Functional Analysis,
Approximation and
Computation
4:1 (2012), 65–70

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/faac

A note on the orthogonality of bounded linear operators

Barraa Mohamed[a], Boumazgour Mohamed[b]

[a]*Départment de Mathématiques, Faculté des Sciences Semlalia, B.P 2390 Marrakesh, Morocco*
[b]*Faculté des Sciences Juridiques et Sociales, Univerité Ibn Zohr, B. P 8658 Agadir, Morocco*

Abstract. Let A and B be bounded linear operators on a Hilbert space H. The operator B is said to be orthogonal to A whenever $\|B + \lambda A\| \geq \|B\|$ for all $\lambda \in \mathbb{C}$. We give a theorem on the orthogonality and introduce the center of mass of a pair of operators. Some distance formulas are also given.

1. Introduction

Let H be a complex separable Hilbert space and let $\mathcal{L}(H)$ denote the algebra of all bounded linear operators on H. If $A, B \in \mathcal{L}(H)$, then B is said to be orthogonal to A (in the Birkhoff-James sense [7]) whenever

$$\|B + \lambda A\| \geq \|B\| \text{ for all } \lambda \in \mathbb{C}.$$

Orthogonality of matrices (more generally bounded linear operators on an infinite dimensional Hilbert space) has been studied by many authors [2, 7, 9, 13, 15, 16].

In [15], for $A \in \mathcal{L}(H)$, Stampfli defined the center of mass of A to be the scalar $c(A)$ that satisfies the equality

$$\|A - c(A)\| = \min\{\|A - \lambda\| : \lambda \in \mathbb{C}\}.$$

We refer to [3, 4, 15, 16] for basic properties of the center of mass of an operator.

In this paper, we study the orthogonality of a given pair of operators and investigate the properties of its center of mass.

The paper is organized as follows. In Section 1, we give a characterization of orthogonality of any pair of operators on a Hilbert space. Then we introduce the center of mass of such a pair and study its continuity.

In Section 2, we give some formulas for the distance of an operator to the class of scalar multiples of another one.

For $A \in \mathcal{L}(H)$, let $\sigma(A)$ denote the spectrum of A and denote $m(A) = \inf\{\|Ax\| : x \in H, \|x\| = 1\}$.

2010 *Mathematics Subject Classification*. Primary 47A12; Secondary 47A30, 47B47.
Keywords. Birkhoff-James orthogonality; norm; numerical range; distance formulas.
Received: March 21, 2012; Accepted: April 4, 2012
Communicated by Dragan S. Djordjević
Email addresses: barraa@ucam.ac.ma (Barraa Mohamed), boumazgour@hotmail.com (Boumazgour Mohamed)

2. Orthogonality and center of mass

Let $A, B \in \mathcal{L}(H)$. Following [10], the numerical range of A^*B relatively to B, denoted by $W_B(A^*B)$, is defined to be the set

$$\{\lambda \in \mathbb{C} : \exists \{x_n\} \subseteq H, \|x_n\| = 1, \lambda = \lim_n < A^*Bx_n, x_n >, \lim_n \|Bx_n\| = \|B\|\}.$$

$W_B(A^*B)$ is a compact convex subset of the complex plane. In case $A = I$, the identity operator on H, it reduces to the maximal numerical range of the operator B (see [15]).

The following theorem gives a characterization of Birkhoff-James's orthogonality for a pair of bounded operators.

Theorem 2.1. *For $A, B \in \mathcal{L}(H)$, the following conditions are equivalent*

i) $0 \in W_B(A^*B)$,

ii) $\|B\| \leq \|B + \lambda A\|$ *for all* $\lambda \in \mathbb{C}$,

iii) $\|B\|^2 + |\lambda|^2 m^2(A) \leq \|B + \lambda A\|^2$ *for all* $\lambda \in \mathbb{C}$.

Proof. $iii) \to ii)$: obvious.

$i) \to iii)$: If $0 \in W_B(A^*B)$, then there exists a unit sequence $\{x_n\}_n$ in H such that $\lim_n < A^*Bx_n, x_n >= 0$ and $\lim_n \|Bx_n\| = \|B\|$. For $\lambda \in \mathbb{C}$ and $n \in \mathbb{N}$ we have

$$\|B + \lambda A\|^2 \geq \|Bx_n + \lambda Ax_n\|^2 = \|Bx_n\|^2 + |\lambda|^2 \|Ax_n\|^2 + 2Re(\overline{\lambda} < A^*Bx_n, x_n >).$$

Hence

$$\|B + \lambda A\|^2 \geq \|B\|^2 + |\lambda|^2 \limsup_n \|Ax_n\|^2 \geq \|B\|^2 + |\lambda|^2 m^2(A).$$

$ii) \to i)$: Let us suppose that $0 \notin W_B(A^*B)$. By rotation, we may suppose that $W_B(A^*B)$ is contained in the right half-plane, and therefore there is a line which separates 0 from $W_B(A^*B)$. So there exists $\tau > 0$ such that $Re(W_B(A^*B)) \geq \tau$.

Let $\mathcal{S} = \{x \in H : \|x\| = 1 \text{ and } Re(< A^*Bx, x >) \leq \frac{\tau}{2}\}$ and let $\eta = \sup\{\|Bx\| : x \in \mathcal{S}\}$. Then $\eta < \|B\|$, otherwise $\|B\| = \lim_n \|Bx_n\|$, where $\{x_n\}_n$ is a sequence of elements of $\mathcal{S}$. We may suppose that the sequence $\{< A^*Bx_n, x_n >\}_n$ is convergent since it is bounded. If we set $\lambda = \lim_n < A^*Bx_n, x_n >$ then $Re(\lambda) \leq \frac{\tau}{2}$ and, this is a contradiction.

Let $\mu = \min\{\frac{\tau}{2\|A\|^2}, \frac{\|B\|-\eta}{2\|A\|}\}$. We claim that $\|B - \mu A\| < \|B\|$. Let x be a unit vector in H. We consider two cases. First, if $x \in \mathcal{S}$ then

$$\|Bx - \mu Ax\| \leq \|Bx\| + \mu\|A\| \leq \eta + \frac{\|B\| - \eta}{2} < \|B\|.$$

Next, if $x \notin \mathcal{S}$, write $Bx = (a + ib)Ax + y$ with $a, b \in \mathbb{R}$, $< Ax, y >= 0$ and $y \in H$. Then

$$\begin{aligned} \|Bx - \mu Ax\|^2 &= ((a-\mu)^2 + b^2)\|Ax\|^2 + \|y\|^2 \\ &= \|Bx\|^2 + (\mu^2 - 2a\mu)\|Ax\|^2 \\ &\leq \|B\|^2 + (\mu^2 - 2a\mu)\inf\{\|Ax\|^2 : \|x\| = 1, x \notin \mathcal{S}\} \\ &< \|B\|^2 \end{aligned}$$

since $\mu^2 - 2a\mu < 0$ and $\inf\{\|Ax\|^2 : \|x\| = 1, x \notin \mathcal{S}\} > 0$. Therefore we deduce that $\|B - \mu A\| < \|B\|$ which contradicts our hypothesis. This ends the proof.

□

The special case when $B = I$ has been considered by Kittaneh [8], and used to characterize zero-trace matrices. The case $A = I$ was studied by Stampfli [15], and used to compute the norm of an inner derivation on a Hilbert space. Orthogonality of a pair of matrices with respect to Schatten p-norms ($1 \leq p \leq \infty$) has been characterized in [2, 9].

As a consequence of Theorem 2.1, we have

Corollary 2.2. *Let $A, B \in \mathcal{L}(\mathcal{H})$ with $m(A) > 0$. Then there exists a unique $z_0 \in \mathbb{C}$ such that $\|B - z_0A\|^2 + |\lambda|^2 m^2(A) \leq \|B - z_0A + \lambda A\|^2$ for all λ. Moreover, $0 \in W_B(A^*B)$ if and only if $z_0 = 0$.*

Proof. Since $\|B - \lambda A\|$ is large for λ large, $\inf\{\|B - \lambda A\| : \lambda \in \mathbb{C}\}$ must be attained at some point, say z_0. Hence $\|B - z_0A\| \leq \|B - \lambda A\|$ for all $\lambda \in \mathbb{C}$. By Theorem 2.1, it follows that

$$\|B - z_0A\|^2 + |\lambda|^2 m^2(A) \leq \|B - z_0A + \lambda A\|^2 \text{ for all } \lambda.$$

Next we show that z_0 is unique. To do this suppose that z_1 is another point satisfying the inequality

$$\|B - z_1A\|^2 + |\lambda|^2 m^2(A) \leq \|B - z_1A + \lambda A\|^2 \text{ for all } \lambda.$$

Specializing to $\lambda = z_1 - z_0$, we get

$$\|B - z_1A\|^2 + |z_1 - z_0|^2 m^2(A) \leq \|B - z_0A\|^2 \leq \|B - z_1\|^2.$$

Thus $|z_1 - z_0|^2 m^2(A) \leq 0$. Since $m(A) > 0$, we deduce that $z_1 = z_0$.

The remainder of the proof follows easily from the above theorem . □

Definition 2.3. *For $A, B \in \mathcal{L}(\mathcal{H})$ with $m(A) > 0$, we define the center of mass of A^*B relatively to B to be the point z_0 specified in Corollary 2.2, and designate it by $c_B(A^*B)$.*

Clearly, the center of mass of A^*B relatively to B is the scalar satisfying

$$\|B - c_B(A^*B)A\| = \inf\{\|B - \lambda A\| : \lambda \in \mathbb{C}\}.$$

The special case when A is the identity is of special interest, see [3, 4, 15, 16] and the references cited therein. When the operator B is normal, then its center of mass is exactly the center of the smallest disk containing the spectrum $\sigma(B)$.

Using Corollary 2.2, we prove the following continuity theorem

Theorem 2.4. *Let $A, B, C \in \mathcal{L}(\mathcal{H})$ with $m(A) > 0$. If $\delta \geq \|B - C\|$, then*

$$|c_B(A^*B) - c_C(A^*C)| \leq \frac{1}{2m^2(A)}\left(\delta\|A\| + \sqrt{\delta^2\|A\|^2 + 8m^2(A)\delta\|C - c_C(A^*C)A\|}\ \right).$$

Consequently, the map $B \longmapsto c_B(A^*B)$ is uniformly continuous.

Proof. Without loss of generality we may assume that $c_C(A^*C) = 0$. Denote $c_B(A^*B) = z_0$. For $\delta_0 = \|B - C\|$, we have by Theorem 2.1,

$$\begin{aligned}
\|B\|^2 \geq& \|B - z_0A\|^2 + |z_0|^2 m^2(A)\\
&\geq |z_0|^2 m^2(A) + (\|C - z_0A\| - \delta_0)^2\\
&= |z_0|^2 m^2(A) + \|C - z_0A\|^2 - 2\delta_0\|C - z_0A\| + \delta_0^2\\
&\geq 2|z_0|^2 m^2(A) + \|C\|^2 - 2\delta_0(\|C\| + |z_0|\|A\|) + \delta_0^2\\
&\geq 2|z_0|^2 m^2(A) + \|B\|^2 - 2\delta_0\|C\| - 2\delta_0(\|C\| + |z_0|\|A\|).
\end{aligned}$$

Hence

$$|z_0|^2 m^2(A) - \delta_0|z_0|\|A\| - 2\delta_0\|C\|| \leq 0.$$

From this we derive that

$$|c_B(A^*B) - c_C(A^*C)| \leq \frac{1}{2m^2(A)}(\delta_0\|A\| + \sqrt{\delta_0^2\|A\|^2 + 8m^2(A)\|C - c_C(A^*C)A\|}\ )$$
$$\leq \frac{1}{2m^2(A)}(\delta\|A\| + \sqrt{\delta^2\|A\|^2 + 8m^2(A)\|C - c_C(A^*C)A\|}\ ).$$

□

3. Distance formulas

In this section we give some formulas for the distance of an operator to the class of multiple scalars of another one. Our results cover and extend those of Ando [1, Theorem 12.59], Bhatia and semrl [2], and Stolov [14] (see also [11, 12]).

For simplicity of notation, we denote $d(B, \mathbb{C}A) = \inf\{\|B - \lambda A\| : \lambda \in \mathbb{C}\}$.

Theorem 3.1. *If $A, B \in \mathcal{L}(\mathcal{H})$ with $m(A) > 0$, then*

$$d(B, \mathbb{C}A) = \sup\{| < Bx, y > | : x, y \in H, \|x\| = \|y\| = 1 \text{ and } < Ax, y >= 0\}.$$

Proof. Let $\lambda \in \mathbb{C}$ and let $x, y \in H$ with $\|x\| = \|y\| = 1$ and $< Ax, y >= 0$. We have

$$| < Bx, y > | = | < (B - \lambda A)x, y > | \leq \|B - \lambda A\|.$$

Thus

$$d(B, \mathbb{C}A) \geq \sup\{| < Bx, y > | : x, y \in H, < Ax, y >= 0 \text{ and } \|x\| = \|y\| = 1\}.$$

Now, let $\mu_0 \in \mathbb{C}$ such that $\|B - \mu_0 A\| = d(B, \mathbb{C}A)$. Without loss of generality we may suppose that $\mu_0 = 0$. Then $\|B\| = d(B, \mathbb{C}A)$. By virtue of Theorem 2.1, we have $0 \in W_B(A^*B)$. Thus one can find a unit sequence $\{x_n\} \subseteq H$ such that

$$\lim_n < A^*Bx_n, x_n >= 0 \text{ and } \lim_n \|Bx_n\| = \|B\|.$$

Write

$$Bx_n = \alpha_n Ax_n + \beta_n y_n$$

with $< y_n, Ax_n >= 0$, $\|y_n\| = 1$ and $\alpha_n, \beta_n \in \mathbb{C}$. Since

$$\lim_n \|Bx_n\|^2 = \lim_n |\alpha_n|^2 \|Ax_n\|^2 + |\beta_n|^2$$

and

$$\lim_n \alpha_n \|Ax_n\|^2 = \lim_n < A^*Bx_n, x_n >= 0,$$

then

$$\lim_n |\alpha_n|^2 \|Ax_n\|^2 = 0,$$

That is,

$$\|B\| = \lim_n |\beta_n| = \lim_n | < Bx_n, y_n > | \leq \sup\{| < Bx, y > | :< Ax, y >= 0, \|x\| = \|y\| = 1\}.$$

This completes the proof. □

In case $A = I$, a different proof attributed to Ando can be found in [1, Theorem 12.59]. The proof presented here is shorter.

In [11], for a matrix B, Parker proved that $\|B - m(B)\| = \max_{\|x\|=1}\{\|Bx\|^2 - | < Bx, x > |^2\}^{\frac{1}{2}}$ (see also [12, 14]). For a pair (A, B) of operators, we have the following generalization.

Theorem 3.2. *If $A, B \in \mathcal{L}(\mathcal{H})$ with $m(A) > 0$, then*

$$d(B, \mathbb{C}A) = \sup_{\|x\|=1}\{\|Bx\|^2 - \frac{|<Bx, Ax>|^2}{\|Ax\|^2}\}^{\frac{1}{2}}.$$

Proof. Let $x \in H$ and let $\lambda \in \mathbb{C}$. We have

$$\begin{aligned}\|B - \lambda A\|^2 \geq \|(B - \lambda A)x\|^2 &= \|Bx\|^2 + |\lambda|^2 \|Ax\|^2 - 2Re(\overline{\lambda} < Bx, Ax >)\\ &= \|Bx\|^2 + |\lambda \|Ax\| - \frac{<Ax, Bx>}{\|Ax\|}|^2 - \frac{|<Ax, Bx>|^2}{\|Ax\|^2}\\ &\geq \|Bx\|^2 - \frac{|<Ax, Bx>|^2}{\|Ax\|^2}.\end{aligned}$$

Consequently, we derive that

$$d(B, \mathbb{C}A) \geq \sup_{\|x\|=1}\{\|Bx\|^2 - \frac{|<Ax, Bx>|^2}{\|Ax\|^2}\}^{\frac{1}{2}}.$$

Conversely, let $\mu_0 \in \mathbb{C}$ such that $\|B - \mu_0 A\| = d(B, \mathbb{C}A)$. By Theorem 2.1, there exists a unit sequence $\{x_n\}$ such that $\lim_n < A^*(B - \mu_0 A)x_n, x_n >= 0$ and $\lim_n \|(B - \mu_0 A)x_n\| = \|B - \mu_0 A\|$. As before, we may assume that $\mu_0 = 0$. We have

$$\|B\|^2 = d^2(B, \mathbb{C}A) = \lim_n \|Bx_n\|^2 - \frac{|<Ax_n, Bx_n>|^2}{\|Ax_n\|^2}.$$

Whence

$$d(B, \mathbb{C}A) = \sup_{\|x\|=1}\{\|Bx\|^2 - \frac{|<Bx, Ax>|^2}{\|Ax\|^2}\}^{\frac{1}{2}}.$$

□

If K is a compact subset of the plane, then among all closed disks containing K there is a unique (possibly degenerate) disk of least radius. For $B \in \mathcal{L}(\mathcal{H})$, we denote R_B the radius of the smallest disk containing $\sigma(B)$.

As a consequence of Theorem 3.2, we have the next inequality (see [5]).

Corollary 3.3. *If $B \in \mathcal{L}(\mathcal{H})$, then*

$$R_B \leq \sup_{\|x\|=1}\{\|Bx\|^2 - |<Bx, x>|^2\}^{\frac{1}{2}}.$$

Proof. Let $\lambda \in \mathbb{C}$. Since $\sigma(B - \lambda) = \sigma(B) - \lambda$, it follows that $R_{B-\lambda} = R_B$. But $R_{B-\lambda} \leq \|B - \lambda\|$. Thus we conclude that $R_B \leq d(B, \mathbb{C}I)$, which ends the proof. □

Remark 3.4. *Recall that a bounded operator $B \in \mathcal{L}(\mathcal{H})$ is normaloid whenever $r(B) = \|B\|$([6]), where r denotes the spectral radius. If $B - \lambda$ is normaloid for all λ, then $R_B = \sup_{\|x\|=1}\{\|Bx\|^2 - |<Bx, x>|^2\}^{\frac{1}{2}}$. Indeed, for each $\lambda \in \mathbb{C}$, $\sigma(B - \lambda) \subseteq \{z : |z| \leq r(B - \lambda)\}$; hence $R_B \leq r(B - \lambda)$ and therefore $R_B \leq \inf\{r(B - \lambda) : \lambda \in \mathbb{C}\}$. On the other hand, let λ_0 be the center of the smallest disk containing $\sigma(B)$. We have $R_B = R_{B-\lambda_0}$ and $\sigma(B - \lambda)$ is contained in the disk centered at the origin and with radius R_B. Then*

$$\inf\{r(B - \lambda) : \lambda \in \mathbb{C}\} \leq r(B - \lambda_0) \leq R_B.$$

Consequently

$$R_B = \inf\{r(B - \lambda) : \lambda \in \mathbb{C}\} = \sup_{\|x\|=1}\{\|Bx\|^2 - |<Bx, x>|^2\}^{\frac{1}{2}}.$$

One of the methods to compute the center of mass of an operator is Williams's theorem [4, 16]. However, it is not usually easy to determine the exact value of it even in the finite dimensional case, see [3] for examples.

Let $C_2(H)$ denote the class of Hilbert-Schmidt operators on H. Define the center $c_{2,B}(A^*B)$ $(A, B \in C_\in(\mathcal{H}))$ by

$$\left\|B - c_{2,B}(A^*B)A\right\|_2 = \inf\{\|B - \lambda A\|_2 : \lambda \in \mathbb{C}\},$$

where $\|X\|_2^2 = tr(X^*X)$, $(X \in C_\in(\mathcal{H}))$ and tr is the usual functional trace. In this case, we may give a formula for $c_{2,B}(A^*B)$. Indeed, by orthogonal projection theorem, we have $B - c_{2,B}(A^*B)A$ is orthogonal to A, that is, $0 =< B - c_{2,B}(A^*B)A >= tr(A^*(B - c_{2,B}(A^*B)A)$. From this it follows that

$$c_{2,B}(A^*B) = \frac{tr(A^*B)}{\|A\|_2^2}.$$

References

[1] C. Apostol, L. Fialkow, D. Herrero, and D. Voiculescu, Approximation of Hilbert space operators, Volume II. Research Notes in Math. 102, Pitman, 1984.
[2] R. Bhatia and P. Semrl, Orthogonality of matrices and some distance problems, Linear Algebra Appli. **287**(1999), 77-85.
[3] J.I. Fujii and Y. Seo, Graphs and numerical center of mass, Math. Japon. **38**(1993), 351-359.
[4] J.I. Fujii and M. Fujii, Theorems of Williams and Prasanna, Math. Japon. **38**(1993), 35-37.
[5] G. Garske, An inequality concerning the smallest disk that contains the spectrum of an operator, Proc. Amer. Math. Soc. **78**(1980), 529-532.
[6] P.R. Halmos, A Hilbert Space Problem Book, Van Nostrand, Princeton, N.J., 1967.
[7] R.C. James, Orthogonality and linear functionals in normed linear spaces, Trans. Amer. Math. Soc. **61**(1947), 265-292.
[8] F. Kittaneh, On zero-trace matrices, Linear Algebra Appli. **51**(1991), 119-124.
[9] C-K. Li and H. Schneider, Orthogonality of matrices, Linear Algebra Appli. **347**(2002), 115-122.
[10] B. Magajna, On the distance to finite dimensional subspaces in operator algebras, J. London. Math. Soc. **47**(1993), 516-532.
[11] W.V. Parker, Sets of complex numbers associated with a matrix, Duke Math. J. **15**(1948), 711-715.
[12] S. Prasanna, The norm of a derivation and the B$\ddot{j}$ork-Thomee-Istratescu theorem, Math. Japon. **26**(1981), 585-588.
[13] I. Singer, Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces, Springer, New York, 1970.
[14] E.L. Stolov, The Hausdorff set of a matrix, Izv. Vyssh. Ucheben. Zaved. Math. **10**(1979), 98-100.
[15] J. Stampfli, The norm of a derivation, Pac. J. Math., **33**(1970), 737-747.
[16] J.P. Williams, Finite operators, Proc. Amer. Math. Soc, **26**(1970), 129-136.