Pinchings and Norms of Scaled Triangular Matrices

RAJENDRA BHATIA [a,†], WILLIAM KAHAN [b,†] AND REN-CANG LI [c,*]

[a] *Indian Statistical Institute, New Delhi 110 016, India;* [b] *Computer Science Division and Department of Mathematics, University of California at Berkeley, Berkeley, CA 94720, USA;* [c] *Department of Mathematics, University of Kentucky, Lexington, KY 40506, USA*

Communicated by W. Watkins

(Received 15 December 2000)

Suppose U is an upper-triangular matrix, and D a nonsingular diagonal matrix whose diagonal entries appear in nondescending order of magnitude down the diagonal. It is proved that

$$\|D^{-1}UD\| \geq \|U\|$$

for any matrix norm that is reduced by a pinching. In addition to known examples --weakly unitarily invariant norms -- we show that any matrix norm defined by

$$\|A\| \stackrel{\text{def}}{=} \max_{x\neq 0,\, y\neq 0} \frac{\text{Re}\,(x^* Ay)}{\phi(x)\psi(y)},$$

where $\phi(\cdot)$ and $\psi(\cdot)$ are two absolute vector norms, has this property. This includes ℓ_p operator norms as a special case.

Keywords: Triangular matrix; Scaling; ℓ_p operator norm; Unitarily invariant norm; Pinching inequality

AMS Subject Classifications: 15A42, 15A60, 65F35

1 INTRODUCTION

Suppose $U = (u_{ij})$ is an upper-triangular matrix, i.e., $u_{ij} = 0$ if $i > j$, and $D = \text{diag}(\delta_1, \delta_2, \ldots, \delta_n)$ a nonsingular diagonal matrix whose diagonal entries appear in ascending order of magnitude down the diagonal, i.e., $0 < |\delta_i| \leq |\delta_j|$ if $i < j$. The matrix $D^{-1}UD$ is still upper-triangular and its nonzero entries satisfy

$$|\delta_i^{-1} u_{ij} \delta_j| = |\delta_j/\delta_i|\,|u_{ij}| \geq |u_{ij}|. \tag{1.1}$$

since $i \le j$. As a consequence, we would expect that for certain matrix norms $\|\cdot\|$

$$\|D^{-1}UD\| \ge \|U\|. \tag{1.2}$$

Inequality (1.2) is clearly equivalent to $\|U\| \ge \|DUD^{-1}\|$; and if norm $\|\cdot\|$ is invariant under either transpose or conjugate transpose, then (1.2) is equivalent to $\|DLD^{-1}\| \ge \|L\|$ and $\|L\| \ge \|D^{-1}LD\|$ which are also equivalent to each other, where L is a lower-triangular matrix.

Inequality (1.2) may be used to bound the *componentwise relative condition number* $\||U^{-1}|\cdot|U|\|$ [1, p. 37], [2] for solving a linear triangular system, where $|\cdot|$ of a matrix is understood entrywise. In fact it implies

$$\||U^{-1}|\cdot|U|\| = \min_{D=\operatorname{diag}(\delta_i)\text{ with }0<|\delta_i|\le|\delta_j|\text{ for }i<j} \||D^{-1}U^{-1}|\cdot|UD|\|.$$

(1.2) also has applications in highly relative accurate singular value decompositions [3].

Inequality (1.2) is obviously true for any *absolute matrix norms* since absolute matrix norms are increasing functions of absolute values of the matrix's entries [4, Definition 5.5.9 and Theorem 5.5.10, p. 285]. Frequently used absolute matrix norms are

$$|A|_p \equiv |(a_{ij})|_p \stackrel{\text{def}}{=} \sqrt[p]{\sum_{i,j} |a_{ij}|^p} \quad \text{for } p \ge 1,$$

including the Frobenius norm ($p = 2$) as a special one. What about other matrix norms, especially the often used ℓ_p operator norms and the unitarily invariant norms? These norms are not, in general, increasing functions of the absolute values of the matrix entries. However, we shall prove that Inequality (1.2) is valid for them. This is true because they all satisfy the pinching inequality (see Section 2 below). More generally, we shall consider matrix norms defined by

$$\|A\|_{\phi,\psi} \stackrel{\text{def}}{=} \max_{x\ne 0,\, y\ne 0} \frac{\operatorname{Re}(x^*Ay)}{\phi(x)\psi(y)}, \quad \text{and}$$
$$\phi(\cdot) \quad \text{and} \quad \psi(\cdot) \text{ are two absolute vector norms.} \tag{1.3}$$

Here $\operatorname{Re}[\cdot]$ denotes the real part of a complex number. These include ℓ_p operator norms defined by

$$\|A\|_p \stackrel{\text{def}}{=} \max_{y\ne 0} \frac{\|Ay\|_p}{\|y\|_p} = \max_{x\ne 0,\, y\ne 0} \frac{\operatorname{Re}(x^*Ay)}{\|x\|_q\|y\|_p}, \quad \text{for } p \ge 1,$$

where q is p's dual defined by $1/p + 1/q = 1$, and

$$\|x\|_p = \sqrt[p]{\sum_j |\xi_j|^p}$$

for any vector $x = (\xi_1, \xi_2, \ldots, \xi_n)^*$.

The assumption that both $\phi(\cdot)$ and $\psi(\cdot)$ are absolute is essential. Here is a counter example: $n = 2$, $\phi(\cdot) = \|\cdot\|_2$, $\psi(\cdot) = \|B\cdot\|_2$, and

$$U = \begin{pmatrix} -3 & -1 \\ 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & \\ & 2 \end{pmatrix}, \quad \text{and} \quad B = \begin{pmatrix} 8 & 5 \\ -2 & -1 \end{pmatrix}.$$

Then $\phi(\cdot)$ is absolute; while $\psi(\cdot)$ is not. Calculations lead to

$$\begin{aligned} \|U\|_{\phi,\psi} &= \|UB^{-1}\|_2 = 5.42\cdots > 4.16\cdots = \|D^{-1}UDB^{-1}\|_2 \\ &= \|D^{-1}UD\|_{\phi,\psi}. \end{aligned}$$

With this example, counterexamples with none of $\phi(\cdot)$ and $\psi(\cdot)$ being absolute can be constructed: one way is to take $\psi(\cdot)$ as above and

$$\phi_\epsilon(\cdot) = \|I_\epsilon \cdot\|_2, \quad \text{for } I_\epsilon = \begin{pmatrix} 1 & \epsilon \\ \epsilon & 1 \end{pmatrix}$$

for small $|\epsilon|$. It is easy to verify that $\phi_\epsilon(\cdot)$ is not absolute for any $\epsilon \neq 0$. Then $\phi_0(\cdot) = \|\cdot\|_2$, the same as the $\phi(\cdot)$ above. Since both $\|U\|_{\phi_\epsilon,\psi}$ and $\|D^{-1}UD\|_{\phi_\epsilon,\psi}$ are continuous at $\epsilon = 0$, for sufficiently small $\epsilon \neq 0$ we have

$$\|U\|_{\phi_\epsilon,\psi} > \|D^{-1}UD\|_{\phi_\epsilon,\psi}.$$

2 PINCHINGS AND NORMS

Let A be an $n \times n$ matrix partitioned conformally into blocks as

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & \cdots & A_{NN} \end{pmatrix}, \tag{2.1}$$

where the diagonal blocks are square. The block diagonal matrix

$$\mathcal{C}(A) \stackrel{\text{def}}{=} \begin{pmatrix} A_{11} & & & \\ & A_{22} & & \\ & & \ddots & \\ & & & A_{NN} \end{pmatrix} \tag{2.2}$$

is called a *pinching* of A. It is known that pinching reduces A's weakly unitarily invariant norms* (and thus unitarily invariant norms), i.e.,

$$\text{The Pinching Inequality:} \quad \|\mathcal{C}(A)\| \leq \|A\| \tag{2.3}$$

holds for every weakly unitarily invariant norm. The following lemma enlarges the domain of this inequality.

LEMMA 2.1 *Every norm $\|\cdot\|_{\phi,\psi}$ defined as in (1.3) satisfies the pinching inequality.*

Proof We follow the idea used in [8]. Let ω be a primitive Nth root of unity, and let X be the diagonal matrix that in a partitioning conformal with (2.1) has the form

$$X = \operatorname{diag}(I, \omega I, \omega^2 I, \ldots, \omega^{N-1} I),$$

where the I's are the identity matrices of appropriate sizes. It can be seen that

$$\mathcal{C}(A) = \frac{1}{N}\sum_{j=0}^{N-1} X^{*j} A X^j. \tag{2.4}$$

Let x, y be vectors such that $\phi(x) = 1 = \psi(y)$ and $\|\mathcal{C}(A)\|_{\phi,\psi} = \mathrm{Re}[x^*\mathcal{C}(A)y]$. Then by (2.4) we have

$$\begin{aligned}\|\mathcal{C}(A)\|_{\phi,\psi} &= \frac{1}{N}\sum_{j=0}^{N-1} \mathrm{Re}[x^* X^{*j} A X^j y] \\ &= \frac{1}{N}\sum_{j=0}^{N-1} \mathrm{Re}[(X^j x)^* A (X^j y)].\end{aligned} \tag{2.5}$$

Now note that each X^j is a diagonal matrix with unimodular diagonal entries. Since ϕ and ψ are absolute norms, $\phi(X^j x) = 1 = \psi(X^j y)$ for all j. Hence, each of the summands on the right hand side of (2.5) is bounded by $\|A\|_{\phi,\psi}$. Consequently $\|\mathcal{C}(A)\|_{\phi,\psi} \leq \|A\|_{\phi,\psi}$. ■

In the next section, we shall need a special corollary of the pinching inequality:

LEMMA 2.2 *Let*

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

---

*A matrix norm $\|\cdot\|_{wu}$ is *weakly unitarily invariant*, if it satisfies, besides the usual properties of any norm, also

$$\|XAX^*\|_{wu} = \|A\|_{wu} \quad \text{for any unitary matrix } X.$$

See [5, pp. 97–107] or [6]. A matrix norm $\|\cdot\|$ is *unitarily invariant* if it satisfies, besides the usual properties of any norm, also

1. $\|XAY\| = \|A\|$ for any unitary matrices X and Y;
2. $\|A\| = \|A\|_2$ for any A having rank one.

See Bhatia [5, p. 91] or Stewart and Sun [7, pp. 74–88]. We denote by $\|\cdot\|$ a general *unitarily invariant norm*. The most frequently used unitarily invariant norms are the spectral norm $\|\cdot\|_2$ (also called the ℓ_2 operator norm) and the Frobenius norm $\|\cdot\|_F$.

be any block matrix in which the diagonal blocks are square. Let $0 \le \alpha \le 1$. *Then*

$$\left\| \begin{pmatrix} A_{11} & \alpha A_{12} \\ \alpha A_{21} & A_{22} \end{pmatrix} \right\| \le \|A\|$$

for every norm that satisfies the pinching inequality (2.3).

Proof Write

$$\begin{pmatrix} A_{11} & \alpha A_{12} \\ \alpha A_{21} & A_{22} \end{pmatrix} = \alpha A + (1-\alpha)\mathcal{C}(A),$$

and then use the triangle inequality and the pinching inequality. ■

3 NORMS OF SCALED TRIANGULAR MATRICES

THEOREM 3.1 *Suppose* U *is an upper-triangular matrix, and* D *a nonsingular diagonal matrix whose diagonal entries appear in ascending order of magnitude down the diagonal. Then the inequality*

$$\|D^{-1}UD\| \ge \|U\| \tag{1.2}$$

holds for any matrix norm that satisfies the pinching inequality (2.3).

Proof Without loss of any generality, we may assume $D = \mathrm{diag}(\delta_1, \delta_2, \ldots, \delta_n)$ with

$$1 = \delta_1 \le \delta_2 \le \cdots \le \delta_n.$$

Decompose $D = F_2 \cdots F_n$, where

$$F_i = \begin{pmatrix} I_{i-1} & \\ & \gamma_i I_{n-i+1} \end{pmatrix},$$

$\gamma_i = \delta_i/\delta_{i-1} \ge 1$ for $i = 2, 3, \ldots, n$, and I_k is the $k \times k$ identity matrix. Inequality (1.2) is proved if it is proved for every $D = F_i$. This is what we shall do.

Let

$$F = \begin{pmatrix} I & \\ & \gamma I \end{pmatrix},$$

where $\gamma \ge 1$ and the I's generally have different dimensions. Conformally partition

$$U = \begin{pmatrix} R & S \\ & T \end{pmatrix}.$$

To prove the inequality (1.2) for $D = F$, we have to show

$$\left\| \begin{pmatrix} R & S \\ & T \end{pmatrix} \right\| \le \left\| \begin{pmatrix} R & \gamma S \\ & T \end{pmatrix} \right\|$$

for every norm that satisfies the pinching inequality. This follows from Lemma 2.2. ■

4 NORMS AND WIPING OUT OF ENTRIES

The idea that we have used in proving Lemma 2.1 depends on expressing $C(A)$ as an average of special diagonal similarities of A. This can be done for some other parts of a matrix, too.

Let $\mathbb{S}$ be an equivalence relation on $\{1, 2, \ldots, n\}$ with equivalence classes $\mathbb{I}_1, \mathbb{I}_2, \ldots, \mathbb{I}_N$. Let $\mathcal{F}$ be the map on the space of $n \times n$ matrices that sets all those (i, j) entries – $(i, j) \notin \mathbb{S}$ – to zero and keeps the others unchanged. It is shown [8] that such a map has a representation

$$\mathcal{F}(A) = \frac{1}{N} \sum_{j=0}^{N-1} X^{*j} A X^j, \tag{4.1}$$

where X is a diagonal matrix whose ith diagonal entry is ω^{k-1} if $i \in \mathbb{I}_k$, and ω is a primitive Nth root of unity. Using the argument in the proof of Lemma 2.1 we see that

$$\|\mathcal{F}(A)\|_{\phi,\psi} \leq \|A\|_{\phi,\psi}, \tag{4.2}$$

for every norm defined in (1.3). A particularly interesting example of such a map $\mathcal{F}$ is the one that wipes out all entries of a matrix except those on the dexter and the sinister diagonals, i.e., $\mathcal{F}$ sets all A's entries to zero except those (i, j) entries for which either $i = j$ or $i + j = n + 1$, where nothing is changed.

For A and $C(A)$ as in (2.1) and (2.2), let

$$\mathcal{O}(A) = A - C(A)$$

be the off-diagonal part of A. A straightforward application of the pinching inequality shows that

$$\|\mathcal{O}(A)\|_{\phi,\psi} \leq 2\|A\|_{\phi,\psi}.$$

However, using the representation (2.4) we can show, as in [8], that

$$\|\mathcal{O}(A)\|_{\phi,\psi} \leq 2(1 - 1/N)\|A\|_{\phi,\psi}. \tag{4.3}$$

This inequality is sharp for the norm $\|\cdot\|_2$ [8].

The same idea can be used to estimate the norms $\|A - \mathcal{F}(A)\|_{\phi,\psi}$ using the representation (4.1).

Finally we remark that for the ℓ_p operator norms $(1 \leq p \leq \infty)$, we could obtain better bounds than (4.3) by an interpolation argument. To see this, we notice that the operator norms $\|A\|_1$ and $\|A\|_\infty$ are diminished if any entry of A is wiped out. Hence, we have

$$\|\mathcal{O}(A)\|_p \leq \|A\|_p \quad \text{for } p = 1, \infty. \tag{4.4}$$

Since $A \mapsto \mathcal{O}(A)$ is a linear map on the space of matrices, we get by interpolating between the values $p = 1, 2, \infty$, the inequality

$$\|\mathcal{O}(A)\|_p \leq (2 - 2/N)^{2/\max\{p,q\}} \|A\|_p, \tag{4.5}$$

where q is p's dual. For details of the interpolation method, the reader is referred to Reed and Simon [9, pp. 7–45]. (The reader should be aware that $\|A\|_p$ denotes the Schatten p-norm in [8], but here it denotes the ℓ_p operator norm.)

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.

References

[1] J. Demmel (1997). *Applied Numerical Linear Algebra.* SIAM, Philadelphia.

[2] N.J. Higham (1996). *Accuracy and Stability of Numerical Algorithms.* SIAM, Philadelphia.

[3] J.W. Demmel, M. Gu, S.C. Eisenstat, I. Slapničar, K. Veselić, and Z. Drmač (1999). Computing the singular value decomposition with high relative accuracy. *Linear Algebra and its Applications*, **299**, 21–80.

[4] R.A. Horn and C.R. Johnson (1985). *Matrix Analysis.* Cambridge University Press, Cambridge.

[5] R. Bhatia (1996). *Matrix Analysis. Graduate Texts in Mathematics*, Vol. 169. Springer, New York.

[6] R. Bhatia and J.A.R. Holbrook (1987). Unitary invariance and spectral variation. *Linear Algebra and Its Applications*, **95**, 43–68.

[7] G.W. Stewart and J.-G. Sun (1990). *Matrix Perturbation Theory.* Academic Press, Boston.

[8] R. Bhatia, M.-D. Choi and C. Davis (1989). Comparing a matrix to its off-diagonal part. *Operator Theory: Advances and Applications*, **40**, 151–164.

[9] M. Reed and B. Simon (1975). *Methods of Modern Mathematical Physics, II: Fourier Analysis, Self-Adjointness.* Academic Press, New York.