q\,\mathrm{up}\big(R^{-\mathrm{T}}F^{\mathrm{T}}R^{\mathrm{T}} + RFR^{-1}\big) \\ &\quad - (Q + \Delta Q)\mathrm{up}\big(R^{-\mathrm{T}}F^{\mathrm{T}}R^{\mathrm{T}}RFR^{-1}\big) + (Q + \Delta Q)\mathrm{up}\big(R^{-\mathrm{T}}\Delta R^{\mathrm{T}}\Delta R R^{-1}\big).\end{aligned} \tag{3.30}$$

To bound the first term on the right hand side of (3.30), we write

$$\begin{aligned}&D = \mathrm{diag}(D_{n-1}, \delta_n) \in \mathbb{D}_n, \quad D_{n-1} \in \mathbb{D}_{n-1}, \\ &R = D\hat{R}, \quad \hat{R} = \begin{bmatrix}\hat{R}_{n-1} & \hat{r} \\ 0 & \hat{r}_{nn}\end{bmatrix}, \quad \hat{R}_{n-1} \in \mathbb{R}^{(n-1)\times(n-1)}.\end{aligned}$$

It is easy to verify that the matrix formed by the first $n-1$ columns of $\hat{R}F\hat{R}^{-1}$ is $\hat{R}F_{n-1}\hat{R}_{n-1}^{-1}$. Then, by (2.9) we have

$$\begin{aligned}\|\mathrm{low}(RFR^{-1}) &- (\mathrm{low}(RFR^{-1}))^{\mathrm{T}}\|_{\mathrm{F}} \\ &= \|D\mathrm{low}(\hat{R}F\hat{R}^{-1})D^{-1} - D^{-1}(\mathrm{low}(\hat{R}F\hat{R}^{-1}))^{\mathrm{T}}D\|_{\mathrm{F}} \\ &\le \sqrt{2}\,\zeta_D\|\hat{R}F_{n-1}\hat{R}_{n-1}^{-1}\|_{\mathrm{F}} \\ &\le \sqrt{2}\,\zeta_D\|D^{-1}R\|_2\|R_{n-1}^{-1}D_{n-1}\|_2\|F_{n-1}\|_{\mathrm{F}}.\end{aligned}$$

The remaining terms in (3.30) can be bounded using (2.6), (2.7) and (3.29). Thus we obtain

$$\begin{aligned}\|\Delta Q\|_{\mathrm{F}} &\le \sqrt{2}\,\zeta_D\|D^{-1}R\|_2\|R_{n-1}^{-1}D_{n-1}\|_2\|F_{n-1}\|_{\mathrm{F}} + \sqrt{2}\|RFR^{-1}\|_{\mathrm{F}}\|\Delta Q\|_{\mathrm{F}} \\ &\quad + \frac{1}{\sqrt{2}}\|RFR^{-1}\|_{\mathrm{F}}^2 + \frac{1}{\sqrt{2}}\|\Delta RR^{-1}\|_{\mathrm{F}}^2 \\ &\le \sqrt{2}\,\zeta_D\|D^{-1}R\|_2\|R_{n-1}^{-1}D_{n-1}\|_2\|F_{n-1}\|_{\mathrm{F}} + \sqrt{2}\,\kappa_2(R)\|F\|_{\mathrm{F}}\|\Delta Q\|_{\mathrm{F}} \\ &\quad + \frac{1}{\sqrt{2}}\kappa_2^2(R)\|F\|_{\mathrm{F}}^2 + \frac{1}{\sqrt{2}}(\sqrt{2}+\sqrt{3})^2\kappa_2^2(R)\|F\|_{\mathrm{F}}^2.\end{aligned}$$

Then, with (3.20) it follows that

$$\begin{aligned}\|\Delta Q\|_{\mathrm{F}} &\le \frac{\sqrt{2}\,\zeta_D\|D^{-1}R\|_2\|R_{n-1}^{-1}D_{n-1}\|_2\|F_{n-1}\|_{\mathrm{F}} + (3\sqrt{2}+2\sqrt{3}\,)\kappa_2^2(R)\|F\|_{\mathrm{F}}^2}{1-\sqrt{2}\,\kappa_2(R)\|F\|_{\mathrm{F}}} \\ &\le \frac{1}{2}(1+\sqrt{2}+\sqrt{3})\zeta_D\|D^{-1}R\|_2\|R_{n-1}^{-1}D_{n-1}\|_2\|F_{n-1}\|_{\mathrm{F}} \\ &\quad + \frac{1}{2}(3+6\sqrt{2}+5\sqrt{3}+\sqrt{6})\kappa_2^2(R)\|F\|_{\mathrm{F}}^2,\end{aligned}$$

leading to (3.21) and (3.22). □

REMARK **3.5.** For the R-factor, the perturbation bound (3.24) does not have improvement over the perturbation bound by (2.13), noticing

$$\|\Delta A\|_{\mathrm{F}} = \|AF\|_2 \le \|A\|_2\|F\|_{\mathrm{F}}.$$

For the Q-factor, the bound (3.22) is interesting. Note that the coefficient of $\|F\|_{\mathrm{F}}$ in (3.22) can sometimes be very small because of the tininess of ζ_D. For example, for $R=\mathrm{diag}(1,\epsilon)$ with small $\epsilon>0$, take $D=\mathrm{diag}(1,\epsilon)$ to get

$$\zeta_D\|D^{-1}R\|_2\|R_{n-1}^{-1}D_{n-1}\|_2 = \epsilon.$$

Thus it is possible for the second order term in (3.22) to dominate the bound.

REMARK **3.6.** From (3.21) and (3.23) we obtain the following first-order bounds:

$$\begin{aligned}\|\Delta Q\|_{\mathrm{F}} &\le \sqrt{2}\left(\inf_{D\in\mathbb{D}_n}\ \zeta_D\|D^{-1}R\|_2\|R_{n-1}^{-1}D_{n-1}\|_2\right)\|F_{n-1}\|_{\mathrm{F}} + O(\|F\|_{\mathrm{F}}^2), \\ \frac{\|\Delta R\|_{\mathrm{F}}}{\|R\|_2} &\le \left(\inf_{D\in\mathbb{D}_n}\ \rho_D\kappa_2(D^{-1}R)\right)\|F\|_{\mathrm{F}} + O(\|F\|_{\mathrm{F}}^2).\end{aligned}$$

REMARK **3.7.** We would like to choose D such that $\zeta_D\|D^{-1}R\|_2\|R_{n-1}^{-1}D_{n-1}\|_2$ is a good approximation to $\inf_{D\in\mathbb{D}_n}\zeta_D\|D^{-1}R\|_2\|R_{n-1}^{-1}D_{n-1}\|_2$ in (3.22) and also choose D such that $\rho_D\kappa_2(D^{-1}R)$ is a good approximation to $\inf_{D\in\mathbb{D}_n}\rho_D\kappa_2(D^{-1}R)$ in (3.24). For the latter, numerical experiments in [7] indicated that a good choice for $D = \operatorname{diag}(\delta_1, \ldots, \delta_n)$ is to equilibrate the rows of $R = (r_{ij})$ while keeping $\zeta_D \le 1$. Specifically, we take

$$\delta_0 = 0, \quad \text{and} \quad \delta_i = \min\left\{\delta_{i-1}, \sqrt{\sum_{j=i}^{n} r_{ij}^2}\right\} \quad \text{for } 1 \le i \le n. \tag{3.31}$$

Obviously this choice should also be good for the former.

REMARK **3.8.** Suppose we use the standard column pivoting (see, e.g., [10, section 5.4.1]) in the QR factorization of A and the same permutation matrix is applied to the QR factorization of $A(I+F)$, then it is easy to see that the bounds (3.22) and (3.24) still hold. If we choose $D = \operatorname{diag}(r_{11}, r_{22}, \ldots, r_{nn})$, then according to [7, section 5.1],

$$\rho_D \kappa_2(D^{-1}R) \le 2^{n-1}\sqrt{n(n+1)}.$$

By the same proof given in [7, section 5.1], we see that

$$\begin{aligned}\zeta_D \|D^{-1}R\|_2 \|R_{n-1}^{-1} D_{n-1}\|_2 &\le \left(\max_{2\le i\le n} \frac{r_{ii}}{r_{i-1,i-1}}\right) \cdot 2^{n-2} \cdot \sqrt{n(n+1)/2} \\ &\le 2^{n-2}\sqrt{n(n+1)/2}.\end{aligned}$$

These inequalities suggest that the standard column pivoting is likely to decrease the size of the perturbations in the R-factor and the size of the perturbations in the Q-factor as well if the second-order bound in (3.22) does not dominate the perturbation bound.

We can also refine the perturbation bound in (3.22) by using the same approach as we did for a left multiplicative perturbation. The inequality (3.22) still holds if $\|F\|_F$ is replaced by

$$\min_{\alpha} \|\alpha D_R - 1\|_F = \|\hat{\alpha} D_R - I\|_F.$$

Theorem 3.5. *With the notation and assumption of* Theorem 3.4*,*

$$\begin{aligned}\|\Delta Q\|_F \le{}& \frac{1}{2}(1+\sqrt{2}+\sqrt{3})\left(\inf_{D\in\mathbb{D}_n} \zeta_D \|D^{-1}R\|_2 \|R_{n-1}^{-1} D_{n-1}\|_2\right) \|\hat{\alpha} D_R - I\|_F \\ &+ \frac{1}{2}(3+6\sqrt{2}+5\sqrt{3}+\sqrt{6})\kappa_2^2(R)\|\hat{\alpha} D_R - I\|_F^2,\end{aligned} \tag{3.32}$$

where $\hat{\alpha} = \operatorname{trace}(D_R)/\|D_R\|_F^2$.

REMARK **3.9.** When D_R itself in (3.19) is upper triangular and nonsingular:

$$\widetilde{Q} = Q\Lambda, \quad \widetilde{R} = \Lambda R D_R,$$

where Λ is the diagonal matrix whose ith diagonal entry is the sign of the ith diagonal entry of D_R. In particular if also D_R has positive diagonal entries, then $\Lambda = I$ and $\widetilde{Q} = Q$ and $\widetilde{R} = RD_R$. This observation in principle can be used to derive sharper bounds on the R-factor as we did in Theorem 3.3. But the gain is not substantial, however. So we omit the detail.

S	Lemma 2.2		Theorems 3.1		Theorems 3.2		Theorems 3.3	
	(2.11)	(2.13)	(3.2)	(3.3)	(3.11)	(3.12)	(3.17)	(3.18)
S_0	$1 \cdot 10^{-1}$	$3 \cdot 10^{-7}$	$3 \cdot 10^{-7}$	$4 \cdot 10^{-7}$	$3 \cdot 10^{-7}$	$4 \cdot 10^{-7}$	$2 \cdot 10^{-7}$	$3 \cdot 10^{-7}$
	$\|\Delta Q\|_F = 2 \cdot 10^{-7}$ and $\|\Delta R\|_F / \|R\|_2 = 2 \cdot 10^{-8}$							
S_1	$7 \cdot 10^{-2}$	$1 \cdot 10^{+0}$	$3 \cdot 10^{-7}$	$4 \cdot 10^{-7}$	$3 \cdot 10^{-7}$	$4 \cdot 10^{-7}$	$2 \cdot 10^{-7}$	$3 \cdot 10^{-7}$
	$\|\Delta Q\|_F = 2 \cdot 10^{-7}$ and $\|\Delta R\|_F / \|R\|_2 = 2 \cdot 10^{-7}$							
US_0V^T	$1 \cdot 10^{-1}$	$1 \cdot 10^{-5}$	$3 \cdot 10^{-7}$	$4 \cdot 10^{-7}$	$3 \cdot 10^{-7}$	$4 \cdot 10^{-7}$	$2 \cdot 10^{-7}$	$3 \cdot 10^{-7}$
	$\|\Delta Q\|_F = 2 \cdot 10^{-7}$ and $\|\Delta R\|_F / \|R\|_2 = 5 \cdot 10^{-8}$							

TABLE 4.1. Example 1: Column 2 to column 9 in the 3rd, 5th, and 7th row are the right-hand sides of the corresponding inequalities. The "exact" errors are given in the 4th, 6th and 8th row.

4. **Examples.** By making D_L and/or D_R not close to the identity matrices but to some scalar multiples of the identity matrices or even some orthogonal matrices, it is almost trivial to create an example for these "optimized" bounds in Theorems 3.2, 3.3, 3.5 to be arbitrarily tighter than their counterparts in Theorems 3.1 and 3.4. So in what follows, we will only consider examples in which D_L and/or D_R are close to the identity matrices. Consequently, we expect the "optimized" bounds are only marginally better.

Example 1. We construct this example for left multiplicative perturbations by letting $A = BS$, where B is an $m \times n$ well-conditioned matrix and S is some scaling matrix whose entries vary widely in magnitude. Then we perturb B to $B + \Delta B$ and finally

$$\widetilde{A} = (B + \Delta B)S = [I_m + (\Delta B)B^\dagger]BS \equiv D_L A,$$

where $B^\dagger = (B^T B)^{-1} B^T$, $D_L = I_m + (\Delta B)B^\dagger \equiv I_m + E$. Take $m = 9$, $n = 5$, and

$$B = \begin{bmatrix} 8 & 2 & -12 & -5 & -10 \\ 6 & 0 & 0 & -3 & 13 \\ -8 & -10 & -11 & -12 & 3 \\ -3 & -9 & -13 & -13 & 15 \\ -12 & -4 & -3 & 9 & 11 \\ -22 & -12 & 10 & 0 & -7 \\ 10 & -11 & 1 & -6 & -13 \\ -5 & 15 & 7 & 8 & -1 \\ 3 & 1 & -12 & 2 & -3 \end{bmatrix}, \tag{4.1}$$

$$S_0 = \begin{bmatrix} 10^6 & & & & \\ & 10^4 & & & \\ & & 10^3 & & \\ & & & 10^2 & \\ & & & & 1 \end{bmatrix}, \quad S = S_0,\ S_1, \text{ or } US_0V^T, \tag{4.2}$$

where S_1 is obtained from rearranging the diagonal entries of S_0 in the increasing order, $U, V \in \mathbb{R}^{n \times n}$ are two random orthogonal matrices. We take

$$\Delta B = \texttt{1.0e} - \texttt{6} * (\texttt{2} * \texttt{rand(m,n)} - \texttt{1})$$

in MATLAB-like notation. In Table 4.1, we list various error bounds in Lemma 2.2, Theorems 3.1 – 3.3. In computing these bounds, we treat the computed QR factorizations of A by MATLAB's `qr(A,0)` as exact ones. This is justifiable because $\kappa_2(A)$ is about 10^6. Therefore the computed Q and R will have at least about 10

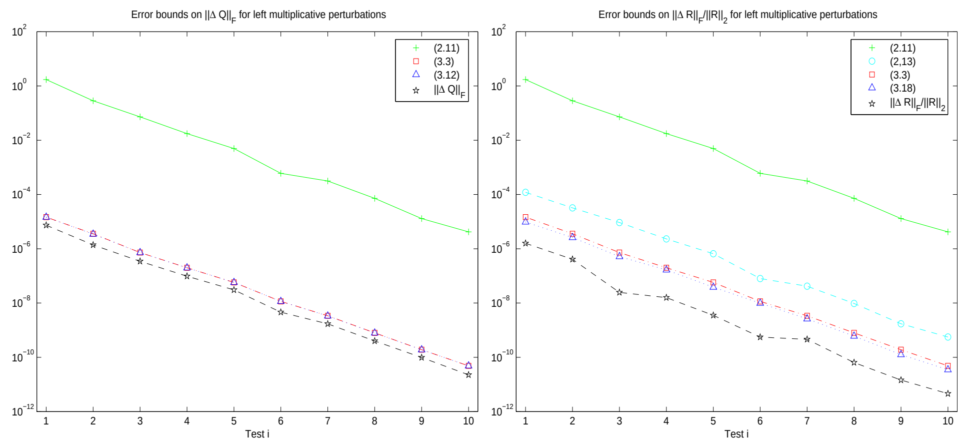

FIGURE 4.1. Example 2: Error bounds on the changes of the QR-factors under left multiplicative perturbations.

correct decimal digits which is good enough to be treated as exact for the given perturbation $\Delta A = (\Delta B)S$. We used $D = \text{diag}(\delta_1, \dots, \delta_n)$ for (2.13) computed by (3.31) to get

$$\begin{aligned} &\zeta_D = 9.80 \cdot 10^{-2}, \quad \rho_D = 1.00, \quad \kappa(D^{-1}R) = 1.05, && \text{for } S = S_0; \\ &\zeta_D = 1.00 \cdot 10^{-2}, \quad \rho_D = 1.41, \quad \kappa(D^{-1}R) = 5 \cdot 10^6, && \text{for } S = S_1; \\ &\zeta_D = 5.88 \cdot 10^{-1}, \quad \rho_D = 1.16, \quad \kappa(D^{-1}R) = 4.19 \cdot 10^1, && \text{for } S = US_0V^{\mathrm{T}}. \end{aligned}$$

It can be seen from this table that the bounds by (2.11) in Lemma 2.2 are very poor, but amazingly the one by (2.13) is comparable to those by Theorems 3.1 – 3.3 for $S = S_0$ and only about 25 to 50 times bigger for $S = US_0V^{\mathrm{T}}$ but extremely poor for $S = S_1$.

We point out that starting with $A = BS$ is purely for the convenience of constructing an example of left multiplicative perturbations. Once $\widetilde{A} = D_{\mathrm{L}}A$ is done, our bounds by Theorems 3.1 – 3.3 do not need to know this structure in A in order to give the sharper bounds as in the table. ◇

Example 2. We simply take $A = U\Sigma V^{\mathrm{T}}$, where $U \in \mathbb{R}^{m\times n}$ is random and has orthonormal columns, $V \in \mathbb{R}^{n\times n}$ is a random orthogonal matrix, and Σ is diagonal to make $\kappa_2(A)$ about 10^5. Again $m = 9$ and $n = 5$. We set $D_{\mathrm{L}} = I_m + \epsilon \times \texttt{randn(m)}$ for the case of left multiplicative perturbations and $D_{\mathrm{R}} = I_n + \epsilon \times \texttt{randn(n)}$ for the case of right multiplicative perturbations, where `randn` is MATLAB's random matrix generator, and $\epsilon = 10^{-6}/4^i$ for $i = 0, 1, \dots, 9$. Again we treat the computed QR factorizations of A by MATLAB's `qr(A,0)` as exact ones. This is justifiable because $\kappa_2(A)$ is about 10^5, and so the computed Q and R will have at least about 11 correct decimal digits. Figure 4.1 shows the error bounds on the changes in the QR-factors under left multiplicative perturbations while Figure 4.2 presents the same information but under right multiplicative perturbations. Also shown are the bounds in Lemma 2.2 which were established under the general additive perturbation assumption and the "exact" errors between the QR factors computed by MATLAB's `qr` function. From the figures, we may come to the following conclusions:

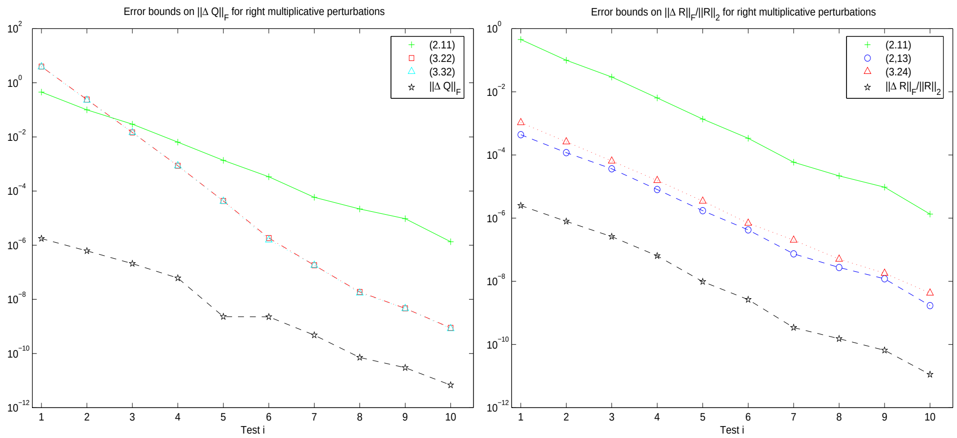

FIGURE 4.2. Example 2: Error bounds on the changes of the QR-factors under right multiplicative perturbations. In the left plot, the curves for (3.22) and (3.32) overlap each other.

1. In all cases, the existing (2.11) gives the worst bounds, except that the bounds by (3.22) and (3.32) for ΔQ (under right multiplicative perturbations) give the worst bounds for larger perturbations but the best bounds for smaller perturbations.
2. The existing (2.13), even though established under the general additive perturbation assumption, is quite competitive to those established here under multiplicative perturbations, although there are examples for which (2.13) is very poor as in Example 1 when $S = S_1$.
3. The behaviors of the bounds by (3.22) and (3.32) for ΔQ are rather interesting. In the left plot of Figure 4.2, the two curves begin above the one for (2.11) for larger perturbations and then moves down to below it as perturbations become smaller. This seems to reflect the comment we made in Remark 3.5.
4. The bound by (3.24) for right multiplicative perturbations are actually worse than the one by (2.13), even though (2.13) was established under the general additive perturbation assumption. This confirms the comment we made in Remark 3.5.
5. As expected, the optimized versions – those in Theorems 3.2, 3.3, and Theorem 3.5 – are only marginally sharper than their counterparts in these tests because $D_{\rm L}$ and $D_{\rm R}$ are made close to the identity matrices.

5. Concluding remarks.

We have performed a multiplicative perturbation analysis for the QR factorization, designed to take advantage of the structure that comes with the perturbations. Several bounds for the relative changes in the QR factors have been established. They imply that both the QR factors are well conditioned with respect to a left multiplicative perturbation, but the same claim can not be said for the bounds for a right multiplicative perturbation since these bounds are dependent on the condition number of the scaled R-factor.

Multiplicative perturbations arise naturally from matrix scaling, a commonly used technique to improve the conditioning of a matrix. More can be said for left multiplicative perturbations which arise every time we compute QR factorizations by the Householder transformations. Let

$$A = BS, \quad S = \operatorname{diag}(\|A_{(:,1)}\|_2, \dots, \|A_{(:,n)}\|_2).$$

Suppose that $\kappa_2(B)$ is modest (and thus B has full column rank). Let $\widehat{Q}$ and $\widetilde{R}$ be the computed QR factors by the Householder transformations. According to Theorem 19.4 in [11, p.360], there exists an orthonormal $\widetilde{Q} \in \mathbb{R}^{m\times n}$ such that

$$A + \Delta A = \widetilde{Q}\widetilde{R}, \quad \widehat{Q} = \widetilde{Q} + O(\boldsymbol{u}), \tag{5.1a}$$

$$\|(\Delta A)_{(:,j)}\|_2 = O(\boldsymbol{u})\,\|A_{(:,j)}\|_2 \text{ for } 1 \le j \le n, \tag{5.1b}$$

where $\boldsymbol{u}$ is the unit machine roundoff. But the computed $\widetilde{R}$ does not necessarily have positive diagonal entries. If we let $\Lambda = \operatorname{diag}(\pm 1) \in \mathbb{R}^{n\times n}$ such that $\Lambda\widetilde{R}$ does have positive diagonal entries, then after the following substitutions:

$$\widetilde{R} \leftarrow \Lambda\widetilde{R},\ \widehat{Q} \leftarrow \widehat{Q}\Lambda,\ \widetilde{Q} \leftarrow \widetilde{Q}\Lambda,$$

the equations in (5.1) still hold and $A + \Delta A = \widetilde{Q}\widetilde{R}$ is now the QR factorization in the sense of what we specified at the beginning of this paper and is unique. Assume that such post-substitutions are performed, e.g., in (5.1) $\widetilde{R}$ is made to have positive diagonal entries. It follows from (5.1) that

$$\widetilde{A} \equiv A + \Delta A = [B + (\Delta A)S^{-1}]S = (I+E)\,BS = (I+E)A,$$

where $E = (\Delta A)S^{-1}B^{\dagger}$ and $\|E\|_p \le O(\boldsymbol{u})\,\|B^{\dagger}\|_p$ for $p = 2, \mathrm{F}$. Let $A = QR$ be the unique QR factorization of A, and apply Theorem 3.1 to conclude

$$Q = \widetilde{Q} + O(\boldsymbol{u})\,\kappa_2(B) = \widehat{Q} + O(\boldsymbol{u})\,\kappa_2(B), \quad (\widetilde{R} - R)R^{-1} = O(\boldsymbol{u})\,\kappa_2(B), \tag{5.2}$$

where all $O(\boldsymbol{u})$ are no bigger than $\boldsymbol{u}$ times some low degrees of polynomials in m and n. The equations in (5.2) are also implied by the results in [23].

REFERENCES

[1] R. Bhatia, *Matrix factorizations and their perturbations*, Linear Algebra Appl., **1**97/198 (1994), 245–276.

[2] R. Bhatia, "Matrix Analysis," Graduate Texts in Mathematics, Vol. 169. Springer, New York, 1997.

[3] R. Bhatia and K. Mukherjea, *Variation of the unitary part of a matrix*, SIAM J. Matrix Anal. Appl., **1**5 (1994), 1007–1014.

[4] Åke Björck, "Numerical Methods for Least Squares Problems," SIAM, Philadelphia, 1996.

[5] X.-W. Chang, *On the perturbation of the Q-factor of the QR factorization*, to appear in Numer. Linear Algebra Appl., 13 pages.

[6] X.-W. Chang and C. C. Paige, *Componentwise perturbation analyses for the QR factorization*, Numer. Math., **8**8 (2001), 319–345.

[7] X.-W. Chang, C. C. Paige and G. W. Stewart, *Perturbation analyses for the qr factorization*, SIAM J. Matrix Anal. Appl., **1**8 (1997), 775–791.

[8] X.-W. Chang and D. Stehlé, *Rigorous perturbation bounds of some matrix factorizations*, SIAM J. Matrix Anal. Appl., **3**1 (2010), 2841–2859.

[9] X.-W. Chang, D. Stehlé and G. Villard, *Perturbation analysis of the QR factor R in the context of LLL lattice basis reduction*, to appear in Mathematics of Computation, 25 pages.

[10] G. H. Golub and C. F. Van Loan, "Matrix Computations," Johns Hopkins University Press, Baltimore, Maryland, 3rd edition, 1996.

[11] N. J. Higham, "Accuracy and Stability of Numerical Algorithms," SIAM, Philadephia, 2nd edition, 2002.

[12] R.-C. Li, *Relative perturbation bounds for the unitary polar factor*, BIT, **37** (1997), 67–75.

[13] R.-C. Li, *Relative perturbation theory: I. eigenvalue and singular value variations*, SIAM J. Matrix Anal. Appl., **1**9 (1998), 956–982.

[14] R.-C. Li, *Relative perturbation theory: II. eigenspace and singular subspace variations*, SIAM J. Matrix Anal. Appl., 20 (1999), 471–492.
[15] R.-C. Li, *Relative perturbation bounds for positive polar factors of graded matrices*, SIAM J. Matrix Anal. Appl., **2**5 (2005), 424–433.
[16] R.-C. Li and G. W. Stewart, *A new relative perturbation theorem for singular subspaces*, Linear Algebra Appl., **3**13 (2000), 41–51.
[17] G. W. Stewart, *Perturbation bounds for the QR factorization of a matrix*, SIAM J. Numer. Anal., **1**4 (1977), 509–518.
[18] G. W. Stewart, *On the perturbation of LU, Cholesky and QR factorizations*, SIAM J. Matrix Anal. Appl., **1**4 (1993), 1141–1146.
[19] G. W. Stewart and J. G. Sun, "Matrix Perturbation Theory," Academic Press, Boston, 1990.
[20] J. G. Sun, *Perturbation bounds for the Cholesky and QR factorizations*, BIT, **3**1 (1991), 341–352.
[21] J. G. Sun, *Componentwise perturbation bounds for some matrix decompositions*, BIT, **3**2 (1992), 702–714.
[22] J. G. Sun, *On perturbation bounds for the QR factorization*, Linear Algebra Appl., **2**15 (1995), 95–112.
[23] H. Zha, *A componentwise perturbation analysis of the QR decomposition*, SIAM J. Matrix Anal. Appl., **1**4 (1993), 1124–1131.

Received December 2010; revised May 2011.

E-mail address: chang@cs.mcgill.ca
E-mail address: rcli@uta.edu