

Two Submatrix Properties of Certain Induced Norms

Charles R. Johnson*

Institute for Basic Standards, National Bureau of Standards, Washington, D.C. 20234

(May 7, 1975)

Induced matrix norms N , defined consistently for all orders up to a given order, which have the property:

$$N(\hat{A}) \leq N(A)$$

for all principal submatrices \hat{A} of an arbitrary $m \times m$ complex matrix A are characterized. Conditions are also given which insure that

$$N(A \oplus B) = \max \{N(A), N(B)\}$$

Key words: Consistency; matrix norm; monotonicity; principal submatrix; spectral radius, vector norm.

It is well known that for $A \in M_k(C)$, the $k \times k$ complex matrices,

$$\rho(A) \leq \|A\| \quad (1)$$

where $\rho(A)$ is the *spectral radius* of A and $\|\cdot\|$ is any vector-norm-induced *matrix norm*. However, if \hat{A} is a principal submatrix of A , it is *not* in general the case that

$$\rho(\hat{A}) \leq \rho(A), \quad (2)$$

nor is it the case that

$$\rho(\hat{A}) \leq \|A\|. \quad (3)$$

EXAMPLE 1: Let $\|\cdot\|_\infty$ be the infinity vector norm on C^2 , $\|x\|_\infty \equiv \max \{|x_1|, |x_2|\}$ and define

$$\|x\|_S \equiv \|Sx\|_\infty$$

where

$$S = \begin{pmatrix} 1 & 0 \\ 2 & 2 \end{pmatrix}.$$

Then the vector norm $\|\cdot\|_S$ induces a matrix norm $\|\cdot\|_S$ on $M_2(C)$ which is related to the matrix norm

$$\left\| \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right\|_\infty \equiv \max \{|a| + |b|, |c| + |d|\}$$

by

$$\|A\|_S = \|SAS^{-1}\|_\infty.$$

AMS Subject Classification: 15A60, 15A42, 65F35.

* Present address: Institute for Fluid Dynamics and Applied Mathematics, University of Maryland, College Park, Md. 20742.

Now, if

$$A = \begin{pmatrix} a & a \\ -a & -a \end{pmatrix},$$

where $a > 0$ then

$$\rho(A) = 0$$

while

$$\rho(\hat{A}) = a$$

where \hat{A} is either principal submatrix of A . However,

$$\|A\|_s = \frac{1}{2}a$$

so that

$$\rho(\hat{A}) \not\leq \rho(A) \text{ and } \rho(\hat{A}) \not\leq \|A\|_s.$$

Also, the spectral radius enjoys the property that

$$\rho(A \oplus B) = \max \{ \rho(A), \rho(B) \} \quad (4)$$

whereas *not* all matrix norms $\|\cdot\|$ satisfy

$$\|A \oplus B\| = \max \{ \|A\|, \|B\| \}. \quad (5)$$

For example, the Euclidean matrix norm $\|A\| = (\sum |a_{ij}|^2)^{1/2}$ fails condition (5). However, the matrix norms induced by the p -norms

$$\|x\|_p = \left(\sum_{i=1}^k |x_i|^p \right)^{1/p} \quad (6)$$

on vectors do satisfy (3) and (5). It is the goal of this note to point out general characteristics of classes of induced matrix norms which insure the validity of (3) or (5).

By a *vector norm* we mean a map $\|\cdot\|: C^k \rightarrow R$, the real numbers, which satisfies for all $x, y \in C^k$

$$\|x\| \geq 0, \text{ and } = 0 \iff x = 0; \quad (7)$$

$$\|cx\| = |c| \|x\| \quad \text{for any complex number } c; \quad (8)$$

and

$$\|x+y\| \leq \|x\| + \|y\|. \quad (9)$$

By a *matrix norm* we mean a map $\|\cdot\|: M_k(C) \rightarrow R$, which satisfies for all $A, B \in M_k(C)$

$$\|A\| \geq 0, \text{ and } = 0 \iff A = 0; \quad (10)$$

$$\|cA\| = |c| \|A\| \quad \text{for any complex number } c; \quad (11)$$

$$\|A+B\| \leq \|A\| + \|B\| \quad (12)$$

$$\|AB\| \leq \|A\| \|B\|. \quad (13)$$

Any vector norm $\|\cdot\|$ on C^k uniquely induces a matrix norm $\|\cdot\|$ on $M_k(C)$ by

$$\|A\| \equiv \max_{\|x\|=1} \|Ax\| \quad (14)$$

Finally, because of the homogeneity (8) it is clear that a vector norm $\|\cdot\|$ is precisely determined by its *unit ball*

$$B_{\|\cdot\|} \equiv \{x \in C^k : \|x\| \leq 1\} \quad (15)$$

which is well known to be a convex, compact set containing the origin.

Since we compare the norm of a matrix to norms of principal submatrices, we need to consider classes of norms defined for all orders less than or equal to a given order.

Let

$$n \equiv \{n_k, k=1, \dots, m : n_k \text{ is a vector norm on } C^k\} \quad (16)$$

and if $x \in C^k$, $k \leq m$, we abbreviate $n_k(x)$ by $n(x)$. Also let N_k be the matrix norm which n_k induces on $M_k(C)$ via (14) and

$$N \equiv \{N_k : k=1, \dots, m\}. \quad (17)$$

If $A \in M_k(C)$, $k \leq m$, we abbreviate $N_k(A)$ by $N(A)$.

DEFINITION 1: We say that n is consistent if for each $k=2, \dots, m$ and each $j=1, \dots, k$

$$n_k(x_1, \dots, x_{j-1}, 0, x_{j+1}, \dots, x_k) = n_{k-1}(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_k). \quad (18)$$

We further say that n is strongly consistent if for all $k=2, \dots, m$ and all $x^1 \in C^{k_1}$, $\sum_{i=1}^l k_i = k$, we have

$$n_k(x^1, \dots, x^l) \equiv n_l(\|x^1\|, \dots, \|x^l\|). \quad (19)$$

REMARK 1: The norm set n is consistent if and only if B_{n_k} is the intersection of any k -dimensional coordinate plane with $B_{n_{k+1}}$, $k=1, \dots, m-1$.

DEFINITION 2: We say that n_k is monotone if for $x = (x_1, \dots, x_k)$ and all $j=1, \dots, k$

$$n_k(x_1, \dots, x_{j-1}, 0, x_{j+1}, \dots, x_k) \leq n_k(x_1, \dots, x_k). \quad (20)$$

We say that n is monotone if and only if n_k is monotone, $k=2, \dots, m$.

REMARK 2: From the convexity of B_{n_k} it follows that if n_k is monotone then

$$n_k(x_1, \dots, x_{j-1}, \alpha x_j, x_{j+1}, \dots, x_k) \leq n_k(x_1, \dots, x_k) \quad (21)$$

for $0 \leq \alpha \leq 1$.

REMARK 3: Geometrically, it is clear that n_k is monotone if and only if B_{n_k} satisfies " $x \in$ boundary (B_{n_k}) implies that the projection of x onto any $k-1$ dimensional coordinate plane is in B_{n_k} ". Moreover, if n is consistent, then n is monotone if and only if $x \in$ boundary (B_{n_m}) implies that the projection of x onto any k dimensional coordinate plane, $k < m$, is in B_{n_m} .

EXAMPLE 2: Each of the p -norms defined by (6) is consistent, strongly consistent and monotone.

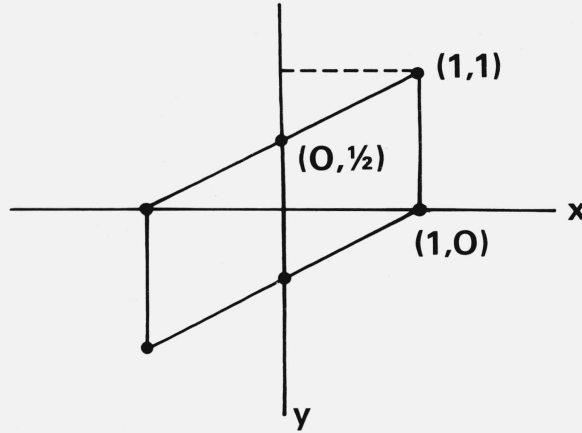
EXAMPLE 3: If $\|\cdot\|_1$ is the p -norm with $p=1$ in (6) on C^2 and we define

$$\|(x, y)\|_T = \|T \begin{pmatrix} x \\ y \end{pmatrix}\|_1$$

where

$$T = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix},$$

then $B_{\|\cdot\|_T}$ is



In this case $\|(0, 1)\|_T = 2 \not\leq 1 = \|(1, 1)\|_T$, and $\|\cdot\|_T$ is *not* monotone. Visually, the projection of $(1, 1)$ from the boundary of $B_{\|\cdot\|_T}$ onto the y axis leaves $B_{\|\cdot\|_T}$.

Our first main result concerns norms of principal submatrices.

THEOREM 1: *If n is consistent and \hat{A} denotes any principal submatrix, we have:*

$$N(\hat{A}) \leq N(A) \quad (22)$$

for all $A \in M_k(C)$, $2 \leq k \leq m$, if and only if n is monotonic.

PROOF: First of all, suppose n is consistent and monotonic. Because of the consistency it suffices to suppose $\hat{A} = A_1$, that principal submatrix determined by deleting the first row and column of A . (The argument for any other case would be entirely analogous.) We then have

$$N(A_1) = \max_{n(x)=1} n(A_1 x) \quad \text{by (14)}$$

$$= \max_{n(0, x)=1} n \left(\begin{pmatrix} 0 & 0 \\ 0 & A_1 \end{pmatrix} \begin{pmatrix} 0 \\ x \end{pmatrix} \right) \quad \text{by (18)}$$

$$= \max_{n(y)=1} n \left(\begin{pmatrix} 0 & 0 \\ 0 & A_1 \end{pmatrix} y \right) \quad \text{by (20)}$$

$$\begin{aligned} &= N \left(\begin{pmatrix} 0 & 0 \\ 0 & A_1 \end{pmatrix} \right) = N \left(\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} A \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \right) \\ &\leq N \left(\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \right) N(A) N \left(\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \right) \quad \text{by (13)} \end{aligned}$$

$$= 1 N(A) 1 \quad \text{by (20)}$$

$$= N(A),$$

which completes the first portion of the proof.

Conversely, suppose that n is consistent and that $N(\hat{A}) \leq N(A)$ for all $A \in M_k(C)$, $2 \leq k \leq m$. This implies that

$$N\left(\begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \end{pmatrix}\right) = 1 \quad (23)$$

since

$$1 \leq N\left(\begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \end{pmatrix}\right) \leq N(I) = 1$$

which by definition means that

$$\frac{n(x_1, \dots, x_{j-1}, 0, x_{j+1}, \dots, x_k)}{n(x)} \leq 1 \quad (24)$$

or, equivalently, that n is monotonic (20). This completes the proof.

COROLLARY 1: If n is consistent and monotonic, then for any $A \in M_m(C)$

$$\rho(\hat{A}) \leq N(A) \quad (25)$$

for all principal submatrices \hat{A} of A .

PROOF: That $\rho(\hat{A}) \leq N(\hat{A})$ is always the case if N is induced. From theorem 1, $N(\hat{A}) \leq N(A)$.

THEOREM 2: If n is strongly consistent and monotone, and $A \in M_{m_1}(C)$, $B \in M_{m_2}(C)$, $m_1 + m_2 = m$, then

$$N(A \oplus B) = \max \{N(A), N(B)\}. \quad (26)$$

PROOF: Suppose $x \in C^{m_1}$ and $y \in C^{m_2}$ satisfy $n(x) = n(y) = 1$ and $n(Ax) = \alpha$, $n(By) = \beta$. Then $tx \oplus (1-t)y \in C^m$ and $(A \oplus B)(tx \oplus (1-t)y) = tAx \oplus (1-t)By$ so that

$$\frac{n((A \oplus B)(tx \oplus (1-t)y))}{n(tx \oplus (1-t)y)} = \frac{n((t\alpha, (1-t)\beta))}{n((t, (1-t)))} \quad (27)$$

since n is strongly consistent. Since n is monotone, it follows from remark 2 that, for $0 \leq t \leq 1$, the maximum attained by the right-hand side of (27) is the greater of α and β (when either $t = 1$ or $t = 0$). We may thus conclude that

$$\max_{n(z)=1} (A \oplus B)z \equiv N(A \oplus B)$$

is attained when z is of the form either

$$z = z^1 \oplus 0 \text{ or } z = 0 \oplus z^2,$$

$z^1 \in C^{m_1}$, $z^2 \in C^{m_2}$. Then $N(A \oplus B)$ is either $N(A)$ or $N(B)$, whichever is larger, as was to be shown.

By an entirely similar approach, Theorem 2 may be generalized to:

COROLLARY 2: If n is strongly consistent and monotone, and $A_i \in M_{m_i}(C)$, $i = 1, \dots, l$ where $m_1 + \dots + m_l = m$, then

$$N(A_1 \oplus \dots \oplus A_l) = \max_{1 \leq i \leq l} N(A_i). \quad (28)$$

As noted previously, the p -norms (6) are consistent, strongly consistent and monotone. It would be of further interest to determine which vector norms of the type (16) are consistent, strongly consistent and/or monotone.

Author's Note (added in proof):

Theorem 2 of our note should be compared with a result of a similar nature (theorem 2) contained in "Norms on Direct Sums and Tensor Products," Math. Comp. **26**, 401–414 (1972). In fact it follows from the present work together with that paper (whose definitions we follow for new terms) that the following conditions on a consistent vector norm n are equivalent:

- (i) n is monotone;
- (ii) n is absolute;
- (iii) the matrix norm N induced by n satisfies (22); and
- (iv) the matrix norm N induced by n satisfies (26).

(Paper 79B3&4–424)