On a Connection between Infima of Norms and Eigenvalues of Associated Operators*

RICHARD S. VARGA
Kent State University
Kent, Ohio

1. INTRODUCTION

If $\mathcal{F}_1$ is the family of all vector norms on $\mathbb{C}^n$, i.e.,

$$\mathcal{F}_1 \equiv \{ \phi: \phi \text{ is any vector norm on } \mathbb{C}^n \},$$

(1.1)

and if $A: \mathbb{C}^n \to \mathbb{C}^n$ is any (bounded) linear transformation, i.e., $A \in [\mathbb{C}^n]$, then it is very well known (cf. Householder [1, p. 46]) that

$$\inf\{\|A\|_\phi: \phi \in \mathcal{F}_1\} = \rho(A),$$

(1.2)

where $\|A\|_\phi \equiv \sup_{\|x\|_\phi \leq 1} \phi(Ax)$ denotes the operator norm of $A$ with respect to the vector norm $\phi$, and

$$\rho(A) \equiv \max\{|\lambda_i|: \lambda_i \text{ is an eigenvalue of } A\}$$

denotes the spectral radius of $A$. On the other hand, if, for the canonical basis $\{e_i\}_{i=1}^n$ in $\mathbb{C}^n$, $\mathcal{F}_2$ is the particular family of vector norms on $\mathbb{C}^n$ of the form

$$\mathcal{F}_2 \equiv \left\{ \phi: \text{ there exist positive real numbers } \phi_1, \phi_2, \ldots, \phi_n \text{ such that for all } x = \sum_{i=1}^n c_i e_i \text{ in } \mathbb{C}^n, \phi(x) = \max_{1 \leq i \leq n} \{|c_i|/\phi_i\} \right\},$$

(1.3)

it is less well known [cf. Eq. (5.3)] that

$$\inf\{\|A\|_\phi: \phi \in \mathcal{F}_2\} = \rho(A),$$

(1.4)

* Research supported in part by the Atomic Energy Commission under Grant AT(11-1)-2075.

where if $A = (a_{i,j})$ is the matrix representation of $A \in \mathbb{C}^n$, relative to the canonical basis $\{e_j\}_{j=1}^n$, then $|A| \in \mathbb{C}^n$ is defined analogously by $|A| = (|a_{i,j}|)$. The expression in Eq. (1.4) is in fact an interesting consequence of the Perron-Frobenius theory of nonnegative matrices.

The expressions of Eqs. (1.2) and (1.4) are similar in that each, loosely speaking, states that the infimum of the norm of a fixed element $A$ in $\mathbb{C}^n$ over a family of norms is related either to the eigenvalues of $A$, or to the eigenvalues of a $B \in \mathbb{C}^n$, in some way associated with $A$. One aim here is to obtain a generalization of Eqs. (1.2) and (1.4) for arbitrary families $\mathcal{F}$ of norms. In so doing, we are naturally led to problems which connect with the theory of invariant cones and with the Perron-Frobenius theory of nonnegative matrices.

2. A GENERALIZATION

Let $\mathcal{F}$ be any fixed (finite or infinite) nonvoid family of vector norms on $\mathbb{C}^n$, $n \geq 1$, and let $A$ be a fixed element in $\mathbb{C}^n$. Relative to $\mathcal{F}$ and $A$, set

$$\hat{\Omega}_{\mathcal{F}}(A) \equiv \{B \in \mathbb{C}^n: ||B||_\phi \leq ||A||_\phi \quad \forall \phi \in \mathcal{F}\}. \quad (2.1)$$

This set is, of course, nonempty since it contains $\omega A$ for any complex number $\omega$ with $|\omega| \leq 1$.

We now prove

**Theorem 1.** For any arbitrary family of norms $\mathcal{F}$ on $\mathbb{C}^n$, and for an arbitrary $A \in \mathbb{C}^n$,

$$\inf\{||A||_\phi: \phi \in \mathcal{F}\} = \sup\{\rho(B): B \in \hat{\Omega}_{\mathcal{F}}(A)\}. \quad (2.2)$$

In particular, there is a $B \in \hat{\Omega}_{\mathcal{F}}(A)$ such that

$$\inf\{||A||_\phi: \phi \in \mathcal{F}\} = \rho(B). \quad (2.3)$$

**Proof.** Since it is classical (cf. Householder [1, p. 45]) that $||D||_\phi \geq \rho(D)$ for any $D \in \mathbb{C}^n$ and any vector norm $\phi$ on $\mathbb{C}^n$, it follows from Eq. (2.1) that

$$||A||_\phi \geq \rho(B), \quad \forall \phi \in \mathcal{F}, \quad \forall B \in \hat{\Omega}_{\mathcal{F}}(A),$$
from which it is evident that

$$\inf \{ \|A\|_\phi : \phi \in \mathcal{F} \} \geq \sup \{ \rho(B) : B \in \hat{\Omega}_\mathcal{F}(A) \}.$$ 

To show that equality holds in the above inequality, let $\eta$ be the nonnegative real number defined by $\eta \equiv \inf \{ \|A\|_\phi : \phi \in \mathcal{F} \}$, and consider $B \equiv \eta I$. By definition, $\|B\|_\phi = \eta \leq \|A\|_\phi$ for all $\phi \in \mathcal{F}$, so that $B \in \Omega_\mathcal{F}(A)$. But, as $\rho(B) = \eta$, the desired results of Eqs. (2.2) and (2.3) follow immediately. Q.E.D.

3. CHARACTERIZATION

With respect to the equality in Eq. (2.2), it is natural to ask if there is a $\phi \in \mathcal{F}$ such that

$$\|A\|_\phi = \inf \{ \|A\|_\phi : \phi \in \mathcal{F} \}. \quad (3.1)$$

The answer to this question is in general negative, as we shall see. However, what we seek now is a characterization of $A \in \mathcal{C}^n$ and a norm $\phi$ in $\mathcal{F}$ for which Eq. (3.1) is valid.

For notation, let $K_\phi$ denote the unit ball in $\mathcal{C}^n$ for the particular vector norm $\phi$ in $\mathcal{F}$, i.e., $K_\phi \equiv \{ x \in \mathcal{C}^n : \phi(x) \leq 1 \}$. Clearly, $K_\phi$ is a closed, bounded, convex, and balanced set with interior points in $\mathcal{C}^n$, for each $\phi \in \mathcal{F}$. For any $\phi \in \mathcal{F}$ and for any $A \in \mathcal{C}^n$, we have by definition that

$$B : K_\phi \to \|A\|_\phi K_\phi, \quad \forall B \in \hat{\Omega}_\mathcal{F}(A),$$

or equivalently, if $BK_\phi \equiv \{ y \in \mathcal{C}^n : y = Bx \text{ for some } x \in K_\phi \}$, then

$$BK_\phi \subseteq \|A\|_\phi K_\phi, \quad \forall B \in \hat{\Omega}_\mathcal{F}(A).$$

Since $\|A\|_\phi K_\phi$ is a balanced convex set in $\mathcal{C}^n$, it is then evident that the balanced convex hull of all $BK_\phi$, for $B$ in $\hat{\Omega}_\mathcal{F}(A)$, is in $\|A\|_\phi K_\phi$, i.e., if $\text{conv}\{\hat{\Omega}_\mathcal{F}(A)K_\phi\}$ denotes the smallest balanced convex set containing all $BK_\phi$ for $B$ in $\hat{\Omega}_\mathcal{F}(A)$, so that

$$\text{conv}\{\hat{\Omega}_\mathcal{F}(A)K_\phi\} = \left\{ \sum_{i=1}^m \lambda_i x_i : m \text{ finite, each } x_i \text{ is in some } BK_\phi \right\}.$$
with $B$ in $\hat{\Omega}_\mathcal{F}(A)$, and \( \sum_{i=1}^{n} |\lambda_i| \leq 1 \), \( i = 1 \).

then

\[
\text{conv}\{\hat{\Omega}_\mathcal{F}(A)K_\phi\} \subset \|A\|_\phi K_\phi, \quad \forall \phi \in \mathcal{F}.
\]

Next, for a fixed $\phi \in \mathcal{F}$ with unit ball $K_\phi$, we define the unit ball of $\psi \in \mathcal{F}$, relative to $\phi$, by

\[
K_\psi^\phi = \{x \in \mathbb{C}^n : \psi(x) \leq \alpha_\phi\},
\]

where $\alpha_\phi \equiv \sup\{\psi(x) : \phi(x) < 1\}$. Clearly, $K_\phi \subset K_\psi^\phi$ but $K_\phi \subset (1 - \varepsilon)K_\psi^\phi$ for any $0 < \varepsilon < 1$, and $K_\psi^\phi$ is just a multiple of the usual unit ball for $\psi$. It is clear from $K_\phi \subset K_\psi^\phi$ that

\[
BK_\phi \subset BK_\psi^\phi \subset \|A\|_\psi K_\psi^\phi, \quad \forall B \in \hat{\Omega}_\mathcal{F}(A),
\]

so that

\[
\text{conv}\{\hat{\Omega}_\mathcal{F}(A)K_\phi\} \subset \|A\|_\psi K_\psi^\phi, \quad \forall \psi \in \mathcal{F}.
\]

Hence, as this holds for all $\psi \in \mathcal{F}$, then

\[
\text{conv}\{\hat{\Omega}_\mathcal{F}(A)K_\phi\} \subset \bigcap_{\psi \in \mathcal{F}} \{\|A\|_\psi K_\psi^\phi\}.
\]

Since

\[
\bigcap_{\psi \in \mathcal{F}} \{\|A\|_\psi K_\psi^\phi\} \subset \|A\|_\phi K_\phi,
\]

then from the above inclusion,

\[
\text{conv}\{\hat{\Omega}_\mathcal{F}(A)K_\phi\} \subset \bigcap_{\psi \in \mathcal{F}} \{\|A\|_\psi K_\psi^\phi\} \subset \|A\|_\phi K_\phi.
\]

This is used in proving

**Theorem 2.** For an arbitrary family $\mathcal{F}$ of norms on $\mathbb{C}^n$ and for an arbitrary $A \in [\mathbb{C}^n]$, the following relations are equivalent:
(i) there is a $\phi \in \mathcal{F}$ such that
\[ \|A\|_\phi = \inf\{\|A\|_\psi : \psi \in \mathcal{F}\}; \quad (3.5) \]

(ii) there is a $\phi \in \mathcal{F}$ such that
\[ \text{conv}\{\hat{\mathcal{Q}}_{\mathcal{F}}(A)K_\phi\} = \|A\|_\phi K_\phi; \quad (3.6) \]

(iii) there is a $\phi \in \mathcal{F}$ such that $K_\phi \subset (1 - \varepsilon)K_\phi$ for any $0 < \varepsilon < 1$, and
\[ \|A\|_\phi K_\phi = \bigcap_{\phi \in \mathcal{F}} \{\|A\|_\psi K_\psi\}. \quad (3.7) \]

Proof. Assuming (i), then $\|A\|_\phi \leq \|A\|_\psi$, $\forall \psi \in \mathcal{F}$. If $\hat{B} \equiv \|A\|_\phi \cdot I$, then by definition, $\hat{B} \in \hat{\mathcal{Q}}_{\mathcal{F}}(A)$, and as such, $\hat{B}K_\phi \subset \text{conv}\{\hat{\mathcal{Q}}_{\mathcal{F}}(A)K_\phi\}$. Thus,
\[ \|A\|_\phi K_\phi \subset \text{conv}\{\hat{\mathcal{Q}}_{\mathcal{F}}(A)K_\phi\} \subset \|A\|_\phi K_\phi, \]
the last inclusion following from Eq. (3.4), i.e., $\|A\|_\phi K_\phi = \text{conv}\{\hat{\mathcal{Q}}_{\mathcal{F}}(A)K_\phi\}$, and (i) implies (ii). Assuming (ii), (iii) follows from Eq. (3.4). Next, assuming (iii), it follows from (3.7) that $\|A\|_\phi K_\phi \subset \|A\|_\phi K_\phi$, $\forall \psi \in \mathcal{F}$. But as $K_\phi \subset K_\phi$ and $K_\phi \subset (1 - \varepsilon)K_\phi$ for any $0 < \varepsilon < 1$, it is clear that $\|A\|_\phi \leq \|A\|_\phi$ for any $\psi \in \mathcal{F}$, i.e., $\|A\|_\phi = \inf\{\|A\|_\psi : \psi \in \mathcal{F}\}$, and (iii) implies (i).

Corollary. Given a family $\mathcal{F}$ of norms on $\mathbb{C}^n$, assume that $A \in [\mathbb{C}^n]$ satisfies one (and hence all) of the relations (3.5)–(3.7). Then, for any $B \in \hat{\mathcal{Q}}_{\mathcal{F}}(A)$ with $\rho(B) = \inf\{\|A\|_\psi : \psi \in \mathcal{F}\}$, each eigenvalue $\lambda$ of $B$ with $|\lambda| = \rho(B)$ is such that the number of linearly independent eigenvectors belonging to $\lambda$ is equal to the multiplicity of $\lambda$ for $B$, i.e., the Jordan block associated with $\lambda$ in the Jordan normal form of $B$ is diagonal.

Proof. First, we note from Theorem 1 that the set of matrices $B \in \hat{\mathcal{Q}}_{\mathcal{F}}(A)$ with $\rho(B) = \inf\{\|A\|_\psi : \psi \in \mathcal{F}\}$ is not empty. Next, as $\rho(B) \leq \|B\|_\phi$ and as $\|B\|_\phi \leq \|A\|_\phi$ because $B \in \hat{\mathcal{Q}}_{\mathcal{F}}(A)$, then the assumption that $\|A\|_\phi = \inf\{\|A\|_\psi : \psi \in \mathcal{F}\}$ gives us that $\rho(B) = \|B\|_\phi$ for every such $B \in \hat{\mathcal{Q}}_{\mathcal{F}}(A)$,
and the eigenvalue and eigenvector properties of the Corollary follow from a well known result of Householder (cf. [1, p. 47]). Q.E.D.

As a remark, consider the particular family $\mathcal{F}_1$ of norms of Eq. (1.1). For any $A \in [\mathbb{C}^n]$, we know from Eq. (1.2) that

$$\rho(A) = \inf\{|A|_\phi : \phi \in \mathcal{F}_1\}.$$ 

Suppose that $A$ has an eigenvalue $\lambda$ with $|\lambda| = \rho(A)$ for which the Jordan block associated with $\lambda$ in the Jordan normal form of $A$ is not diagonal. Then, as $A \in \hat{\Omega}_\mathcal{F}_1(A)$ does not satisfy the conclusions of the above corollary, we obtain the known result (cf. [2]) that

$$|A|_\phi > \inf\{|A|_\phi : \phi \in \mathcal{F}_1\} = \rho(A), \quad \forall \phi \in \mathcal{F}_1,$$

i.e., Eq. (3.1) cannot hold for any $\phi \in \mathcal{F}_1$.

4. CONNECTIONS WITH CONES

For an arbitrary family $\mathcal{F}$ of norms on $\mathbb{C}^n$, and an arbitrary but fixed $A \in [\mathbb{C}^n]$, choose any $B \in \hat{\Omega}_\mathcal{F}(A)$ for which

$$\rho(B) = \inf\{|A|_\phi : \phi \in \mathcal{F}\}.$$  \hspace{1cm} (4.1)

Next, by way of normalization, choose any eigenvalue $\lambda$ of $B$ with $|\lambda| = \rho(B)$, for which the Jordan block associated with $\lambda$ (in the Jordan normal form) is maximal in size, and rotate $\lambda$ into $\rho(B)$, i.e., if $Bx = \lambda x$ where $x \neq 0$ and $\exp(i\theta) \lambda = \rho(B)$, then $\hat{B}x = \rho(B)x$ where $\hat{B} \equiv \exp(i\theta) B$ is also an element of $\hat{\Omega}_\mathcal{F}(A)$. Thus, we are considering all $B \in \hat{\Omega}_\mathcal{F}(A)$ for which

(i) $\rho(B) = \inf\{|A|_\phi : \phi \in \mathcal{F}\}$ is an eigenvalue of $B$;

(ii) the maximal Jordan block associated with $\rho(B)$ in the Jordan normal form of $B$ is no smaller than the Jordan block of any eigenvalue $\nu$ of $B$ with $|\nu| = \rho(B)$.  \hspace{1cm} (4.2)

It is interesting to note that the particular matrix $\hat{B} = \eta \cdot I$ where $\eta \equiv \inf\{|A|_\phi : \phi \in \mathcal{F}\}$ trivially satisfies the conditions of (4.2), and is, in addition, a real matrix. Thus, for any real $B \in \hat{\Omega}_\mathcal{F}(A)$ satisfying Eq. (4.2),
there is (cf. Vandergraft [3, Theorem 3.1]) a real solid cone $\mathcal{K}_B$ in real Euclidean space $E^n$ for which $B: \mathcal{K}_B \rightarrow \mathcal{K}_B$, and consequently [3], $\mathcal{K}_B$ contains an eigenvector corresponding to $\rho(B)$. We state this as

**Theorem 3.** For an arbitrary family $\mathcal{F}$ of norms on $C^n$, and an arbitrary $A \in [C^n]$, there exist $B \in \hat{Q}_\mathcal{F}(A)$ satisfying Eq. (4.2), and hence, for each such real $B$, there is a real solid cone $\mathcal{K}_B$ in $E^n$ for which $B: \mathcal{K}_B \rightarrow \mathcal{K}_B$, and $\mathcal{K}_B$ contains an eigenvector of $B$ corresponding to $\rho(B)$.

It is interesting to note that if $A \in [C^n]$ maps a real solid cone $\mathcal{K}$ in $E^n$ into itself, i.e., $A: \mathcal{K} \rightarrow \mathcal{K}$, then it is known [3] that $\rho(A)$ is an eigenvalue of $A$, and that (4.2ii) holds with $B = A$. Next, it is always possible to choose a family $\mathcal{F}$ of norms on $[C^n]$ for which

$$\rho(A) = \inf\{\|A\|_{\phi}: \phi \in \mathcal{F}\}.$$ 

For $A$ and this family $\mathcal{F}$, (4.2) is evidently satisfied with $B = A$, and Theorem 3 regenerates the cone property of $A$.

5. An Example

To illustrate some of the above results, consider the particular family $\mathcal{F}_2$ of norms on $C^n$ given in Eq. (1.3). For $\psi \in \mathcal{F}_2$, we associate via Eq. (1.3) with $\psi$ a positive numbers $\psi_1, \psi_2, \ldots, \psi_n$, and write $\psi \sim (\psi_1, \psi_2, \ldots, \psi_n)$. For a fixed $A \in [C^n]$, it is easy to verify that

$$\|A\|_{\phi} = \max \left\{ \sum_{j=1}^n |a_{i,j}|(\psi_j/\psi_i) : 1 \leq i \leq n \right\},$$  

(5.1)

where $A = (a_{i,j})$ is the matrix representation of $A$. Note that the norm $\|A\|_{\phi}$ depends only on the moduli $|a_{i,j}|$ of $A = (a_{i,j})$. Because of this, if $|A| \equiv (|a_{i,j}|) \in [C^n]$, then $|A| \in \hat{Q}_{\mathcal{F}_2}(A)$, and we have as a well known consequence of the Perron-Frobenius theory of nonnegative matrices (see [4, p. 32] for the irreducible case) that

$$\rho(|A|) = \max \left\{ \max_{1 \leq i \leq n} \left[ \sum_{j=1}^n |a_{i,j}|(\psi_j/\psi_i) : \psi_i > 0, i = 1, 2, \ldots, n \right] \right\},$$  

(5.2)

from which it follows from Eq. (5.1) that [cf. Eq. (1.4)]

$$\rho(|A|) = \inf\{\|A\|_{\phi}: \psi \in \mathcal{F}_2\}.$$  

(5.3)
Moreover, it is also known from the Perron-Frobenius theory of non-negative matrices that if $|A|$ is irreducible, then there exist $n$ positive numbers $\phi_i, i = 1, 2, \ldots, n$, such that (cf. [4, p. 32])

$$\rho(|A|) = \sum_{i=1}^{n} |a_{ii}|(\phi_i/\phi_n) \quad \text{for all} \quad i = 1, 2, \ldots, n, \quad (5.4)$$

so that $\rho(|A|) = ||A||_\phi$ where $\phi \sim (\phi_1, \phi_2, \ldots, \phi_n) \in \mathbb{R}_+$. In this case, we have that

$$\rho(|A|) = ||A||_\phi = \inf\{||A||_\psi: \psi \in \mathbb{F}_2\}. \quad (5.5)$$

In this irreducible case, one can draw a stronger conclusion about $|A|$ than that given in the Corollary of Theorem 2, i.e., that $\rho(|A|)$ is a simple eigenvalue of $|A|$. However, it is easy to see that Eq. (5.5) can hold for certain reducible matrices, such as

$$|A| = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix},$$

for which $\rho(|A|)$ is not a simple eigenvalue of $|A|$.

REFERENCES


Received November 23, 1971