J. Math. Anal. Appl. 407 (2013) 350–358

Contents lists available at SciVerse ScienceDirect

Journal of Mathematical Analysis and Applications

journal homepage: www.elsevier.com/locate/jmaa

Characterization of Birkhoff–James orthogonality

Tirthankar Bhattacharyya [a], Priyanka Grover [b,*]

[a] Department of Mathematics, Indian Institute of Science, Bangalore-560012, India

[b] Theoretical Statistics and Mathematics Unit, Indian Statistical Institute, Delhi Centre, 7, S.J.S. Sansanwal Marg, New Delhi-110016, India

ARTICLE INFO

Article history:
Received 7 February 2013
Available online 16 May 2013
Submitted by R.M. Aron

Keywords:
Birkhoff–James orthogonality
Subdifferential
Hausdorff–Toeplitz theorem
Numerical range
Separating Hyperplane theorem
Singular value decomposition
Operator tuples
C^*-algebra
Hilbert C^*-module
State
Variance

ABSTRACT

The Birkhoff–James orthogonality is a generalization of Hilbert space orthogonality to Banach spaces. We investigate this notion of orthogonality when the Banach space has more structures. We start by doing so for the Banach space of square matrices moving gradually to all bounded operators on any Hilbert space, then to an arbitrary C^*-algebra and finally a Hilbert C^*-module.

1. Introduction

Let X be a complex Banach space. An element $x \in X$ is said to be *Birkhoff–James orthogonal* to another element $y \in X$ if $\|x + \lambda y\| \geq \|x\|$ for all complex numbers λ. It is easy to see that Birkhoff–James orthogonality is equivalent to the usual orthogonality in case X is a Hilbert space. When $X = \mathbb{M}(n)$, the Banach space of all $n \times n$ complex square matrices, a very tractable condition of Birkhoff–James orthogonality was found by Bhatia and Šemrl in [4]. They showed that an $n \times n$ matrix A is Birkhoff–James orthogonal to an $n \times n$ matrix B if and only if there is a unit vector $x \in \mathbb{C}^n$ such that $\|Ax\| = \|A\|$ and $\langle Ax, Bx\rangle = 0$. Here $\|A\|$ denotes the operator norm of A. Later Benitez, Fernandez and Soriano [3] showed that a necessary and sufficient condition for the norm of a real finite dimensional normed space X to be induced by an inner product is that for any $A, B \in \mathscr{B}(X)$, A is Birkhoff–James orthogonal to B if and only if there exists a unit vector $x \in X$ such that $\|Ax\| = \|A\|$ and $\langle Ax, Bx\rangle = 0$.

Motivated by these, in this note we explore Birkhoff–James orthogonality in the setting of Hilbert C^*-modules. All inner products in this note are conjugate linear in the first component and linear in the second component. We start by giving a new proof of the Bhatia–Šemrl theorem using tools of convex analysis. This is more illuminating because it involves minimization of a certain convex function and is therefore a geometric approach. This is a new way of looking at the theorem and might be useful elsewhere. Another method was given by Kečkič in [7]. He first computes the φ-Gateaux derivative $D_{\varphi,A}(B)$ of the norm at A, in the B and φ directions, which is defined by $D_{\varphi,A}(B) = \lim_{t\to 0^+} \frac{\|A+te^{i\varphi}B\|-\|A\|}{t}$. Then he uses the property that A

* Corresponding author.
E-mail addresses: tirtha@member.ams.org (T. Bhattacharyya), priyanka.grover17@gmail.com, pgrover8r@isid.ac.in (P. Grover).

0022-247X/$ – see front matter
http://dx.doi.org/10.1016/j.jmaa.2013.05.022

is Birkhoff–James orthogonal to B if and only if $\inf_\varphi \mathrm{D}_{\varphi,A}(B) \geq 0$ to prove the Bhatia–Šemrl theorem. The unit vector x that crops up naturally suggests that this theorem should be generalizable to Hilbert C^*-modules with x replaced by a state of the underlying C^*-algebra. We handle the special module $\mathscr{B}(\mathcal{H}, \mathcal{K})$ in Section 3 with several applications to operator tuples. Section 4 is on arbitrary Hilbert C^*-modules.

These results can be applied to obtain some distance formulas in $\mathbb{M}(n)$ and other C^*-algebras. These are also important in problems related to derivations and operator approximations. In approximation theory the condition that A is Birkhoff–James orthogonal to B can be interpreted as follows. Suppose $A \in \mathbb{M}(n)$ is not in $\mathbb{C}B$, the subspace spanned by the matrix B. Then the zero matrix is the best approximation to A among all matrices in $\mathbb{C}B$.

Recently, Birkhoff–James orthogonality in Hilbert C^*-modules has been studied in [1] as our work was in progress. Some of our results overlap with them. Our approach is very different from that in [1]. We proceed gradually from square matrices to Hilbert C^*-modules. This is a natural development.

2. Bhatia–Šemrl theorem

The statement of the theorem is as follows.

Theorem 2.1. *Let $A, B \in \mathbb{M}(n)$. Then $\|A + \lambda B\| \geq \|A\|$ for all $\lambda \in \mathbb{C}$ if and only if there is a unit vector x such that $\|Ax\| = \|A\|$ and $\langle Ax, Bx\rangle = 0$.*

We first deal with $\|A + tB\| \geq \|A\|$ for all $t \in \mathbb{R}$. Our approach will revolve around the function $f(t) = \|A + tB\|$ mapping $\mathbb{R}$ into $\mathbb{R}_+$. To say that $\|A + tB\| \geq \|A\|$ for all $t \in \mathbb{R}$ is to say that f attains its minimum at the point 0. Since f is a convex function, the tools of convex analysis are available. The crux of our argument lies in calculating the *subdifferential* $\partial f(t)$ of f, and then showing that the point 0 is in the set $\partial f(0)$.

Definition 2.2. Let $f : X \to \mathbb{R}$ be a convex function. The subdifferential of f at a point $x \in X$ is the set $\partial f(x)$ of continuous linear functionals $v^* \in X^*$ such that

$$f(y) - f(x) \geq \mathrm{Re}\ v^*(y - x) \quad \text{for all } y \in X.$$

It is a convex subset of X^*. Its importance lies in the following proposition.

Proposition 2.3. *A convex function $f : X \to \mathbb{R}$ attains its minimum value at $x \in X$ if and only if $0 \in \partial f(x)$.*

We want to apply this to the function $f(t) = \|A + tB\|$ which we shall realize as the composition of two functions. The first of them is $t \to A + tB$ from $\mathbb{R}$ into $\mathbb{M}(n)$. The second is from $\mathbb{M}(n)$ to $\mathbb{R}_+$, sending any $T \in \mathbb{M}(n)$ to $\|T\|$. Thus we need to find subdifferentials of compositions. The subdifferential of the norm function has been calculated in [11]. We need it at a positive semidefinite matrix.

Proposition 2.4. *Let A be a positive semidefinite matrix. Then*

$$\partial\|A\| = \text{convex hull of } \{uu^* : \|u\| = 1, Au = \|A\|u\}. \tag{2.1}$$

To handle the composition maps, we need a chain rule.

Proposition 2.5. *Consider the composite map*

$$\mathbb{R}^n \xrightarrow{L} \mathbb{M}(n) \xrightarrow{g} \mathbb{R},$$

where g is a convex map and $L(x) = A + S(x)$ for all $x \in \mathbb{R}^n$, with $S : \mathbb{R}^n \to \mathbb{M}(n)$ being a linear map. Then the subdifferential of $g \circ L$ at a point $x \in \mathbb{R}^n$ is given by

$$\partial(g \circ L)(x) = S^*\partial g(L(x)) \quad \text{for all } x \in \mathbb{R}^n, \tag{2.2}$$

where $S^* : \mathbb{M}(n) \to \mathbb{R}^n$ is the adjoint of S satisfying

$$(S^*(T))'y = \mathrm{Re}\ \mathrm{tr}\ T^*S(y) \quad \text{for all } T \in \mathbb{M}(n),\ y \in \mathbb{R}^n.$$

(Here $(S^*(T))'$ means the transpose of the vector $S^*(T)$.)

These elementary facts can be found in [6]. We are now ready to prove a real version of the Bhatia–Šemrl theorem using these concepts.

Theorem 2.6. *Let $A, B \in \mathbb{M}(n)$. Then $\|A + t B\| \geq \|A\|$ for all $t \in \mathbb{R}$ if and only if there exists a unit vector x such that $\|Ax\| = \|A\|$ and $\mathrm{Re}\langle Ax, Bx\rangle = 0$.*