Norms of Interval Matrices

Farhadsefat, R.
2011

Institute of Computer Science
Academy of Sciences of the Czech Republic

Norms of Interval Matrices

Raena Farhadsefat, Jiří Rohn and Taher Lotfi

Technical report No. V-1122

24.08.2011

Pod Vodárenskou věží 2, 182 07 Prague 8, phone: +420 266 051 111, fax: +420 286 585 789,
e-mail:rohn@cs.cas.cz

Institute of Computer Science
Academy of Sciences of the Czech Republic

Norms of Interval Matrices

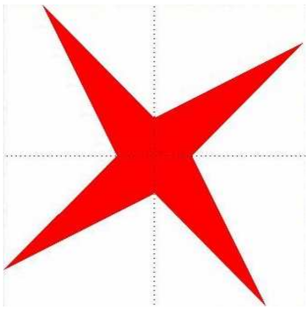
Raena Farhadsefat[1], Jiří Rohn[2] and Taher Lotfi[3]

Technical report No. V-1122

24.08.2011

Abstract:

Interval matrix norms induced by point matrix norms are introduced in the space of interval matrices. It is shown that evaluating the interval matrix norm induced by a point matrix norm $\|\cdot\|_p$ is exponential (probably NP-hard) for $p = 2$ and requires computation of only one point matrix norm for $p \in \{1, \infty, (1, \infty), F\}$.



Keywords:
Interval matrix, interval norm, norm, absolute norm.[4]

[1]Department of Applied Mathematics and Computer Sciences, Hamedan Branch, Islamic Azad University, Hamedan, Iran, e-mail: rftmath@gmail.com.

[2]Institute of Computer Science, Czech Academy of Sciences, Prague, Czech Republic, e-mail: rohn@cs.cas.cz. This author's work was supported by the Institutional Research Plan AV0Z10300504.

[3]Department of Applied Mathematics and Computer Sciences, Hamedan Branch, Islamic Azad University, Hamedan, Iran, e-mail: lotfi@iauh.ac.ir.

[4]Above: logo of interval computations and related areas (depiction of the solution set of the system $[2,4]x_1 + [-2,1]x_2 = [-2,2]$, $[-1,2]x_1 + [2,4]x_2 = [-2,2]$ (Barth and Nuding [3])).

1 Introduction and notation

This paper is dedicated to definition and properties of interval matrix norms, a topic not studied so far as far as known to the authors. In Section 2 we describe the space $\mathbb{IR}^{m\times n}$ of all $m\times n$ interval matrices enriched with operations of sum and scalar multiplication. Then in Section 3 we prove that given a matrix norm $\|\cdot\|$ in $\mathbb{R}^{m\times n}$, the mapping $|||\cdot||| : \mathbb{IR}^{m\times n} \to \mathbb{R}$ given by

$$|||\boldsymbol{A}||| = \sup\{\, \|A\| \mid A \in \boldsymbol{A}\,\} \tag{1.1}$$

is a norm in $\mathbb{IR}^{m\times n}$ which we call an interval matrix norm induced by the point matrix norm $\|\cdot\|$. After summing up some properties of induced interval matrix norms in Sections 3 and 4, we prove in Section 5 that the norm (1.1) can always be computed as maximum of the original norm $\|\cdot\|$ over the set of 2^{mn} so-called vertex matrices, and in Section 6 we show that this number can be decreased to 2^{m+n} for the interval matrix norm $|||\cdot|||_2$ induced by the point matrix norm $\|\cdot\|_2$. Both these results look pessimistic. But in the main result of this paper we prove in Section 7 that for each $\boldsymbol{A} = [A_c - \Delta, A_c + \Delta] \in \mathbb{IR}^{m\times n}$ there holds

$$|||\boldsymbol{A}|||_p = \|\, |A_c| + \Delta\,\|_p$$

for each $p \in \{1, \infty, (1,\infty), F\}$, so that for four of the five most (and, in fact, almost exclusively) used norms the induced interval matrix norm can be computed by evaluation of a *single* point matrix norm. This nice property is due to the fact that all the four norms are absolute. In Conclusion we formulate two problems associated with induced norms that remain to be solved.

The basic notation used is the following. Matrix inequalities, as $A \le B$, $A > 0$, and the absolute value $|A|$ of A are understood entrywise, and $A \circ B$ denotes the Hadamard (entrywise) product of two matrices of the same size. Other notation is introduced throughout the paper whenever needed.

2 The interval matrix space $\mathbb{IR}^{m\times n}$

Let m, n be fixed positive integers. If $\underline{A}$ and $\overline{A}$ are two matrices in $\mathbb{R}^{m\times n}$, $\underline{A} \le \overline{A}$, then the set of matrices

$$\boldsymbol{A} = [\underline{A}, \overline{A}] = \{\, A \mid \underline{A} \le A \le \overline{A}\,\}$$

is called an *interval matrix*. In many cases it is advantageous to express the data in terms of the center matrix $A_c = \frac{1}{2}(\underline{A} + \overline{A})$ and of the radius matrix $\Delta = \frac{1}{2}(\overline{A} - \underline{A})$, so that $\boldsymbol{A}$ can also be given as

$$\boldsymbol{A} = [A_c - \Delta, A_c + \Delta].$$

The set of all $m \times n$ interval matrices

$$\mathbb{IR}^{m\times n} = \{\, [\underline{A}, \overline{A}] \mid \underline{A} \le \overline{A},\ \underline{A}, \overline{A} \in \mathbb{R}^{m\times n}\,\}$$

is called the $m\times n$ interval matrix space. Notice the difference between $\mathbb{R}$ and $\mathbb{IR}$, the latter symbol being constructed by merging the letters $\mathbb{I}$ (for "interval") and $\mathbb{R}$. In particular, $[A, A] \in \mathbb{IR}^{m\times n}$ for each $A \in \mathbb{R}^{m\times n}$. To underline the distinction, we call the matrices from $\mathbb{R}^{m\times n}$ *point matrices* as opposites to interval matrices. Thus, $A \in \mathbb{R}^{m\times n}$ is a point matrix whereas $[A, A]$ is an interval matrix.

Next we introduce the sum and scalar multiplication operations in $\mathbb{IR}^{m\times n}$. For each $\boldsymbol{A} = [\underline{A}, \overline{A}] = [A_c - \Delta, A_c + \Delta] \in \mathbb{IR}^{m\times n}$, $\boldsymbol{B} = [\underline{B}, \overline{B}] = [B_c - \Delta', B_c + \Delta'] \in \mathbb{IR}^{m\times n}$ and for each $\alpha \in \mathbb{R}$ we define

$$\begin{aligned} \boldsymbol{A} + \boldsymbol{B} &= \{ A + B \mid A \in \boldsymbol{A},\ B \in \boldsymbol{B} \}, \\ \alpha \boldsymbol{A} &= \{ \alpha A \mid A \in \boldsymbol{A} \}. \end{aligned}$$

It can be easily proved that

$$\boldsymbol{A} + \boldsymbol{B} = [A_c + B_c - (\Delta + \Delta'), A_c + B_c + (\Delta + \Delta')] = [\underline{A} + \underline{B}, \overline{A} + \overline{B}], \tag{2.1}$$

$$\alpha \boldsymbol{A} = [\alpha A_c - |\alpha|\Delta, \alpha A_c + |\alpha|\Delta] = \begin{cases} [\alpha\underline{A}, \alpha\overline{A}] & \text{if } \alpha \geq 0, \\ [\alpha\overline{A}, \alpha\underline{A}] & \text{if } \alpha < 0 \end{cases} \tag{2.2}$$

(see Alefeld and Herzberger [2] or Neumaier [8]). From (2.1), (2.2), it follows that the sum and scalar multiplication of interval matrices possess the following properties for each $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C} \in \mathbb{IR}^{m\times n}$ and each $\alpha, \beta \in \mathbb{R}$:

$$\begin{aligned} \boldsymbol{A} + \boldsymbol{B} &= \boldsymbol{B} + \boldsymbol{A}, \\ (\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C} &= \boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C}), \\ \boldsymbol{A} + [0, 0] &= \boldsymbol{A}, \\ 1 \cdot \boldsymbol{A} &= \boldsymbol{A}, \\ \alpha(\boldsymbol{A} + \boldsymbol{B}) &= \alpha\boldsymbol{A} + \alpha\boldsymbol{B}, \\ \alpha(\beta\boldsymbol{A}) &= (\alpha\beta)\boldsymbol{A}. \end{aligned}$$

These are six of the eight properties defining a linear (vector) space, see [4]. The seventh property, namely

$$(\alpha + \beta)\boldsymbol{A} = \alpha\boldsymbol{A} + \beta\boldsymbol{A},$$

unfortunately does not hold in general: indeed, in view of (2.2) this property is equivalent to

$$\begin{aligned} & [(\alpha + \beta)A_c - |\alpha + \beta|\Delta, (\alpha + \beta)A_c + |\alpha + \beta|\Delta] \\ = \ & [(\alpha + \beta)A_c - (|\alpha| + |\beta|)\Delta, (\alpha + \beta)A_c + (|\alpha| + |\beta|)\Delta] \end{aligned}$$

which implies

$$|\alpha + \beta|\Delta = (|\alpha| + |\beta|)\Delta,$$

and this is evidently not true: it is sufficient to take $\alpha\beta < 0$ and $\Delta > 0$ to get a counterexample. The eighth property does not hold because for any $\boldsymbol{A} = [\underline{A}, \overline{A}] \in \mathbb{IR}^{m\times n}$ with $\underline{A} \neq \overline{A}$ there *does not* exist a $\boldsymbol{B} = [\underline{B}, \overline{B}] \in \mathbb{IR}^{m\times n}$ satisfying

$$\boldsymbol{A} + \boldsymbol{B} = [0, 0].$$

Indeed, this would mean $\underline{A} + \underline{B} = \overline{A} + \overline{B} = 0$ and hence $(\overline{A} - \underline{A}) + (\overline{B} - \underline{B}) = 0$, where both the summands are nonnegative matrices, implying $\underline{A} = \overline{A}$ and $\underline{B} = \overline{B}$, a contradiction. Thus, $\mathbb{IR}^{m\times n}$ with the two operations introduced is not a linear, but a "semilinear" space.

We do not introduce an interval matrix product in $\mathbb{IR}^{m\times n}$. The "obvious" way of defining it as

$$\boldsymbol{A}\boldsymbol{B} = \{\, AB \mid A \in \boldsymbol{A},\, B \in \boldsymbol{B} \,\}$$

(which would require that $m = n$) generally would not produce an interval matrix as a result, and defining it as $\boldsymbol{A} \odot \boldsymbol{B}$ (the interval arithmetic product, see [2]) would not suit our purposes.

3 Interval matrix norms

The fact that $\mathbb{IR}^{m\times n}$ is a sublinear, not a linear, space does not preclude a possibility of introducing a norm there. A function $|||\cdot||| : \mathbb{IR}^{m\times n} \to \mathbb{R}$ is called an *interval matrix norm* in $\mathbb{IR}^{m\times n}$ if for each $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{IR}^{m\times n}$, $\alpha \in \mathbb{R}$ it satisfies (a)-(c):

(a) $|||\boldsymbol{A}||| \geq 0$, and $|||\boldsymbol{A}||| = 0$ if and only if $\boldsymbol{A} = [0, 0]$,

(b) $|||\boldsymbol{A} + \boldsymbol{B}||| \leq |||\boldsymbol{A}||| + |||\boldsymbol{B}|||$,

(c) $|||\alpha\boldsymbol{A}||| = |\alpha|\,|||\boldsymbol{A}|||$.

As the reader might have noticed, we use the notation $|||\cdot|||$ for interval matrix norms in $\mathbb{IR}^{m\times n}$ as opposed to point matrix norms in $\mathbb{R}^{m\times n}$ denoted by $\|\cdot\|$.

The following theorem shows a way how to construct interval matrix norms from point matrix norms.

Theorem 1. *For any point matrix norm $\|\cdot\|$ in $\mathbb{R}^{m\times n}$, the function $|||\cdot||| : \mathbb{IR}^{m\times n} \to \mathbb{R}$ defined by*

$$|||\boldsymbol{A}||| = \sup\{\, \|A\| \mid A \in \boldsymbol{A} \,\} \tag{3.1}$$

is an interval matrix norm in $\mathbb{IR}^{m\times n}$.

Proof. We shall prove that the function $|||\cdot|||$ given by (3.1) possesses the above properties (a)-(c). Let $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{IR}^{m\times n}$ and $\alpha \in \mathbb{R}$. Then:

(a) Taking an arbitrary $A \in \boldsymbol{A}$, we have $|||\boldsymbol{A}||| \geq \|A\| \geq 0$. If $|||\boldsymbol{A}||| = 0$, then by (3.1) we have $\|A\| = 0$ for each $A \in \boldsymbol{A}$, hence $\boldsymbol{A} = \{0\}$, which is only possible if $\boldsymbol{A} = [0, 0]$.

(b) By definition of the sum of interval matrices and by the triangle inequality for the norm $\|\cdot\|$ we have

$$\begin{aligned} |||\boldsymbol{A} + \boldsymbol{B}||| &= \sup\{\, \|A + B\| \mid A \in \boldsymbol{A},\, B \in \boldsymbol{B} \,\} \leq \sup\{\, \|A\| + \|B\| \mid A \in \boldsymbol{A},\, B \in \boldsymbol{B} \,\} \\ &\leq \sup\{\, \|A\| \mid A \in \boldsymbol{A} \,\} + \sup\{\, \|B\| \mid B \in \boldsymbol{B} \,\} = |||\boldsymbol{A}||| + |||\boldsymbol{B}|||. \end{aligned}$$

(c) Similarly, by the definition of scalar multiplication and by the homogeneity of the norm $\|\cdot\|$ we have

$$|||\alpha\boldsymbol{A}||| = \sup\{\, \|\alpha A\| \mid A \in \boldsymbol{A} \,\} = |\alpha| \sup\{\, \|A\| \mid A \in \boldsymbol{A} \,\} = |\alpha|\, |||\mathbf{A}|||,$$

which concludes the proof. □

We shall say that the interval matrix norm $|||\cdot|||$ defined by (3.1) is *induced* by the point matrix norm $\|\cdot\|$. An interval matrix norm induced by some point matrix norm is called simply an induced interval matrix norm. If the point matrix norm is denoted by $\|A\|_\alpha$ for some α, then the induced interval matrix norm is denoted again by $|||\boldsymbol{A}|||_\alpha$.

Since each norm is a continuous function (Horn and Johnson [7]) and each interval matrix is a compact set, the supremum in (3.1) is attained as maximum, hence (3.1) can be equivalently rewritten as

$$|||\boldsymbol{A}||| = \max\{\, \|A\| \mid A \in \boldsymbol{A}\,\}. \tag{3.2}$$

It follows from the definition of an induced matrix norm that for each $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{IR}^{m\times n}$, $\boldsymbol{A} \subseteq \boldsymbol{B}$ implies $|||\boldsymbol{A}||| \le |||\boldsymbol{B}|||$. Also, for every two induced matrix norms $|||\boldsymbol{A}|||_\alpha$ and $|||\boldsymbol{A}|||_\beta$ there exist positive constants c and d such that

$$c\,|||\boldsymbol{A}|||_\beta \le |||\boldsymbol{A}|||_\alpha \le d\,|||\boldsymbol{A}|||_\beta$$

holds for each $\boldsymbol{A} \in \mathbb{IR}^{m\times n}$. This is simply the assertion of equivalence of norms, see [7].

If $m = n$, then a point matrix norm is called consistent if it satisfies $\|AB\| \le \|A\|\|B\|$ for each $A, B \in \mathbb{R}^{n\times n}$ ($m = n$ is needed here for the matrix product to be feasible). Since we have not introduced matrix multiplication in the space $\mathbb{IR}^{n\times n}$, we define this notion in another way. We say that an interval matrix norm $|||\cdot|||$ induced by a point matrix norm $\|\cdot\|$ is *consistent* if it satisfies

$$\max_{A\in\boldsymbol{A},\, B\in\boldsymbol{B}} \|AB\| \le |||\boldsymbol{A}|||\,|||\boldsymbol{B}||| \tag{3.3}$$

for each $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{IR}^{n\times n}$.

Theorem 2. *If $\|\cdot\|$ is a consistent point matrix norm, then $|||\cdot|||$ is a consistent induced interval matrix norm.*

Proof. If $A \in \boldsymbol{A} \in \mathbb{IR}^{n\times n}$, $B \in \boldsymbol{B} \in \mathbb{IR}^{n\times n}$, then

$$\|AB\| \le \|A\|\,\|B\| \le |||\boldsymbol{A}|||\,|||\boldsymbol{B}|||$$

and taking the maximum over $\boldsymbol{A}, \boldsymbol{B}$ gives (3.3). □

4 Characterization of induced norms

In this section we are going to characterize induced matrix norms in terms of their own (without use of inducing point matrix norms). As the first step towards this goal we show that the inducing point matrix norm can be reconstructed from the induced interval matrix norm.

Theorem 3. *If $|||\cdot|||$ is an induced interval matrix norm in $\mathbb{IR}^{m\times n}$, then the inducing matrix norm satisfies*

$$\|A\| = |||[A, A]||| \tag{4.1}$$

for each $A \in \mathbb{R}^{m\times n}$.

Proof. For each $A \in \mathbb{R}^{m\times n}$ we have $[A, A] = \{A\}$, hence $|||[A, A]||| = \|A\|$ by (3.1), which is (4.1). □

Hence, the inducing norm can be fully reconstructed from the induced norm and therefore it is unique: different point matrix norms induce different interval matrix norms.

Now we give a characterization of induced interval matrix norms.

Theorem 4. *A norm $|||\cdot|||$ in $\mathbb{IR}^{m\times n}$ is an induced interval matrix norm if and only if it satisfies*

$$|||\boldsymbol{A}||| = \max\{\, |||[A, A]||| \mid A \in \boldsymbol{A}\,\} \tag{4.2}$$

for each $\boldsymbol{A} \in \mathbb{IR}^{m\times n}$.

Proof. If $|||\cdot|||$ is an induced interval matrix norm, then

$$\|A\| = |||[A, A]||| \tag{4.3}$$

is the inducing point matrix norm by Theorem 3 and (3.2) implies (4.2). Conversely, if (4.2) holds for each $\boldsymbol{A} \in \mathbb{IR}^{m\times n}$, then the function $\|\cdot\|$ defined by (4.3) is a point matrix norm in $\mathbb{R}^{m\times n}$ (because $|||\cdot|||$ is a norm in $\mathbb{IR}^{m\times n}$ by assumption), and (4.2) implies (3.2), so that $|||\cdot|||$ is induced. □

5 Computing the norms I: the general case

Let E denote the $m \times n$ matrix of all ones. Given an $\boldsymbol{A} = [A_c - \Delta, A_c + \Delta] = [\underline{A}, \overline{A}] \in \mathbb{IR}^{m\times n}$, for each Z satisfying $|Z| = E$ (i.e., a ± 1-matrix in $\mathbb{R}^{m\times n}$) define

$$A_Z = A_c + Z \circ \Delta,$$

where "$\circ$" denotes the Hadamard (entrywise) product. It is obvious that $(A_Z)_{ij} = \overline{A}_{ij}$ if $Z_{ij} = 1$ and $(A_Z)_{ij} = \underline{A}_{ij}$ if $Z_{ij} = -1$ for each i, j, so that $A_Z \in \boldsymbol{A}$. If we view $\boldsymbol{A}$ as a rectangle in $\mathbb{R}^{mn}$, then the matrices A_Z, $|Z| = E$ are exactly the vertices of this rectangle; this is why they are called *vertex matrices*. Notice that if $\Delta > 0$, then there are exactly 2^{mn} vertex matrices. In the next theorem we show that any induced interval matrix norm can be expressed (and, if mn is small, also computed) via vertex matrices.

Theorem 5. *If an interval matrix norm $|||\cdot|||$ is induced by a point matrix norm $\|\cdot\|$, then*

$$|||\boldsymbol{A}||| = \max_{|Z|=E} \|A_c + Z \circ \Delta\| \tag{5.1}$$

for each $\boldsymbol{A} \in \mathbb{IR}^{m\times n}$.

Proof. If we again view $\boldsymbol{A}$ as a rectangle in $\mathbb{R}^{mn}$, then we can use the fact that each point of this rectangle can be expressed as a convex combination of its vertices. Hence each $A \in \boldsymbol{A}$ can be written in the form

$$A = \sum_{|Z|=E} \lambda_Z A_Z,$$

where all the λ_Z's are nonnegative and satisfy

$$\sum_{|Z|=E} \lambda_Z = 1.$$

Now we have

$$\|A\| \le \sum_{|Z|=E} \lambda_Z \|A_Z\| \le (\max_{|Z|=E} \|A_Z\|) \sum_{|Z|=E} \lambda_Z = \max_{|Z|=E} \|A_Z\|,$$

so that, by definition of the induced interval matrix norm,

$$|||\boldsymbol{A}||| \le \max_{|Z|=E} \|A_Z\|.$$

On the other hand, $A_Z \in \boldsymbol{A}$ for each Z which gives that $\|A_Z\| \le |||\boldsymbol{A}|||$, so that

$$\max_{|Z|=E} \|A_Z\| \le |||\boldsymbol{A}|||.$$

Consequently

$$|||\boldsymbol{A}||| = \max_{|Z|=E} \|A_Z\| = \max_{|Z|=E} \|A_c + Z \circ \Delta\|$$

as claimed. □

The formula (5.1) requires norms of up to 2^{mn} point matrices to be computed. In the next two sections we shall show that this number can be reduced for particular norms.

6 Computing the norms II: the case of $|||\cdot|||_2$

As is well known [6], the matrix norm $\|A\|_2$ is defined by

$$\|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2$$

where $\|x\|_2 = \sqrt{x^T x}$, and satisfies

$$\|A\|_2 = \sqrt{\varrho(A^T A)} = \sigma_{\max}(A),$$

which is why it is called the *spectral norm*. In this section we derive a formula for computing the induced interval matrix norm $|||\boldsymbol{A}|||_2$. To this end we shall need yet another formula for $\|A\|_2$. This formula appears in Golub and van Loan [5], Horn and Johnson [7], and Stewart and Sun [9] in the guise of a formula for $\sigma_{\max}(A)$, everywhere without proof. So we supply a proof here for the sake of completeness.

Theorem 6. *For each matrix $A \in \mathbb{R}^{m\times n}$ we have*

$$\|A\|_2 = \max_{\|x_1\|_2=\|x_2\|_2=1} x_1^T A x_2. \tag{6.1}$$

Proof. Let $A = XSY^T$ be a singular value decomposition of A, $q = \min(m, n)$, and let $x_1 \in \mathbb{R}^m$, $x_2 \in \mathbb{R}^n$ be vectors with $\|x_1\|_2 = 1$ and $\|x_2\|_2 = 1$. Put $x_1' = X^T x_1$, $x_2' = Y^T x_2$, then $\|x_1'\|_2 = 1$ and $\|x_2'\|_2 = 1$ since X and Y are orthogonal matrices, and let $x_1'' = ((x_1')_i)_{i=1}^q$, $x_2'' = ((x_2')_i)_{i=1}^q$. Then

$$\begin{aligned} x_1^T A x_2 &= x_1'^T S x_2' = \sum_{i=1}^q S_{ii}(x_1')_i(x_2')_i \leq S_{11} \sum_{i=1}^q |(x_1')_i||(x_2')_i| = \sigma_{\max}(A)|x_1''|^T|x_2''| \\ &\leq \|A\|_2\|x_1''\|_2\|x_2''\|_2 \leq \|A\|_2\|x_1'\|_2\|x_2'\|_2 = \|A\|_2 \end{aligned}$$

where we have used the Cauchy-Schwarz inequality. Hence,

$$\max_{\|x_1\|_2=\|x_2\|_2=1} x_1^T A x_2 \leq \|A\|_2. \tag{6.2}$$

On the other hand, if we put $x_1''' = X e_m^{(1)}$, $x_2''' = Y e_n^{(1)}$ where $e_m^{(1)}$ is the first column of the $m \times m$ identity matrix and $e_n^{(1)}$ is the first column of the $n \times n$ identity matrix, then again $\|x_1'''\|_2 = \|x_2'''\|_2 = 1$ and

$$x_1'''^T A x_2''' = e_m^{(1)T} S e_n^{(1)} = S_{11} = \|A\|_2,$$

hence the upper bound in (6.2) is attained which proves (6.1). □

The following theorem was proved (again in the guise of $\sigma_{\max}$ on both sides of (6.3)) by Ahn and Chen [1]. We give here another proof based on Theorem 6. We use the notation $e_m = (1, 1, \ldots, 1)^T \in \mathbb{R}^m$, $e_n = (1, 1, \ldots, 1)^T \in \mathbb{R}^n$.

Theorem 7. *For each* $\boldsymbol{A} \in \mathbb{IR}^{m \times n}$ *we have*

$$\|\boldsymbol{A}\|_2 = \max_{|y|=e_m,\,|z|=e_n} \|A_c + (yz^T) \circ \Delta\|_2. \tag{6.3}$$

Proof. Let $A \in \boldsymbol{A}$, and let $x_1 \in \mathbb{R}^m$, $x_2 \in \mathbb{R}^n$ with $\|x_1\|_2 = \|x_2\|_2 = 1$. Then

$$x_1^T A x_2 = x_1^T A_c x_2 + x_1^T (A - A_c) x_2 \leq x_1^T A_c x_2 + |x_1|^T \Delta |x_2|. \tag{6.4}$$

Define y by $y_i = 1$ if $(x_1)_i \geq 0$ and $y_i = -1$ otherwise ($i = 1, \ldots, m$) and similarly z by $z_j = 1$ if $(x_2)_i \geq 0$ and $z_j = -1$ otherwise ($j = 1, \ldots, n$), then $|y| = e_m$, $|z| = e_n$, and $|x_1| = \mathrm{diag}(y)x_1$, $|x_2| = \mathrm{diag}(z)x_2$, hence

$$x_1^T A_c x_2 + |x_1|^T \Delta |x_2| = x_1^T (A_c + \mathrm{diag}(y)\Delta\mathrm{diag}(z)) x_2. \tag{6.5}$$

Now,

$$(\mathrm{diag}(y)\Delta\mathrm{diag}(z))_{ij} = y_i z_j \Delta_{ij} = ((yz^T) \circ \Delta)_{ij}$$

for each i, j, hence $\mathrm{diag}(y)\Delta\mathrm{diag}(z) = (yz^T) \circ \Delta$ and from (6.4), (6.5) and Theorem 6 we obtain

$$x_1^T A x_2 \leq x_1^T (A_c + (yz^T) \circ \Delta) x_2 \leq \|A_c + (yz^T) \circ \Delta\|_2 \leq \max_{|y|=e_m,\,|z|=e_n} \|A_c + (yz^T) \circ \Delta\|_2.$$

Again applying Theorem 6, we get

$$\|A\|_2 \le \max_{|y|=e_m,\,|z|=e_n} \|A_c + (yz^T) \circ \Delta\|_2$$

for each $A \in \boldsymbol{A}$, hence by definition of the induced interval matrix norm we finally have that

$$|||\boldsymbol{A}|||_2 \le \max_{|y|=e_m,\,|z|=e_n} \|A_c + (yz^T) \circ \Delta\|_2. \tag{6.6}$$

On the other hand, each matrix of the form $A_c + (yz^T) \circ \Delta$, $|y| = e_m$, $|z| = e_n$ belongs to $\boldsymbol{A}$, hence

$$\max_{|y|=e_m,\,|z|=e_n} \|A_c + (yz^T) \circ \Delta\|_2 \le |||\boldsymbol{A}|||_2 \tag{6.7}$$

and (6.6), (6.7) finally yield

$$|||\boldsymbol{A}|||_2 = \max_{|y|=e_m,\,|z|=e_n} \|A_c + (yz^T) \circ \Delta\|_2$$

which concludes the proof. □

Since $|y| = e_m$ and $|z| = e_n$, yz^T is a ± 1-matrix and the formula (6.3) is of the same form as (5.1), but it requires computation of "only" 2^{m+n} point matrix norms compared to 2^{mn} of them in (5.1). Nevertheless, the number remains exponential. This leads us to conjecture that computation of $|||\cdot|||_2$ might be NP-hard.

7 Computing the norms III: absolute norms

Beside the 2-norm, the following four norms are used almost exclusively both in theory and practice:

$$\begin{aligned} \|A\|_1 &= \max_j \sum_i |a_{ij}|, \\ \|A\|_\infty &= \max_i \sum_j |a_{ij}|, \\ \|A\|_{1,\infty} &= \max_{ij} |a_{ij}|, \\ \|A\|_F &= \sqrt{\sum_{ij} a_{ij}^2}. \end{aligned}$$

Of them, only the third norm is not consistent, but can be made such by premultiplying by n (Higham [6]). All four of them share a simple, yet for us decisive property:

$$\|\,|A|\,\| = \|A\| \tag{7.1}$$

for each $A \in \mathbb{R}^{m \times n}$ (note: on the left-hand side of (7.1) there stands the norm of the absolute value of A, not an induced norm). Norms satisfying (7.1) are called *absolute*, and they possess the following property [7].

Theorem 8. *A norm $\|\cdot\|$ in $\mathbb{R}^{m\times n}$ is absolute if and only if for each $A, B \in \mathbb{R}^{m\times n}$, $|A| \le |B|$ implies $\|A\| \le \|B\|$.*

Let us define the sign matrix of A_c by $(\text{sgn}(A_c))_{ij} = 1$ if $(A_c)_{ij} \ge 0$ and $(\text{sgn}(A_c))_{ij} = -1$ otherwise, so that $\text{sgn}(A_c)$ is a ± 1-matrix. Now we can formulate a theorem which we consider the main contribution of this paper.

Theorem 9. *If a point matrix norm $\|\cdot\|$ in $\mathbb{R}^{m\times n}$ is absolute, then the induced interval matrix norm $|||\cdot|||$ satisfies*

$$|||\boldsymbol{A}||| = \|A_c + \text{sgn}(A_c) \circ \Delta\| = \|\,|A_c| + \Delta\,\| \tag{7.2}$$

for each $\boldsymbol{A} \in \mathbb{IR}^{m\times n}$.

Proof. For each $A \in \boldsymbol{A}$ we have

$$|A| = |A_c + A - A_c| \le |A_c| + \Delta = |\,|A_c| + \Delta\,|,$$

hence, by Theorem 8,

$$\|A\| = \|\,|A|\,\| \le \|\,|\,|A_c| + \Delta\,|\,\| = \|\,|A_c| + \Delta\,\|$$

and

$$|||\boldsymbol{A}||| = \max_{A\in\boldsymbol{A}} \|A\| \le \|\,|A_c| + \Delta\,\|. \tag{7.3}$$

Now,

$$|A_c| + \Delta = \text{sgn}(A_c) \circ A_c + \Delta = \text{sgn}(A_c) \circ (A_c + \text{sgn}(A_c) \circ \Delta)$$

hence

$$|A_c| + \Delta = |\,|A_c| + \Delta\,| = |\text{sgn}(A_c) \circ (A_c + \text{sgn}(A_c) \circ \Delta)| = |A_c + \text{sgn}(A_c) \circ \Delta|$$

and

$$\|\,|A_c| + \Delta\,\| = \|\,|A_c + \text{sgn}(A_c) \circ \Delta|\,\| = \|A_c + \text{sgn}(A_c) \circ \Delta\|$$

where $A_c + \text{sgn}(A_c) \circ \Delta \in \boldsymbol{A}$, hence the upper bound in (7.3) is attained, which gives (7.2). □

Thus, in (7.2) the right-hand side term gives a simple expression of the result whereas the middle term yields an explicit form of the matrix in $\boldsymbol{A}$ at which $|||\boldsymbol{A}|||$ is attained. Notice that the matrix is again of the form $A_c + Z \circ \Delta$, where Z is a ± 1-matrix, but this time norm of only *one* matrix is to be computed compared with 2^{m+n} of them for the norm $|||\cdot|||_2$.

To be perfectly clear, we give here explicit results for the above-quoted four absolute norms. This is a direct consequence of Theorem 9.

Theorem 10. *For each $\boldsymbol{A} = [A_c - \Delta, A_c + \Delta] \in \mathbb{IR}^{m\times n}$ we have*

$$\begin{aligned} |||\boldsymbol{A}|||_1 &= \|\,|A_c| + \Delta\,\|_1, \\ |||\boldsymbol{A}|||_\infty &= \|\,|A_c| + \Delta\,\|_\infty, \\ |||\boldsymbol{A}|||_{1,\infty} &= \|\,|A_c| + \Delta\,\|_{1,\infty}, \\ |||\boldsymbol{A}|||_F &= \|\,|A_c| + \Delta\,\|_F. \end{aligned}$$

8 Conclusion

We have introduced interval matrix norms and presented formulae for computing those induced by the five most frequently used point matrix norms. This is certainly not all what can be said of the subject. In particular, two problems have remained unsolved:

1. Construct an interval matrix norm which is not induced by any point matrix norm.
2. Prove that computing $|||\cdot|||_2$ is NP-hard (if true).

This can become a subject of future research.

Bibliography

[1] H.-S. Ahn and Y. Chen, *Exact maximum singular value calculation of an interval matrix*, IEEE Transactions on Automatic Control, 52 (2007), pp. 510–514. 8

[2] G. Alefeld and J. Herzberger, *Introduction to Interval Computations*, Academic Press, New York, 1983. 3, 4

[3] W. Barth and E. Nuding, *Optimale Lösung von Intervallgleichungssystemen*, Computing, 12 (1974), pp. 117–125. 1

[4] S. H. Friedberg, A. J. Insel, and L. E. Spence, *Linear Algebra*, Prentice Hall, Upper Saddle River, 1997. 3

[5] G. H. Golub and C. F. van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, 1996. 7

[6] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, SIAM, Philadelphia, 1996. 7, 9

[7] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985. 5, 7, 9

[8] A. Neumaier, *Interval Methods for Systems of Equations*, Cambridge University Press, Cambridge, 1990. 3

[9] G. W. Stewart and J. Sun, *Matrix Perturbation Theory*, Academic Press, San Diego, 1990. 7