Journal of Inequalities in Pure and Applied Mathematics

RELATION BETWEEN BEST APPROXIMANT AND ORTHOGONALITY IN C_1-CLASSES

SALAH MECHERI

King Saud University, College of Sciences
Department of Mathematics
P.O. Box 2455, Riyah 11451
Saudi Arabia

EMail: mecherisalah@hotmail.com

ISSN (electronic): 1443-5756
162-05

volume 7, issue 2, article 57, 2006.

Received 21 May, 2005; accepted 13 March, 2006.

Communicated by: G.V. Milovanović

Abstract

Let E be a complex Banach space and let M be subspace of E. In this paper we characterize the best approximant to $A \in E$ from M and we prove the uniqueness, in terms of a new concept of derivative. Using this result we establish a new characterization of the best-$\mathcal{C}_1$ approximation to $A \in \mathcal{C}_1$ (trace class) from M. Then, we apply these results to characterize the operators which are orthogonal in the sense of Birkhoff.

2000 Mathematics Subject Classification: Primary 41A52, 41A35, 47B47; Secondary 47B10.

Key words: Best approximant, Schatten p-classes, Orthogonality, φ-Gateaux derivative.

I would like to thank the referee for his careful reading of the paper. His valuable suggestions, critical remarks, and pertinent comments made numerous improvements throughout.

This work was supported by the College of Science Research Center Project No. Math/2006/23.

Contents

1. Introduction

Let E be a complex Banach space and Let M be subspace of E. We first define orthogonality in E. We say that $b \in E$ is orthogonal to $a \in E$ if for all complex λ there holds

$$\|a + \lambda b\| \geq \|a\| . \tag{1.1}$$

This definition has a natural geometric interpretation. Namely, $b \perp a$ if and only if the complex line $\{a + \lambda b \mid \lambda \in \mathbb{C}\}$ is disjoint with the open ball $K(0, \|a\|)$, i.e., if and only if this complex line is a tangent line to $K(0, \|a\|)$. Note that if b is orthogonal to a, then a need not be orthogonal to b. If E is a Hilbert space, then from (1.1) follows $\langle a, b \rangle = 0$, i.e, orthogonality in the usual sense. Next we define the best approximant to $A \in E$ from M. For each $A \in E$ there exists a $B \in M$ such that

$$\|A - B\| \leq \|A - C\| \qquad \text{for all } C \in M.$$

Such B (if they exist) are called best approximants to A from M. Let $B(H)$ denote the algebra of all bounded linear operators on a complex separable and infinite dimensional Hilbert space H and let $T \in B(H)$ be compact, and let $s_1(X) \geq s_2(X) \geq \cdots \geq 0$ denote the singular values of T, i.e., the eigenvalues of $|T| = (T^*T)^{\frac{1}{2}}$ arranged in their decreasing order. The operator T is said to belong to the Schatten p-classes C_p $(1 \leq p < \infty)$ if

$$\|T\|_p = \left[\sum_{i=1}^{\infty} s_i(T)^p\right]^{\frac{1}{p}} = [\text{tr}(T)^p]^{\frac{1}{p}} < \infty, \qquad 1 \leq p < \infty,$$

where tr denotes the trace functional. Hence $\mathcal{C}_1$ is the trace class, C_2 is the Hilbert -Schmidt class, and C_∞ corresponds to the class of compact operators with

$$\|T\|_\infty = s_1(T) = \sup_{\|f\|=1} \|Tf\|$$

denoting the usual operator norm. For the general theory of the Schatten p-classes the reader is referred to [10]. Recall that the norm $\|\cdot\|$ of the $B-$space V is said to be Gâteaux differentiable at non-zero elements $x \in V$ if

$$\lim_{\mathbb{R}\ni t\to 0} \frac{\|x+ty\| - \|x\|}{t} = \operatorname{Re} D_x(y),$$

for all $y \in V$. Here $\mathbb{R}$ denotes the set of all reals, Re denotes the real part, and D_x is the unique support functional (in the dual space V^*) such that $\|D_x\| = 1$ and $D_x(x) = \|x\|$. The Gâteaux differentiability of the norm at x implies that x is a smooth point of the sphere of radius $\|x\|$. It is well known (see [4] and the references therein) that for $1 < p < \infty$, $\mathcal{C}_p$ is a uniformly convex Banach space. Therefore every non-zero $T \in \mathcal{C}_p$ is a smooth point and in this case the support functional of T is given by

$$D_T(X) = \operatorname{tr}\left[\frac{|T|^{p-1} U X^*}{\|T\|_p^{p-1}}\right], \tag{1.2}$$

for all $X \in \mathcal{C}_p$, where $T = U|T|$ is the polar decomposition of T. In this section we characterize the best approximant to $A \in E$ from M and we prove the uniqueness, in terms of a new concept of derivative. Using these results we establish a new characterization of the best- $\mathcal{C}_1$ approximation to $A \in \mathcal{C}_1$

from M in all Banach spaces without care of smoothness. Further, we apply these results to characterize the operators which are orthogonal in the sense of Birkhoff. It is very interesting to point out that these results has been done in L_1 and $\mathcal{C}(K)$ (see [9, 5]) but, at least to our knowledge, it has not been given, till now, for $\mathcal{C}_p$-classes.

To approach the concept of an approximant consider a set of mathematical objects (complex numbers, matrices or linear operator, say) each of which is, in some sense, "nice", i.e. has some nice property $\mathcal{P}$ (being real or self-adjoint, say): and let A be some given, not nice, mathematical object: then a $\mathcal{P}$ best approximant of A is a nice mathematical object that is "nearest" to A. Equivalently, a best approximant minimizes the distance between the set of nice mathematical objects and the given, not nice object.

Of course, the terms "mathematical object", "nice", "nearest", vary from context to context. For a concrete example, let the set of mathematical objects be the complex numbers, let "nice"=real and let the distance be measured by the modulus, then the real approximant of the complex number z is the real part of it, $\operatorname{Re} z = \frac{(z+\bar{z})}{2}$. Thus for all real x

$$|z - \operatorname{Re} z| \leq |z - x|.$$

2. Preliminaries

From the Clarckson-McCarthy inequalities it follows that the dual space $C_p^* \cong C_q$ is strictly convex. From this we can derive that every non zero point in C_p is a smooth point of the corresponding sphere. So we can check what is the unique support functional F_X.

However, if the dual space is not strictly convex, there are many points which are not smooth. For instance, it happens in C_1, C_∞ and $B(H)$. The concept of $\varphi-$ Gateaux derivative will be used in order to substitute the usual concept of Gateaux derivative at points which are not smooth in $B(H)$. The concepts of Gateaux derivative and $\varphi-$ Gateaux derivative have also been used in Global minimizing problems, see for instance, [7], [8], [6] and references therein.

Definition 2.1. *Let $(X, \|\cdot\|)$ be an arbitrary Banach space, $x, y \in X$, $\varphi \in [0, 2\pi)$, and $F : X \to \mathbb{R}$. We define the φ-Gâteaux derivative of F at a vector $x \in X$, in $y \in X$ and φ direction by*

$$D_\varphi F(x; y) = \lim_{t \to 0^+} \frac{F(x + te^{i\varphi}y) - F(x)}{t}.$$

We recall (see [3]) that the function $y \mapsto D_{\varphi,x}(y)$ is subadditive,

$$D_{\varphi,x}(y) \leq \|y\| . \tag{2.1}$$

The function $f_{(x,y)}(t) = \|x + te^{i\varphi}y\|$ is convex, $D_{\varphi,x}(y)$ is the right derivative of the function $f_{(x,y)}$ at the point 0 and taking into account the fact that the function $f_{(x,y)}$ is convex $D_{\varphi,x}(y)$ always exists.

The previous simple construction allows us to characterize the best- $\mathcal{C}_1$ approximation to $A \in \mathcal{C}_1$ from M in all Banach spaces without care of smoothness

Note that when $\varphi = 0$ the φ-Gateaux derivative of F at x in direction y coincides with the usual Gateaux derivative of F at x in a direction y given by

$$DF(x;y) = \lim_{t \to 0^+} \frac{F(x+ty) - F(x)}{t}.$$

According to the notation given in [3] we will denote $D_\varphi F(x;y)$ for $F(x) = \|x\|$ by $D_{\varphi,x}(y)$ and for the same function we write $D_x(y)$ for $DF(x;y)$.

The following result has been proved by Keckic in [3].

Theorem 2.1. *The vector y is orthogonal to x in the sense of Birkhoff if and only if*

$$\inf_{\varphi} D_{\varphi,x}(y) \geq 0. \tag{2.2}$$

Now we recall the following theorem proved in [3].

Theorem 2.2. *Let $X, Y \in \mathcal{C}_1(H)$. Then, there holds*

$$D_X(Y) = \operatorname{Re}\{\operatorname{tr}(U^*Y)\} + \|QYP\|_{\mathcal{C}_1},$$

where $X = U|X|$ is the polar decomposition of X, $P = P_{\ker X}$, $Q = Q_{\ker X^*}$ are projections.

The following corollary establishes a characterization of the $\varphi-$ Gateaux derivative of the norm in $\mathcal{C}_1$-classes.

Corollary 2.3. *Let $X, Y \in \mathcal{C}_1(H)$. Then, there holds*

$$D_{\varphi,X}(Y) = \mathrm{Re}\left\{e^{i\varphi}\,\mathrm{tr}(U^*Y)\right\} + \|QYP\|_{\mathcal{C}_1},$$

for all φ, where $X = U|X|$ is the polar decomposition of X, $P = P_{\ker X}$, $Q = Q_{\ker X^*}$ are projections.

3. Main Results

The following Theorem 3.1 has been proved in [5]; for the convenience of the reader we present it and its proof below.

Theorem 3.1. *Let E be a Banach space, M a linear subspace of E, and $A \in E \setminus \overline{M}$. Then the following assertions are equivalent:*

1. *B is a best approximant to A from M;*
2. *for all $Y \in M$, $A - B$ is orthogonal to Y;*
3. $$\inf_{\varphi} D_{\varphi, A-B}(Y) \geq 0, \quad \textit{for all } Y \in M \tag{3.1}$$

Proof. The equivalence between (2) and (3) follows from Theorem 2.1. So we prove the equivalence between (1) and (3). Assume that B is a best approximant to A from M, i.e.,

$$\|A - D\| \geq \|A - B\|, \quad \text{for all } D \in M.$$

Let $\varphi \in [0, 2\pi]$, $t > 0$, and $Y \in M$. Taking $D = B - te^{i\varphi}Y$ in the last inequality gives

$$\|A - B + te^{i\varphi}Y\| \geq \|A - B\|,$$

and so

$$\frac{\|A - B + te^{i\varphi}Y\| - \|A - B\|}{t} \geq 0.$$

Thus, by letting $t \to 0^+$ and taking the infimum over φ we obtain

$$\inf_{\varphi} D_{\varphi,A-B}(Y) \geq 0, \quad \text{for all } Y \in M.$$

Conversely, assume that (3.1) is satisfied. Let $\varphi = 0$ and let $Y \in M$. From the fact that the function $t \mapsto \frac{\|A-B+te^{i\varphi}Y\|-\|A-B\|}{t}$ is nondecreasing on $(0, +\infty)$ we have

$$\frac{\|A - B + Y\| - \|A - B\|}{t} \geq D_{\varphi,A-B}(Y), \quad \text{for all } t > 0, \ Y \in M.$$

Using (3.1) we get

$$\frac{\|A - B + Y\| - \|A - B\|}{t} \geq 0, \quad \text{for all } t > 0, \ Y \in M.$$

Therefore, by taking $t = 1$ and $Y = B + D$, with $D \in M$ (since M is a linear subspace) we get

$$\|A - D\| \geq \|A - B\| \quad \text{for all } D \in M.$$

This ensures that B is a best approximant to A from M and the proof is complete. □

Remark 1. *It is very obvious in Theorem 3.1 that (1) is equivalent to (2)(from the definition of the orthogonality and the best approximant). Rather, it is more important to prove the equivalence between (1) and (3). The same remark applies for Theorem 3.2.*

Using Corollary 2.3 and the previous theorem, we prove the following characterizations of best approximants in $\mathcal{C}_1$-Classes.

Theorem 3.2. *Let M be a subspace of $\mathcal{C}_1(H)$ and $A \in \mathcal{C}_1(H) \setminus \overline{M}$. Then the following assertions are equivalent:*

(i) *B is a best $\mathcal{C}_1(H)$-approximant to A from M:*

(ii) *for all $Y \in M$, $A - B$ is orthogonal to Y;*

(iii)

$$\|QYP\|_{\mathcal{C}_1} \geq |\operatorname{tr}(U^*Y)|, \quad \textit{for all } Y \in M, \tag{3.2}$$

where $A - B = U|A - B|$ is the polar decomposition of $A - B$, $P = P_{\ker(A-B)}$, $Q = Q_{\ker(A-B)^*}$ are projections.

Proof. The equivalence between (ii) and (iii) follows from Corollary 1 in [3]. We have only to prove the equivalence between (i) and (iii). Assume that B is a best $\mathcal{C}_1(H)$-approximant to A from M. Then by the previous theorem we have

$$\inf_{\varphi} D_{\varphi, A-B}(Y) \geq 0, \quad \text{for all } Y \in M,$$

which ensures by Corollary 2.3

$$\inf_{\varphi} \operatorname{Re}\left\{e^{i\varphi} \operatorname{tr}(U^*Y\right\} + \|QYP\|_{\mathcal{C}_1} \geq 0, \quad \text{for all } Y \in M,$$

where $A - B = U|A - B|$ is the polar decomposition of $A - B$ and $P = P_{\ker(A-B)}$, $Q = Q_{\ker(A-B)^*}$ or equivalently

$$\|QYP\|_{\mathcal{C}_1} \geq -\inf_{\varphi} \operatorname{Re}\left\{e^{i\varphi} \operatorname{tr}(U^*Y)\right\}.$$

By choosing the most suitable φ we get

$$\|QYP\|_{\mathcal{C}_1} \geq |\mathrm{tr}(U^*Y|, \quad \text{for all } Y \in M.$$

Conversely, assume that (3.2) is satisfied. Let φ be arbitrary and $Y \in M$. By (3.2) we have

$$\left\|Q\tilde{Y}P\right\|_{\mathcal{C}_1} \geq \left|\mathrm{tr}(U^*\tilde{Y}\right| \geq -\operatorname{Re}\left(\mathrm{tr}(U^*\tilde{Y}\right),$$

with $\tilde{Y} = e^{i\varphi}Y \in M$. Hence,

$$\|QYP\|_{\mathcal{C}_1} \geq -\operatorname{Re}\left(e^{i\varphi}\,\mathrm{tr}(U^*Y\right),$$

for $Y \in M$ and all $\varphi \in [0, 2\pi]$ and so

$$\inf_{\varphi}\left[\|QYP\|_{\mathcal{C}_1} + \operatorname{Re}\left(e^{i\varphi}\,\mathrm{tr}(U^*Y)\right]\right. \geq 0,$$

for $Y \in M$ and all $\varphi \in [0, 2\pi]$. Thus Theorem 3.1 and Corollary 2.3 complete the proof. □

Now we are going to prove the uniqueness of the best approximant. First we need to prove the following proposition. It has its own interest and it will be the key in our proof of the next theorem.

Proposition 3.3. *Let E be a Banach space, M a subspace of E, and $A \in E \backslash \overline{M}$. Assume that B is a best approximant to A from M. Set*

$$\gamma := \inf\left\{D_{\varphi,A-B}(Y);\ \varphi \in [0, 2\pi]; Y \in M,\ \|Y\| = 1\right\}.$$

Then $\gamma \in [0,1]$ *and for all* $Y \in M$*,*

$$\gamma\|Y - B\| \leq \|A - Y\| - \|A - B\|. \tag{3.3}$$

Furthermore, if $\gamma' > \gamma$*, then there exists* $C \in M$ *for which*

$$\gamma'\|C - B\| > \|A - C\| - \|A - B\|.$$

Proof. Since B is a best approximant to A from M, then by Theorem 3.1 we have $\gamma \geq 0$. The fact that $\gamma \leq 1$ follows from the properties of the φ-Gateaux derivative recalled in the Preliminaries. For $\gamma = 0$ the inequality (3.3) is satisfied because B is a best approximant to A from M. Assume now that $\gamma > 0$. By the definition of γ we have for $\varphi = 0$

$$D_{\varphi,A-B}(-Y) \geq \gamma\|Y\|, \quad \text{for all } Y \in M, Y \neq 0.$$

Therefore, for all $t > 0$ we have

$$\frac{\|A - B - tY\| - \|A - B\|}{t} \geq \gamma\|Y\|,$$

for all $Y \in M$, $Y \neq 0$, which is equivalent to

$$\gamma\|tY\| \leq \|A - B - tY\| - \|A - B\|,$$

for all $Y \in M$, $Y \neq 0$. Since M is a linear subspace we get

$$\gamma\|Y - B\| \leq \|A - Y\| - \|A - B\|,$$

for Y belonging to a small ball with center at B, $Y \neq 0$. Since for $Y = 0$ we get $\gamma = 0$ and so the inequality (3.4) is satisfied. Hence

$$\gamma\|Y - B\| \leq \|A - Y\| - \|A - B\|, \quad \text{for all } Y \in M.$$

Assume now that $\gamma' > \gamma$, i.e.,

$$\gamma' > \inf\left\{D_{\varphi,A-B}(Y);\ \varphi \in [0, 2\pi]; Y \in M,\ \|Y\| = 1\right\}.$$

Then there exists $\varphi_0 \in [0, 2\pi]$, $D \in M$ such that $\|D\| = 1$ and

$$\gamma'\|D\| > D_{\varphi_0,A-B}(-D) = \lim_{t\to 0^+} \frac{\|A - B - te^{i\varphi_0}D\| - \|A - B\|}{t}.$$

Consequently, for some t_0 small enough we have

$$\gamma'\|D\| > \frac{\|A - B - t_0e^{i\varphi_0}D\| - \|A - B\|}{t_0},$$

and so

$$\gamma'\|t_0D\| > \|A - B - t_0e^{i\varphi_0}D\| - \|A - B\|.$$

Set $C = B + t_0e^{i\varphi_0}D \in M$. Thus

$$\gamma'\|C - B\| > \|A - C\| - \|A - B\|.$$

This completes the proof. □

Theorem 3.4. *Let M be a subspace of $\mathcal{C}_1(H)$ and $A \in \mathcal{C}_1(H) \setminus \overline{M}$. Let B be a best $\mathcal{C}_1(H)$- approximant to A from M satisfying*

$$\|QYP\|_{\mathcal{C}_1} > |\operatorname{tr}(U^*Y)|, \quad \text{for all } Y \in M, Y \neq 0, \tag{3.4}$$

where $A-B = U|A-B|$ is the polar decomposition of $A-B$, $P = P_{\ker(A-B)}$, $Q = Q_{\ker(A-B)^*}$ are projections. Then B is the unique best $\mathcal{C}_1(H)$-approximant to A from M.

Proof. Assume that (3.4) is satisfied. There exists $\alpha > 0$ such that

$$\|QYP\|_{\mathcal{C}_1} > \alpha > |\operatorname{tr}(U^*Y)|, \quad \text{for all } Y \in M, Y \neq 0. \tag{3.5}$$

Let φ be arbitrary in $[0, 2\pi]$ and $Y \in M$ and put $\tilde{Y} = e^{i\varphi}Y$. Then

$$\alpha > |\operatorname{tr}(U^*\tilde{Y})| \geq -\operatorname{Re}\left(\operatorname{tr}(U^*\tilde{Y})\right) = -\operatorname{Re}\left(e^{i\varphi}\operatorname{tr}(U^*Y)\right).$$

Taking the infimum on φ over $[0, 2\pi]$ yields

$$\alpha \geq \inf_{\varphi}\left[-\operatorname{Re}\left(e^{i\varphi}\operatorname{tr}(U^*Y)\right)\right].$$

This inequality and (3.5) give

$$\|QYP\|_{\mathcal{C}_1} > \inf_{\varphi}\left[-\operatorname{Re}\left(e^{i\varphi}\operatorname{tr}(U^*Y)\right)\right],$$

which is equivalent to

$$\inf_{\varphi}\left[\|QYP\|_{\mathcal{C}_1} + \operatorname{Re}\left(e^{i\varphi}\operatorname{tr}(U^*Y)\right)\right] > 0, \quad \text{for all } Y \in M, Y \neq 0.$$

Now, by Corollary 2.3 and the definition of γ we get $\gamma > 0$. Therefore, by the previous theorem we have

$$\gamma\|Y - B\| \leq \|A - Y\| - \|A - B\|, \quad \text{for all } Y \in M.$$

Assume that C is another best $\mathcal{C}_1(H)$-approximant to A from M such that $C \neq B$. Then

$$\gamma\|C - B\| \leq \|A - C\| - \|A - B\| \leq \|A - B\| - \|A - B\| = 0.$$

This ensures that $\|C - B\| = 0$, which contradicts $C \neq B$. Thus B is the unique best $\mathcal{C}_1(H)$-approximant to A from M. □

References

[1] J. ANDERSON, On normal derivations, *Proc. Amer. Math. Soc.*, **38**(1) (1979), 129–135.

[2] G. BIRKHOFF, Orthogonality in linear metric space, *Duke Math. J.*, **1** (1935), 165–172.

[3] D. KECKIC, Orthogonality in C_1 and C_∞-spaces, *J. Operator Theory*, **51**(1) (2004), 89–104.

[4] F. KITTANEH, Operators that are orthogonal to the range of a derivation, *J. Math. Anal. Appl.*, **203** (1996), 863–873.

[5] S. MECHERI, On the best L^1 approximants, *East Journal on Approximations*, **16**(3) (2004), 383–390.

[6] P.J. MAHER, Commutator approximants, *Proc. Amer. Math. Soc.*, **115** (1992), 995–1000.

[7] S. MECHERI, On minimizing $\|S-(AX-XB)\|_p$, *Serdica Math. J.*, **26**(2) (2000), 119–126.

[8] S. MECHERI, Another version of Maher's Inequality, *Zeitschrift fur und ihre Anwendungen*, **23**(2) (2004), 303–311.

[9] A.M. PINKUS, *On L^1-Approximation*, Cambridge University Press, 1989.

[10] B. SIMON, Trace ideals and their applications, *Lond. Math. Soc Lecture Notes Series* 35, Cambridge University Press, 1979.