DISTANCE BETWEEN HERMITIAN OPERATORS IN SCHATTEN CLASSES

by RAJENDRA BHATIA and PETER ŠEMRL

(Received 17th October 1994)

We consider the distance between a fixed Hermitian operator B and the unitary orbit of another Hermitian operator A and show that in each Schatten p-class, $1<p<\infty$, critical points of this distance function are at operators commuting with B. As a consequence we obtain a perturbation bound for the eigenvalues of Hermitian operators in these Schatten classes.

1991 *Mathematics subject classification*: 47B10.

1. Introduction

Let $\mathscr{H}$ be a complex separable Hilbert space. Let $\mathscr{I}_p$ be the Schatten p-ideal of operators on $\mathscr{H}$ and $\|\cdot\|_p$ the Schatten p-norm. The group of unitary operators on $\mathscr{H}$ will be denoted as $\mathscr{U}$.

The *unitary orbit* of an operator A is the set $\mathscr{U}_A=\{UAU^*\colon U\in\mathscr{U}\}$. For a fixed operator B in $\mathscr{I}_p$ let $f(X)=\|B-X\|_p$ for all X in $\mathscr{I}_p$. Our first theorem says that if $1<p<\infty$, and A, B are Hermitian operators in $\mathscr{I}_p$ then the critical points of f on $\mathscr{U}_A$ all commute with B. Our second theorem, a corollary of the first, gives a perturbation bound for eigenvalues of such operators.

Theorem 1. *Let A, B be Hermitian operators in $\mathscr{I}_p$, $1<p<\infty$. If A_0 is a critical point of the function $f(X)=\|B-X\|_p$ on $\mathscr{U}_A$ then A_0 commutes with B.*

Theorem 2. *Let A, B be Hermitian operators in $\mathscr{I}_p$, $1<p<\infty$. Let $\{\alpha_j\}$ and $\{\beta_j\}$ be the eigenvalues of A and B, where each eigenvalue is enumerated with its proper multiplicity. Then there exists a permutation σ of $\mathbb{N}$ such that*

$$\left(\sum_j |\alpha_j-\beta_{\sigma(j)}|^p\right)^{1/p}\leqq \|A-B\|_p. \tag{1}$$

When the space $\mathscr{H}$ is finite-dimensional, results more general than (1) are known. See [2] for a detailed discussion and for references to the original literature. When $\mathscr{H}$ is infinite-dimensional, the result of Theorem 2 has been proved by Cochran and Hinds in the special case $p=2$ but more generally for normal operators A, B [4]. Other known extensions to infinite dimensions have one common feature: a spurious zero eigenvalue is added, often with infinite multiplicity, to the genuine eigenvalues of A and B. See [5], [6] and [3] where this point is discussed in some detail.

2. Proofs

A formula for the Fréchet derivative of the p-norm, $1<p<\infty$, has been obtained by Aiken, Erdos, and Goldstein [1]. Let $g(X)=\|X\|_p^p$ and let $Dg(X)$ denote the Fréchet derivative of g at X. This is a linear functional on $\mathscr{I}_p$, whose action is given as

$$Dg(X)(Y)=\frac{d}{dt}\bigg|_{t=0} g(X+tY). \tag{2}$$

Let the polar decomposition of X be $X=U|X|$. Theorem 2.1 in [1] then says

$$Dg(X)(Y)=p\mathrm{Re}(\mathrm{tr}(|X|^{p-1}U^*Y)), \tag{3}$$

for $1<p<\infty$.

Proof of Theorem 1. Let A_0 be a critical point of f on $\mathscr{U}_A$. For each skew-Hermitian operator K the family $e^{-tK}A_0e^{tK}$ is a one parameter curve in $\mathscr{U}_A$ passing through A_0. So, we must have

$$\frac{d}{dt}\bigg|_{t=0}\|B-e^{-tK}A_0e^{tK}\|_p^p=0. \tag{4}$$

But, expanding the exponential function in a power series, we obtain

$$B-e^{-tK}A_0e^{tK}=B-A_0+t[K,A_0]+t^2\ldots,$$

where, as usual $[X,Y]$ denotes the commutator $XY-YX$. So, from (4) we obtain

$$\frac{d}{dt}\bigg|_{t=0}\|B-A_0+t[K,A_0]\|_p^p=0. \tag{5}$$

If the polar decomposition of $B-A_0$ is $B-A_0=U|B-A_0|$, then using (2), (3) and (5) we get

$$\mathrm{Re}(\mathrm{tr}(|B-A_0|^{p-1}U^*[K,A_0]))=0. \tag{6}$$

Since $B-A_0$ is Hermitian, U is a unitary operator commuting with $|B-A_0|$. From this it is easy to see that the operator $C=|B-A_0|^{p-1}U^*$ is Hermitian. The condition (6) can be written also as

$$\mathrm{Re}(\mathrm{tr}(K[A_0, C])) = 0. \tag{7}$$

Since A_0 and C are both Hermitian, $[A_0, C]$ is skew-Hermitian. But the trace of the product of two skew-Hermitian operators is always real. So, we have from (7)

$$\mathrm{tr}(K[A_0, C]) = 0. \tag{8}$$

Now recall that K was any arbitrary skew-Hermitian operator. Hence, we must have

$$\mathrm{tr}(X[A_0, C]) = 0 \tag{9}$$

for every operator X. But this can happen only if $[A_0, C] = 0$. In other words, A_0 commutes with $B - A_0$, and hence with B.

Proof of Theorem 2. The assertion of the theorem is true in the finite-dimensional case [2, Theorem 9.7]. Using this and the fact that finite-rank operators are dense in $\mathscr{I}_p$ one can easily prove the following. For every $\varepsilon > 0$ there exists a permutation σ of $\mathbb{N}$ such that

$$\sum_j |\alpha_j - \beta_{\sigma(j)}|^p \leqq \|A - B\|_p^p + \varepsilon. \tag{10}$$

From this it follows that

$$\inf_\sigma \sum |\alpha_j - \beta_{\sigma(j)}|^p \leqq \inf_U \|B - UAU^*\|_p^p, \tag{11}$$

where, σ varies over all permutations of $\mathbb{N}$ and U over all unitary operators. This is a consequence of the fact that the eigenvalues of UAU^* are the same as those of A.

Now two possibilities arise. First, the infimum on the right hand side of (11) could be strictly smaller than $\|B - UAU^*\|_p^p$ for all U in $\mathscr{U}$. In that case, so is the infimum on the left hand side of (11). We then have the desired inequality (1). The second possibility is that the infimum on the right hand side of (11) is actually a minimum. Let $A_0 = U_0AU_0^*$ be this minimal point. Then by Theorem 1, A_0 commutes with B. But then A_0 and B have a common orthonormal basis of eigenvectors. Hence, there is a permutation σ of $\mathbb{N}$ such that

$$\sum_j |\alpha_j - \beta_{\sigma(j)}|^p = \|A_0 - B\|_p^p \leq \|A - B\|_p^p.$$

This proves the theorem.

When $\mathscr{H}$ is finite-dimensional analogues of the inequality (1) are true for all norms induced by symmetric gauge functions. This has been generalized to the infinite-dimensional case by Markus [6] with the modification that a zero eigenvalue with arbitrary multiplicity is added. For $1<p<\infty$ we have shown that it is not necessary to do so, though at the cost of losing the information about the pairing of the two sets of eigenvalues (spurious zeros added) that Markus provides. It would be interesting to know whether the same holds for other symmetric gauge functions.

Acknowledgements. The first author thanks the department of Mathematics of the University of Ljubljana and specially Professor M. Omladič, for their hospitality in July 1994 when this work was done. This research was supported in part by the Ministry of Science of Slovenia. We thank an anonymous referee for bringing the paper [4] to our attention.

REFERENCES

1. J. G. Aiken, J. A. Erdos and J. A. Goldstein, Unitary approximation of positive operators, *Illinois J. Math.* **24** (1980), 61–72.

2. R. Bhatia, *Perturbation bounds for matrix eigenvalues* (Pitman Res. Notes Math. **162** (Longman, Harlow, 1987).

3. R. Bhatia and L. Elsner, The Hoffman–Wielandt inequality in infinite dimensions, *Proc. Indian Acad. Sci. Math. Sci.*, to appear.

4. J. A. Cochran and E. W. Hinds, Improved error bounds for the eigenvalues of certain normal operators, *SIAM J. Numer. Anal.* **9** (1972), 446–453.

5. S. Friedland, Inverse eigenvalue problems, *Linear Algebra Appl.* **17** (1977), 15–51.

6. A. S. Markus, The eigen and singular values of the sum and product of linear operators, *Russian Math. Surveys* **19** (1964), 92–120.

Indian Statistical Institute
New Delhi-110016
India

University of Maribor
Smetanova 17
P.O. Box 224
62000 Maribor
Slovenia