LINEAR ALGEBRA AND ITS APPLICATIONS 6, 249–256 (1973)

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

RICHARD S. VARGA
Kent State University
Kent, Ohio

1. INTRODUCTION

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

$$\mathscr{F}_1 \equiv \{\phi\colon \phi \text{ is any vector norm on } \mathbb{C}^n\}, \tag{1.1}$$

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

$$\inf\{\|A\|_\phi\colon \phi \in \mathscr{F}_1\} = \rho(A), \tag{1.2}$$

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

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

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

$$\mathscr{F}_2 \equiv \Big\{\phi\colon \text{there exist positive real numbers } \phi_1, \phi_2, \ldots, \phi_n \text{ such that for all } x = \sum_{i=1}^n c_i e_i \text{ in } \mathbb{C}^n,\ \ \phi(x) = \max_{1\leqslant i\leqslant n}\{|c_i|/\phi_i\}\Big\}, \tag{1.3}$$

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

$$\inf\{\|A\|_\phi\colon \phi \in \mathscr{F}_2\} = \rho(|A|), \tag{1.4}$$

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

© American Elsevier Publishing Company, Inc., 1973

where if $A = (a_{i,j})$ is the matrix representation of $A \in [\mathbf{C}^n]$, relative to the canonical basis $\{e_j\}_{j=1}^n$, then $|A| \in [\mathbf{C}^n]$ 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 $[\mathbf{C}^n]$ over a family of norms is related either to the eigenvalues of A, or to the eigenvalues of a $B \in [\mathbf{C}^n]$, in some way associated with A. One aim here is to obtain a generalization of Eqs. (1.2) and (1.4) for arbitrary families $\mathscr{F}$ 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 $\mathscr{F}$ be any fixed (finite or infinite) nonvoid family of vector norms on $\mathbf{C}^n$, $n \geqslant 1$, and let A be a fixed element in $[\mathbf{C}^n]$. Relative to $\mathscr{F}$ and A, set

$$\hat{\Omega}_{\mathscr{F}}(A) \equiv \{B \in [\mathbf{C}^n] : ||B||_\phi \leqslant ||A||_\phi \ \forall \phi \in \mathscr{F}\}. \tag{2.1}$$

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

We now prove

THEOREM 1. *For any arbitrary family of norms $\mathscr{F}$ on $\mathbf{C}^n$, and for an arbitrary $A \in [\mathbf{C}^n]$,*

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

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

$$\inf\{||A||_\phi : \phi \in \mathscr{F}\} = \rho(\hat{B}). \tag{2.3}$$

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

$$||A||_\phi \geqslant \rho(B), \qquad \forall \phi \in \mathscr{F}, \qquad \forall B \in \hat{\Omega}_{\mathscr{F}}(A),$$

from which it is evident that

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

To show that equality holds in the above inequality, let η be the nonnegative real number defined by $\eta \equiv \inf\{\|A\|_\phi : \phi \in \mathscr{F}\}$, and consider $\hat{B} \equiv \eta I$. By definition, $\|\hat{B}\|_\phi = \eta \leqslant \|A\|_\phi$ for all $\phi \in \mathscr{F}$, so that $\hat{B} \in \Omega_{\mathscr{F}}(A)$. But, as $\rho(\hat{B}) = \eta$, 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 $\phi \in \mathscr{F}$ such that

$$\|A\|_\phi = \inf\{\|A\|_\psi : \psi \in \mathscr{F}\}. \tag{3.1}$$

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

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

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

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

$$BK_\phi \subset \|A\|_\phi K_\phi, \qquad \forall B \in \hat{\Omega}_{\mathscr{F}}(A).$$

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

$$\operatorname{conv}\{\hat{\Omega}_{\mathscr{F}}(A)K_\phi\} = \left\{\sum_{i=1}^{m} \lambda_i x_i : m \text{ finite, each } x_i \text{ is in some } BK_\phi\right.$$

$$\text{with } B \text{ in } \hat{\Omega}_{\mathscr{F}}(A), \text{ and } \sum_{i=1}^{m} |\lambda_i| \leqslant 1 \Big\}, \tag{3.2}$$

then

$$\operatorname{conv}\{\hat{\Omega}_{\mathscr{F}}(A)K_\phi\} \subset ||A||_\phi K_\phi, \qquad \forall \phi \in \mathscr{F}. \tag{3.3}$$

Next, for a fixed $\phi \in \mathscr{F}$ with unit ball K_ϕ, we define the unit ball of $\psi \in \mathscr{F}$, *relative to* ϕ, by

$$K_\psi{}^\phi = \{x \in \mathbb{C}^n : \psi(x) \leqslant \alpha_\psi\},$$

where $\alpha_\psi \equiv \sup\{\psi(x) : \phi(x) \leqslant 1\}$. Clearly, $K_\phi \subset K_\psi{}^\phi$ but $K_\phi \not\subset (1-\varepsilon)K_\psi{}^\phi$ for any $0 < \varepsilon < 1$, and $K_\psi{}^\phi$ is just a multiple of the usual unit ball for ψ. It is clear from $K_\phi \subset K_\psi{}^\phi$ that

$$BK_\phi \subset BK_\psi{}^\phi \subset ||A||_\psi K_\psi{}^\phi, \qquad \forall B \in \hat{\Omega}_{\mathscr{F}}(A),$$

so that

$$\operatorname{conv}\{\hat{\Omega}_{\mathscr{F}}(A)K_\phi\} \subset ||A||_\psi K_\psi{}^\phi, \qquad \forall \psi \in \mathscr{F}.$$

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

$$\operatorname{conv}\{\hat{\Omega}_{\mathscr{F}}(A)K_\phi\} \subset \bigcap_{\psi \in \mathscr{F}} \{||A||_\psi K_\psi{}^\phi\}.$$

Since

$$\bigcap_{\psi \in \mathscr{F}} \{||A||_\psi K_\psi{}^\phi\} \subset ||A||_\phi K_\phi,$$

then from the above inclusion,

$$\operatorname{conv}\{\hat{\Omega}_{\mathscr{F}}(A)K_\phi\} \subset \bigcap_{\psi \in \mathscr{F}} \{||A||_\psi K_\psi{}^\phi\} \subset ||A||_\phi K_\phi. \tag{3.4}$$

This is used in proving

THEOREM 2. *For an arbitrary family $\mathscr{F}$ of norms on $\mathbb{C}^n$ and for an arbitrary $A \in [\mathbb{C}^n]$, the following relations are equivalent:*

(i) *there is a* $\phi \in \mathscr{F}$ *such that*

$$\|A\|_\phi = \inf\{\|A\|_\psi : \psi \in \mathscr{F}\}; \tag{3.5}$$

(ii) *there is a* $\phi \in \mathscr{F}$ *such that*

$$\operatorname{conv}\{\hat{\Omega}_{\mathscr{F}}(A)K_\phi\} = \|A\|_\phi K_\phi; \tag{3.6}$$

(iii) *there is a* $\phi \in \mathscr{F}$ *such that* $K_\phi \not\subset (1-\varepsilon)K_\psi{}^\phi$ *for any* $0 < \varepsilon < 1$, *and*

$$\|A\|_\phi K_\phi = \bigcap_{\psi \in \mathscr{F}} \{\|A\|_\psi K_\psi{}^\phi\}. \tag{3.7}$$

Proof. Assuming (i), then $\|A\|_\phi \leqslant \|A\|_\psi$, $\forall \psi \in \mathscr{F}$. If $\hat{B} \equiv \|A\|_\phi \cdot I$, then by definition, $\hat{B} \in \hat{\Omega}_{\mathscr{F}}(A)$, and as such, $\hat{B}K_\phi \subset \operatorname{conv}\{\hat{\Omega}_{\mathscr{F}}(A)K_\phi\}$. Thus,

$$\|A\|_\phi K_\phi \subset \operatorname{conv}\{\hat{\Omega}_{\mathscr{F}}(A)K_\phi\} \subset \|A\|_\phi K_\phi,$$

the last inclusion following from Eq. (3.4), i.e., $\|A\|_\phi K_\phi = \operatorname{conv}\{\hat{\Omega}_{\mathscr{F}}(A)K_\phi\}$, and (i) implies (ii). Assuming (ii), (iii) follows from Eq. (3.4). Next, assuming (iii), it follows from (3.7) that $\|A\|_\phi K_\phi \subset \|A\|_\psi K_\psi{}^\phi$, $\forall \psi \in \mathscr{F}$. But as $K_\phi \subset K_\psi{}^\phi$ and $K_\phi \not\subset (1-\varepsilon)K_\psi{}^\phi$ for any $0 < \varepsilon < 1$, it is clear that $\|A\|_\phi \leqslant \|A\|_\psi$ for any $\psi \in \mathscr{F}$, i.e., $\|A\|_\phi = \inf\{\|A\|_\psi : \psi \in \mathscr{F}\}$, and (iii) implies (i). Q.E.D.

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

Proof. First, we note from Theorem 1 that the set of matrices $B \in \hat{\Omega}_{\mathscr{F}}(A)$ with $\rho(B) = \inf\{\|A\|_\psi : \psi \in \mathscr{F}\}$ is not empty. Next, as $\rho(B) \leqslant \|B\|_\phi$ and as $\|B\|_\phi \leqslant \|A\|_\phi$ because $B \in \hat{\Omega}_{\mathscr{F}}(A)$, then the assumption that $\|A\|_\phi = \inf\{\|A\|_\psi : \psi \in \mathscr{F}\}$ gives us that $\rho(B) = \|B\|_\phi$ for every such $B \in \hat{\Omega}_{\mathscr{F}}(A)$, 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 $\mathscr{F}_1$ of norms of Eq. (1.1). For any $A \in [\mathbf{C}^n]$, we know from Eq. (1.2) that

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

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

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

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

4. CONNECTIONS WITH CONES

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

$$\rho(B) = \inf\{||A||_\phi : \phi \in \mathscr{F}\}. \tag{4.1}$$

Next, by way of normalization, choose any eigenvalue λ of B with $|\lambda| = \rho(B)$, for which the Jordan block associated with λ (in the Jordan normal form) is maximal in size, and rotate λ 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 $\hat{\Omega}_{\mathscr{F}}(A)$. Thus, we are considering all $B \in \hat{\Omega}_{\mathscr{F}}(A)$ for which

(i) $\rho(B) = \inf\{||A||_\phi : \phi \in \mathscr{F}\}$ 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 ν of B with $|\nu| = \rho(B)$. (4.2)

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

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

It is interesting to note that if $A \in [\mathbf{C}^n]$ maps a real solid cone $\mathscr{K}$ in E^n into itself, i.e., $A\colon \mathscr{K} \to \mathscr{K}$, 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 $\mathscr{F}$ of norms on $[\mathbf{C}^n]$ for which

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

For A and this family $\mathscr{F}$, (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 $\mathscr{F}_2$ of norms on $\mathbf{C}^n$ given in Eq. (1.3). For $\psi \in \mathscr{F}_2$, we associate via Eq. (1.3) with ψ n positive numbers $\psi_1, \psi_2, \ldots, \psi_n$, and write $\psi \sim (\psi_1, \psi_2, \ldots, \psi_n)$. For a fixed $A \in [\mathbf{C}^n]$, it is easy to verify that

$$\|A\|_\psi = \max\left\{\sum_{j=1}^{n} |a_{i,j}|(\psi_j/\psi_i) \colon 1 \leqslant i \leqslant n\right\}, \tag{5.1}$$

where $A = (a_{i,j})$ is the matrix representation of A. Note that the norm $\|A\|_\psi$ depends only on the moduli $|a_{i,j}|$ of $A = (a_{i,j})$. Because of this, if $|A| \equiv (|a_{i,j}|) \in [\mathbf{C}^n]$, then $|A| \in \hat{\Omega}_{\mathscr{F}_2}(A)$, 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

$$\rho(|A|) = \inf\left\{\max_{1 \leqslant i \leqslant n}\left[\sum_{j=1}^{n} |a_{i,j}|(\psi_j/\psi_i)\right] \colon \psi_i > 0,\ i = 1, 2, \ldots, n\right\}, \tag{5.2}$$

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

$$\rho(|A|) = \inf\{\|A\|_\psi \colon \psi \in \mathscr{F}_2\}. \tag{5.3}$$

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 ϕ_i, $i = 1, 2, \ldots, n$, such that (cf. [4, p. 32])

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

so that $\rho(|A|) = ||A||_\phi$ where $\phi \sim (\phi_1, \phi_2, \ldots, \phi_n) \in \mathscr{F}_2$. In this case, we have that

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

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

1 A. S. Householder, *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York (1964).
2 J. L. Mott and H. Schneider, Matrix algebras and groups relatively bounded in norm, *Archiv der Math.* **10**(1959), 1–6.
3 J. S. Vandergraft, Spectral properties of matrices which have invariant cones, *SIAM J. Appl. Math.* **16**(1968), 1208–1222.
4 R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, New Jersey (1962).

Received November 23, 1971