A STRUCTURE THEOREM FOR THE POLARS OF UNITARILY INVARIANT NORMS

GOVIND S. MUDHOLKAR AND MARSHALL FREIMER

ABSTRACT. The unitarily invariant norms of matrices, or operators, are essentially the symmetric norms of their singular values. A subclass of these norms depending upon only a few largest of the singular values is considered, and the polars of these norms are characterized. The result is then used to obtain generalizations of some well-known inequalities. The implications for operators on infinite-dimensional spaces are discussed.

1. Introduction and summary. The unitarily invariant (UI) norms were introduced by von Neumann [13] for the purpose of metrizing matrix spaces. He, together with his associates, established that the class of UI norms of $n \times n$ complex matrices coincides with the class of symmetric gauge functions (SGF's) of their singular values. These norms have now been variously generalized and utilized in several contexts. For example, Schatten [10, 11] defines norms on two-sided ideals of completely continuous operators on an arbitrary Hilbert space; Fan [2] obtains results relevant to these norms. The von Neumann norms play a crucial role in the study of function spaces and group representations (see, e.g., McCarthy [7] and Kunze [6]) and in obtaining certain bounds of importance in quantum field theory (see Simon [12]). More recently, Bhatia [1] has used the UI norms in the analysis of spectral variation. For historical perspectives and surveys of UI norms, see Schatten [10, 11], Hewitt and Ross [5], Gohberg and Krein [3] and Simon [12].

In his basic paper, von Neumann [13] studied various properties of SGF's and the corresponding UI norms. In particular, he showed that the polar of a UI norm is itself a UI norm. He illustrated the result using the usual p -norms. In this note we present a structure theorem for the polars of a class of UI norms. As an application of this result, we obtain the polars of Ky Fan's k -norms in explicit form and also an extension of Hölder's inequality. Some remarks on the spaces c_ϕ , analogous to the operator spaces c_p (see McCarthy [7]), related to these norms follow.

The basic facts relevant to the purpose of this paper are summarized in §2. §3 contains the main results. §4 is given to some applications and implications.

2. Preliminaries: SGF's and UI norms. Let H be a Hilbert space of finite dimension n . Let $U(H)$ be the group of all unitary operators in $B(H)$, the linear space of all bounded operators on H . A norm $\|\cdot\|$ on $B(H)$ is said to be UI, i.e.,

Received by the editors January 11, 1984 and, in revised form, October 2, 1984.

1980 *Mathematics Subject Classification*. Primary 15A60, 47A30; Secondary 15A39, 15A42, 46B10.

Key words and phrases. Symmetric gauge functions, Hölder's inequality, conjugates, c_p .

unitarily invariant, if for $A \in B(H)$ and $U, V \in U(H)$, $\|A\| = \|UAV\|$. Construction of such norms requires SGF's, i.e., symmetric gauge functions on R^n .

Following von Neumann, a gauge function, i.e., a norm, $\phi: R^n \rightarrow R^1$ is said to be symmetric if (i) $\phi(\mathbf{x}) \geq 0$, with equality iff $\mathbf{x} = \mathbf{0}$, (ii) $\phi(c\mathbf{x}) = |c|\phi(\mathbf{x})$ for any scalar c , and (iii) $\phi(\mathbf{x} + \mathbf{y}) \leq \phi(\mathbf{x}) + \phi(\mathbf{y})$, for $\mathbf{x} = (x_1, x_2, \dots, x_n) \in R^n$, $\mathbf{y} \in R^n$, and if it also satisfies

$$(2.1) \quad \phi(\varepsilon_1 x_{i_1}, \varepsilon_2 x_{i_2}, \dots, \varepsilon_n x_{i_n}) = \phi(\mathbf{x}),$$

for any permutation $(i_1, i_2, \dots, i_n)$ of $(1, 2, \dots, n)$ and any combination $\varepsilon_i = \pm 1$ of the signs.

Let Φ_n denote the class of all SGF's on R^n as defined above. For any $\phi \in \Phi_n$ and $\mathbf{y} \in R^n$, define the dual

$$(2.2) \quad \phi^0(\mathbf{y}) = \max_{\mathbf{x}} \mathbf{x}'\mathbf{y} / \phi(\mathbf{x}).$$

Alternatively, $\phi^0(\mathbf{y})$ is the maximum of $\mathbf{x}'\mathbf{y}$ over the set $\{\mathbf{x} | \phi(\mathbf{x}) \leq 1\}$. Then $\phi^0 \in \Phi_n$ and $(\phi^0)^0 = \phi$. ϕ^0 is known variously as the conjugate (von Neumann, Hewitt and Ross), the associate (Schatten) or the polar (Rockafellar [9]) of ϕ .

Now for $A \in B(H)$ let $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n \geq 0$ be the singular values of A , i.e., the eigenvalues of $(AA^*)^{1/2}$. A major accomplishment of von Neumann's paper [13] is a demonstration of the fact that for any SGF $\phi \in \Phi_n$

$$(2.3) \quad \|A\|_\phi = \phi(\alpha_1, \alpha_2, \dots, \alpha_n)$$

is a UI norm on $B(H)$. Indeed every UI norm is generated in this manner by some SGF (see Schatten [10]). Furthermore,

$$(2.4) \quad \|A\| = \sup_{B \neq 0} \frac{|\text{tr } AB|}{\|B\|_\phi}$$

is the UI norm generated by the polar $\phi^0 \in \Phi_n$ of ϕ (see Hewitt and Ross [5, p. 703]). Call $\|A\|_{\phi^0}$, i.e., $\|A\|$ of (2.4), the polar of $\|A\|_\phi$.

3. The main result. For any $\mathbf{x} = (x_1, x_2, \dots, x_n) \in R^n$, let $x_{(1)} \geq x_{(2)} \geq \dots \geq 0$ denote the ordered values of $|x_i|$, $i = 1, 2, \dots, n$. For any integer m , $m \leq n$, and $\phi \in \Phi_n$ let

$$(3.1) \quad \phi_m(\mathbf{x}) = \phi(x_{(1)}, x_{(2)}, \dots, x_{(m)}).$$

Then it is easy to verify that ϕ_m defines an SGF on R^n . That is, $\phi_m \in \Phi_n$ and, in the notation of §2, the corresponding

$$(3.2) \quad \|A\|_{\phi_m} = \phi(\alpha_1, \alpha_2, \dots, \alpha_m)$$

is a UI norm on $B(H)$. In this section we characterize the polar ϕ_m^0 of $\phi_m \in \Phi_n$ in terms of the polar ϕ^0 of $\phi \in \Phi_n$ and consequently obtain the polar of $\|A\|_{\phi_m}$. In order to describe and prove the main results we need the following two lemmas.

LEMMA 3.1. If $b_1 \geq b_2 \geq \cdots \geq b_n \geq 0$ and m is an integer $1 \leq m \leq n$, then there exists a unique integer k , $0 \leq k < m$, such that

$$(3.3) \quad b_{k+1} \leq \sum_{j=k+1}^n b_j / (m - k) < b_k,$$

the second inequality being inoperative if $k = 0$.

PROOF. For $r = 1, 2, \dots, m$, let

$$(3.4) \quad \beta_r = (m - r)b_r - \sum_{j=r+1}^n b_j.$$

Then

$$\beta_r - \beta_{r+1} = (m - r)(b_r - b_{r+1}) \geq 0,$$

for $r = 1, 2, \dots, m - 1$, and $\beta_m = -\sum_{m+1}^n b_j \leq 0$. Hence, there is a unique k such that $\beta_k > 0 \geq \beta_{k+1}$, i.e. the k which satisfies (3.3).

LEMMA 3.2. Let $b_1 \geq b_2 \geq \cdots \geq b_s \geq 0$, and $\phi \in \Phi_s$. Then the maximum value $\phi^0(\mathbf{b})$ of the sum $\sum_1^s a_i b_i$ subject to $\phi(\mathbf{a}) = \phi(a_1, a_2, \dots, a_s) \leq 1$ is attained at a vector which satisfies $a_1 \geq a_2 \geq \cdots \geq a_s \geq 0$.

PROOF. The maximum must occur when the coordinates of $\mathbf{a}$ and $\mathbf{b}$ are similarly ordered. Otherwise by the rearrangement theorem, e.g., see Hardy, Littlewood and Pólya [4], the sum $\sum_1^s a_i b_i$ could be increased without violating the restriction $\phi(\mathbf{a}) \leq 1$.

THEOREM 3.3. Let $\phi_m \in \Phi_n$ be the SGF derived from $\phi \in \Phi_m$, $m \leq n$, according to (3.1). Then the polar $\phi_m^0 \in \Phi_n$ of ϕ_m is given by

$$(3.5) \quad \phi_m^0(\mathbf{y}) = \phi^0(y_{(1)}, y_{(2)}, \dots, y_{(k)}, \bar{y}, \bar{y}, \dots, \bar{y}),$$

where $\phi^0 \in \Phi_m$ is the polar of ϕ , $y_{(1)} \geq y_{(2)} \geq \cdots \geq y_{(n)} \geq 0$ are the ordered values of the magnitudes $|y_i|$ of the coordinates of $\mathbf{y} \in R^n$, k is obtained by applying Lemma 3.1 to the $y_{(i)}$'s, and $\bar{y} = \sum_{k+1}^n y_{(j)} / (m - k)$.

PROOF. Since we are working with SGF's, assume with no loss of generality that $y_1 \geq y_2 \geq \cdots \geq y_n \geq 0$. By definition, we have

$$(3.6) \quad \phi_m^0(\mathbf{y}) = \max \sum_1^n x_i y_i,$$

where the maximum is over the set $\{\mathbf{x} | \phi_m(\mathbf{x}) = \phi(x_{(1)}, x_{(2)}, \dots, x_{(m)}) \leq 1\}$, and the $x_{(i)}$'s are the ordered values of $|x_i|$, $i = 1, 2, \dots, n$. But

$$(3.7) \quad \sum_1^n x_i y_i \leq \sum_1^{m-1} x_{(i)} y_i + x_{(m)} \sum_m^n y_i,$$

with equality if, and only if, $x_1 \geq x_2 \geq \cdots \geq x_m = x_{m+1} = \cdots = x_n \geq 0$. Hence,

$$(3.8) \quad \phi_m^0(\mathbf{y}) = \max_{S_0} \left\{ \sum_1^{m-1} x_i y_i + x_m y_{[m]} \right\},$$

where $y_{[m]}$ denotes the tail-sum $\sum_m^n y_i$, and

$$(3.9) \quad S_0 = \{ \mathbf{x} \in R^m | x_1 \geq x_2 \geq \cdots \geq x_m \geq 0, \phi(x_1, x_2, \dots, x_m) \leq 1 \}.$$

Now suppose that the maximum in (3.8) is known to occur for $\mathbf{x}$ in the subset of S_0 defined by

$$(3.10) \quad S_j = \{ \mathbf{x} \in S_0 | x_{m-j} = x_{m-j+1} = \cdots = x_m \}.$$

For $j = 0$ this obviously is so. We will now prove that if the maximum is in S_j and $j < m - k - 1$, then in fact the maximum is in S_{j+1} . This will permit us to conclude that the maximum in (3.8) is in S_{m-k-1} .

To prove the inductive step, for any $\mathbf{x} \in S_j$ define $\mathbf{x}^*$ as follows:

$$(3.11) \quad x_i^* = \begin{cases} x_i, & i = 1, 2, \dots, m-j-2, \\ \frac{x_{m-j-1} + (j+1)x_{m-j}}{j+2}, & i = m-j-1, \dots, m. \end{cases}$$

Then by the symmetry and convexity of ϕ ,

$$(3.12) \quad \phi(x_1^*, \dots, x_m^*) \leq \phi(x_1, \dots, x_m) \leq 1,$$

so that $\mathbf{x}^* \in S_{j+1}$. In order to show that

$$(3.13) \quad \sum_1^{m-1} x_i y_i + x_m y_{[m]} \leq \sum_1^{m-1} x_i^* y_i + x_m^* y_{[m]}$$

we substitute the definition of x_i^* . Then the terms with $i \leq m-j-2$ cancel, and it remains to show that

$$(3.14) \quad x_{m-j-1} y_{m-j-1} + x_{m-j} \sum_{m-j}^n y_i \leq \frac{x_{m-j-1} + (j+1)x_{m-j}}{j+2} \sum_{m-j-1}^n y_i.$$

This is equivalent to showing that

$$(3.15) \quad 0 \leq (x_{m-j-1} - x_{m-j}) \left(\sum_{m-j}^n y_i - (j+1)y_{m-j-1} \right).$$

The first factor in (3.15) is nonnegative since $\mathbf{x} \in S_0$. In the notation of the proof of Lemma 3.1, as applied to the y 's, the second factor is $-\beta_{m-j-1}$. This is nonnegative as long as $j < m - k - 1$.

We have just proved that

$$(3.16) \quad \phi_m^0(\mathbf{y}) = \max_{\mathbf{x} \in S_{m-k-1}} \left\{ \sum_1^{m-1} x_i y_i + x_m y_{[m]} \right\}.$$

Since $x_{k+1} = x_{k+2} = \cdots = x_m$, this is equal to

$$\max_{\mathbf{x} \in S_{m-k-1}} \left\{ \sum_1^k x_i y_i + \sum_{k+1}^m x_i \bar{y} \right\}, \quad \text{where } \bar{y} = \frac{1}{m-k} \sum_{k+1}^n y_i.$$

Because the coefficients of $x_{k+1}, \dots, x_m$ are equal, this in turn equals

$$\max_{\mathbf{x} \in S_0} \left\{ \sum_1^k x_i y_i + \sum_{k+1}^m x_i \bar{y} \right\}.$$

But k has been chosen to make $y_1 \geq \cdots \geq y_k \geq \bar{y}$. Hence we can apply Lemma 3.2

to see that the last expression equals

$$(3.17) \quad \max_{\phi(x) \leq 1} \left\{ \sum_1^k x_i y_i + \sum_{k+1}^m x_i \bar{y} \right\} = \phi^0(y_1, y_2, \dots, y_k, \bar{y}, \dots, \bar{y}).$$

This completes the proof of the theorem.

Combining the definitions of a UI norm $\|A\|_\phi$ at (2.3), its polar $\|A\|_{\phi^0}$ at (2.4), and of a trimmed SGF ϕ_m at (3.1), we get

COROLLARY 3.4. *For any $\phi \in \Phi_m$, $m \leq n$, the trimmed SGF ϕ_m defines a UI norm on $B(H)$. Its polar is the UI norm defined by ϕ_m^0 at (3.5).*

4. Applications and implications. The following observations outline some possible uses, further consequences and suggestions resulting from the main structural result presented in the previous section.

(1) Two classes of SGF's commonly occur in literature as illustrations. For $1 \leq p \leq \infty$ and $\mathbf{x} \in R^n$, $\phi(\mathbf{x}) = (\sum_1^n |x_i|^p)^{1/p}$ is an SGF in Φ_n with $\phi^0(\mathbf{x}) = (\sum_1^n |x_i|^q)^{1/q}$, $q^{-1} = 1 - p^{-1}$. The UI norms of the operators generated by these SGF's are known as the Schatten p -norms. If $m \leq n$ is an integer, $\mathbf{x} \in R^n$, then in the notation of §3, $\phi(\mathbf{x}) = \sum_1^m x_{(i)}$ is an SGF, the corresponding UI norm $\|A\|_\phi$ being known as a Ky Fan k -norm. The results in §3 provide the duals of these norms.

(2) If $a_1 \geq a_2 \geq \dots \geq a_n \geq 0$ and $b_1 \geq b_2 \geq \dots \geq b_n \geq 0$, $m \leq n$ and $p \geq 1$, then taking $\phi(a_1, a_2, \dots, a_m) = (\sum_1^m a_i^p)^{1/p}$ we get

$$(4.1) \quad \sum_1^n a_i b_i \leq \left(\sum_1^m a_i^p \right)^{1/p} \left(\sum_1^k b_i^q + (m-k)\bar{b}^q \right)^{1/q},$$

where $q^{-1} = 1 - p^{-1}$, and k is as in §3. This generalization of Hölder's inequality is proved by Mudholkar, Freimer and Subbaiah [8] using arguments of convex analysis.

(3) If $0 < a_1 \leq a_2 \leq \dots \leq a_n$, and $p < 1$, then $(\sum_1^n a_i^p)^{1/p}$ is a concave homogeneous function of $(a_1, a_2, \dots, a_n)$. If $b_1 \geq b_2 \geq \dots \geq b_n \geq 0$, $m \leq n$, and $q^{-1} = 1 - p^{-1}$, then Mudholkar, Freimer and Subbaiah [8] also established the following reversal of Hölder's inequality:

$$(4.2) \quad \sum_1^n a_i b_i \geq \left(\sum_1^m a_i^p \right)^{1/p} \left(\sum_1^k b_i^q + (m-k)\bar{b}^q \right)^{1/q}.$$

The analogous reversal of the more general inequality in §3 can be developed in a straightforward manner.

(4) If f is a real valued convex function on R^n , then

$$f^c(\mathbf{y}) = \sup_{\mathbf{x} \in R^n} [\mathbf{y}'\mathbf{x} - f(\mathbf{x})],$$

the Fenchél conjugate, leads to the inequality $\mathbf{y}'\mathbf{x} \leq f(\mathbf{x}) + f^c(\mathbf{y})$. Analogues of the results in §3 for Fenchél conjugates exist and consequent matrix inequalities follow.

(5) A substantial implication of the structural result pertains to the problem of metrizing the completely continuous operators on an infinite-dimensional Hilbert space (see Schatten [11] and Simon [12]).

If A is a completely continuous operator on a Hilbert space of countable dimension, then the singular values of A (including multiplicities) can be written as a nonincreasing sequence $\{\alpha_i(A)\}$ of nonnegative real numbers which converges to zero. If U and V are unitary operators, then $\alpha_i(UAV) = \alpha_i(A)$ for all i . Conversely, if B is also completely continuous and $\alpha_i(B) = \alpha_i(A)$ for all i , then $B = UAV$ with U and V unitary.

If $1 \leq p < \infty$, we can define

$$(4.3) \quad \|A\|_p = \lim_{n \rightarrow \infty} \left[\sum_{i=1}^n \alpha_i(A)^p \right]^{1/p}.$$

Those completely continuous operators for which this is finite form a Banach space c_p with this as norm (see McCarthy [7]). If $p > 1$ the space is reflexive, and $(c_p)^* = c_q$, where $p^{-1} + q^{-1} = 1$. However, c_1 is not reflexive, and $(c_1)^*$ consists of all bounded operators.

More generally, let ϕ be any symmetric gauge function defined on the space of sequences having only a finite number of nonzero terms. Then we can still use definition (2.2) to define the polar ϕ^0 . (Furthermore, if ϕ_m is defined by (3.1) then ϕ_m^0 is still given by Theorem 3.3.) Both ϕ and ϕ^0 can be extended to the space of sequences converging to zero by taking limits. Of course, for some such sequences, ϕ or ϕ^0 may be equal to $+\infty$.

The unitarily invariant norms for completely continuous operators are of the form $\|A\|_\phi = \phi(\{\alpha_i(A)\})$. Those A for which $\|A\|_\phi < \infty$ form a Banach space c_ϕ . We still have the equation

$$(4.4) \quad \|A\|_{\phi^0} = \sup_{B \neq 0} \frac{|\operatorname{tr} AB|}{\|B\|_\phi}.$$

This leads to the inclusion relation

$$(4.5) \quad (c_\phi)^* \supset c_{\phi^0}$$

with equality holding if c_ϕ is reflexive. It so happens that c_{ϕ_m} is not reflexive but $(c_{\phi_m})^* = c_{\phi_m^0}$ still holds.

REFERENCES

1. R. Bhatia, *Analysis of spectral variation and some inequalities*, Trans. Amer. Math. Soc. **272** (1982), 323–331.
2. K. Fan, *Maximum properties and inequalities for the eigenvalues of completely continuous operators*, Proc. Nat. Acad. Sci. U.S.A. **37** (1951), 760–766.
3. I. C. Gohberg and M. G. Krein, *Introduction to the theory of linear non-self adjoint operators*, Amer. Math. Soc. Transl. **18** (1969).
4. G. H. Hardy, J. E. Littlewood and G. Polya, *Inequalities*, Cambridge Univ. Press, 1934.
5. E. Hewitt and K. A. Ross, *Abstract harmonic analysis*, Vol. II, Springer-Verlag, Berlin, 1970.
6. R. Kunze, *L^p -Fourier transforms on locally compact unimodular groups*, Trans. Amer. Math. Soc. **89** (1958), 519–540.
7. C. A. McCarthy, c_p , Israel J. Math. **5** (1967), 249–271.
8. G. S. Mudholkar, M. Freimer and P. Subbaiah, *An extension of Hölder's inequality*, J. Math. Anal. Appl. **102** (1984), 435–441.

9. R. T. Rockafellar, *Convex analysis*, Princeton Univ. Press, Princeton, N.J., 1970.
10. R. Schatten, *A theory of cross-spaces*, Ann. of Math. Studies, No. 26, Princeton Univ. Press, Princeton, N.J., 1950.
11. ———, *Norm ideals of completely continuous operators*, Springer-Verlag, Berlin, 1960.
12. B. Simon, *Trace ideals and their applications*, Cambridge Univ. Press, 1979.
13. J. von Neumann, *Some matrix inequalities and metrization of matrix spaces*, Tomsk. Univ. Rev. **1** (1937), 286–300 (also Collected Works, Vol. IV).

DEPARTMENT OF STATISTICS AND GRADUATE SCHOOL OF MANAGEMENT, UNIVERSITY OF ROCHESTER,
ROCHESTER, NEW YORK 14627