SELF-COMMUTATOR APPROXIMANTS

P. J. MAHER

(Communicated by Joseph A. Ball)

Abstract. This paper deals with minimizing \( \| B - (X^*X - XX^*) \|_p \), where \( B \) is fixed, self-adjoint and \( B \in \mathcal{C}_p \), and where \( X \) varies such that \( BX = XB \) and \( X^*X - XX^* \in \mathcal{C}_p, 1 \leq p < \infty \). (Here, \( \mathcal{C}_p \), \( 1 \leq p < \infty \), denotes the von Neumann-Schatten class and \( \| \cdot \|_p \) its norm.) The upshot of this paper is that \( \| B - (X^*X - XX^*) \|_p, 1 \leq p < \infty \), is minimized if, and only if, \( X^*X - XX^* = 0 \), and that the map \( X \rightarrow \| B - (X^*X - XX^*) \|_p, 1 < p < \infty \), has a critical point at \( X = V \) if and only if \( V^*V - VV^* = 0 \) (with related results for normal \( B \) if \( p = 1 \) or \( 2 \)).

1. Introduction

This paper is concerned with approximating an operator by a self-commutator \( X^*X - XX^* \) of operators. We study minimizing the quantity

\[
\| B - (X^*X - XX^*) \|_p, \quad 1 \leq p < \infty,
\]

for fixed \( B \) in \( \mathcal{C}_p \) and varying \( X \) such that \( X^*X - XX^* \in \mathcal{C}_p \). (Here \( \mathcal{C}_p \) denotes the von Neumann-Schatten class with norm \( \| \cdot \|_p \), where \( 1 \leq p < \infty \).)

The related topic of approximation by commutators \( AX -XA \), which has attracted much interest, has its roots in quantum theory. The Heisenberg Uncertainty Principle may be mathematically formulated as saying that there exists a pair \( A, X \) of linear transformations and a (non-zero) scalar \( \alpha \) for which

\[
AX -XA = \alpha I.
\]

Clearly, (1.1) cannot hold for square matrices \( A \) and \( X \). (To see this, just take the trace of both sides of (1.1).) Nor can (1.1) hold for bounded linear operators \( A \) and \( X \): two beautiful proofs of this are due to Wielandt [19] and Wintner [20]. This prompts the question: how close can \( AX -XA \) be to the identity?

Halmos [9, 11 Problem 233] proved that if \( A \) is normal, or if \( A \) commutes with \( AX -XA \), then, for all \( X \) in \( \mathcal{L}(H) \),

\[
\| I - (AX -XA) \| \geq \| I \|.
\]

Anderson [2, Theorem 1.7] generalized Halmos’ inequality (1.2): he proved that if \( A \) is normal and commutes with \( B \), then, for all \( X \) in \( \mathcal{L}(H) \),

\[
\| B - (AX -XA) \| \geq \| B \|.
\]

Received by the editors March 5, 2003 and, in revised form, March 25, 2004.

2000 Mathematics Subject Classification. Primary 47B20, 47A30; Secondary 47B10.

Key words and phrases. Self-commutator, von Neumann-Schatten class.
Maher [15, Theorem 3.2] obtained the $C_p$ variant of Anderson’s result: he proved that if the normal operator $A$ commutes with $B$ and if $B \in C_p$, then, for all $X$ such that $AX -XA \in C_p$,

\[(1.4) \quad \|B - (AX -XA)\|_p \geq \|B\|_p, \quad 1 \leq p < \infty,
\]

with equality if, and for $1 < p < \infty$ only if, $AX -XA = 0$.

Using the technique of [15], Bouli and Cherki [4] and Mercheri [16] studied approximation by “generalized commutators” $AX -XC$. They proved [4, Theorem 2.2], [16, Theorem 3.7] that if $B \in C_p$, if $AB = BC$ and if the pair $(A,C)$ has the Putnam-Fuglede $C_p$ property (meaning that $AX = XC \Rightarrow A^*X = XC^*$ if $X \in C_p$), then, for all $X$ such that $AX -XC \in C_p$,

\[(1.5) \quad \|B - (AX -XC)\|_p \geq \|B\|_p, \quad 1 \leq p < \infty,
\]

with equality if, and for $1 < p < \infty$ only if, $AX -XC = 0$. (Other related work includes that of Berens and Finzel [3], Duggal [5, 6, 12, 13] and Mercheri [17, 18].)

In the above results (1.2), (1.3) and (1.4), the zero commutator is, to use Halmos’ terminology [10], a commutator approximant in $C_p$. We consider the critical points of $F: \mathbb{C} \rightarrow C_p$.

The global result guarantees the existence of global minima. Thus, it says that under the same hypotheses ($BX = XB, B^* = B \in C_p$) for all $X$ such that $X^*X -XX^* \in C_p$, where $1 < p < \infty$, 

\[(1.6) \quad \|B - (X^*X -XX^*)\|_p \geq \|B\|_p
\]

with equality in (1.4) if and only if $X^*X -XX^* = 0$.

There is a similar inequality for the trace norm ($p = 1$) proved by different arguments in Theorem 4.2. Examples 4.1 and 4.2 illustrate, and reinforce, the results. Examples 4.1 and 4.2 show that if the (seemingly restrictive) condition $BX = XB$ is dropped, the conclusions of Theorems 4.1 and 4.2 do not hold. Finally, Example 4.3 shows that for $0 < p < 1$ the inequalities may be reversed.

2. Preliminaries

Let $H$ denote a separable, complex Hilbert space. For details concerning the von Neumann-Schatten classes $C_p$ and norms $\|\cdot\|_p$ see [7, Chapter XI], [18, Chapter 2]. The spaces $C_p$ are examples of 2-sided, self-adjoint ideals. Note $C_p \subseteq C_q$ and $\|\cdot\|_p \geq \|\cdot\|_q$ if $1 \leq p \leq q < \infty$. The Fréchet derivative of some real-valued function

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
F at V is denoted by \( D_V F \) and given by
\[
(D_V F)(S) = \lim_{h \to 0} \frac{F(V + hS) - F(V)}{h}
\]
(provided the R.H. limit exists). We state below the Aiken, Erdos and Goldstein differentiation result.

**Theorem 2.1 (11 Theorem 2.1).** Let the map \( \Phi : C_p \to \mathbb{R}^+ \) be given by \( \Phi : X \to ||X||_p^p \). Then:

(a) for \( 1 < p < \infty \), the map \( \Phi \) is Fréchet differentiable at every \( X \) in \( C_p \) with derivative \( D_X \Phi \) given by
\[
(D_X \Phi)(S) = p \Re \tau ||X||^{p-1}U^*S,
\]
where \( \tau \) denotes trace, \( X = U|X| \) is the polar decomposition of \( X \) and \( S \in \mathcal{C}_p \);

(b) for \( 0 < p \leq 1 \), provided \( \dim H < \infty \), the same result holds at every invertible element \( X \).

Observe that the formula for the R.H.S. of \( (D_X \Phi)(S) \) makes sense: for \( |X|^{p-1}U^* \in \mathcal{C}_1 \) since, by [18 Lemma 2.3.1], \( X \in \mathcal{C}_p \iff |X| \in \mathcal{C}_p \iff |X|^p \in \mathcal{C}_1 \), that is, \( |X|^p = (|X|^{p-1}U^*)(U|X|) \in \mathcal{C}_1 \), whence \( |X|^{p-1}U^* \in \mathcal{C}_1 \) as \( X = U|X| \in \mathcal{C}_p \supset \mathcal{C}_1 \).

### 3. Local theory

The local theory of self-commutator approximation is more complicated than the local theory of commutator approximation [15 Theorem 3.2 (b)]; inevitably so, since differentiating a (non-commutative) product is more complicated than differentiating a sum. The local theory centres on Theorem 3.1.

**Theorem 3.1.** Let \( B \) be self-adjoint in \( \mathcal{C}_p \), where \( 1 < p < \infty \). Let \( \mathcal{S} = \{ X : X^*X - XX^* \in \mathcal{C}_p \} \) and let \( F_p : \mathcal{S} \to \mathbb{R}^+ \) be given by
\[
F_p : X \to ||B - (X^*X - XX^*)||_p^p.
\]
Then if \( V \) is a critical point of \( F_p \) such that \( V^*V - VV^* = 0 \) it follows that \( BV = VB \).

**Proof.** As we shall use this proof in that of Theorem 3.2, we adopt the hypothesis that \( V^*V - VV^* = 0 \) only at the last step.

**Step 1.** Let \( V \) be in \( \mathcal{S} \) so that \( B - (V^*V - VV^*) \in \mathcal{C}_p \). (Observe that the set \( \mathcal{S} = \{ X : X^*X - XX^* \in \mathcal{C}_p \} \) properly contains \( \mathcal{C}_p \), for if \( X \in \mathcal{C}_p \) then \( X \in \mathcal{S} \) and, e.g., \( I \in \mathcal{S} \) but \( I \not\in \mathcal{C}_p \).) Let \( \mathcal{S} \) consist of all operators \( S \) for which \( B - [(V + S)^*(V + S) - (V + S)(V + S)^*] \in \mathcal{C}_p \). (Thus, \( \mathcal{S} \) also properly contains \( \mathcal{C}_p \).) Let \( \Phi : X \to ||X||_p^p \) and \( \Psi : X \to B - (X^*X - XX^*) \). Then \( F_p = \Phi \circ \Psi \). Let \( S \) be arbitrary in \( \mathcal{S} \). By considering \( F_p(V + S) - F_p(V) \) it follows from the definition **2.1** of the derivative that the Fréchet derivative of \( F_p \) at \( V \) is given by
\[
(D_V F_p)(S) = (D_{B - (V^*V - VV^*)}\Phi)(VS^* + SV^* - V^*S - S^*V).
\]

Let \( B - (V^*V - VV^*) = U_1|B - (V^*V - VV^*)| \) be the polar decomposition of \( B - (V^*V - VV^*) \) (so that \( \text{Ker} U_1 = \text{Ker} |B - (V^*V - VV^*)| \)). Then by Theorem **2.1** on writing
\[
Y = U_1|B - (V^*V - VV^*)|^p - 1,
\]
we have
\[(D_V F_p)(S) = p \Re \tau[Y^* (VS^* + SV^* - V^* S - S^* V)]\]
for all operators \(S\) in \(S\). Note that \(Y^* = |B - (V^* V - VV^*)|^{p-1} U_1^* \in C_1\) (cf. comments after the statement of Theorem 2.1). Therefore, as \(\Re \tau(T) = \Re \tau(T^*)\) for all \(T\) in \(C_1\), we have \(\Re \tau[Y^* VS^* - Y^* S^* V] = \Re \tau[SV^* Y - V^* SY]\). Hence, by the invariance of trace [18, Theorem 2.2.4],
\[(3.1) \quad (D_V)(F_p)(S) = p \Re \tau[(V^* Y - YV^* + V^* Y^* - Y^* V^*) S].\]

**Step 2.** Let \(V\) be a critical point of \(F_p\), so that \((D_V F_p)(S) = 0\) for all operators \(S\) in \(S\). Take \(S = f \otimes g\), where \(f\) and \(g\) are arbitrary vectors in the underlying Hilbert space \(H\). (The rank one operator \(x \rightarrow \langle x, f \rangle g\), where \(x \in H\), is denoted \(f \otimes g\). Note that \(\tau[T(f \otimes g)] = \langle Tg, f \rangle\) for \(T\) in \(\mathcal{L}(H)\); cf. [18, pp. 73, 90].) Then by (3.1)
\[\Re \langle (V^* Y - YV^* + V^* Y^* - Y^* V^*) g, f \rangle = 0\]
which, since \(f\) and \(g\) are arbitrary, means that \(V^* Y - YV^* + V^* Y^* - Y^* V^* = 0\), that is,
\[(3.2) \quad (\Re Y)V = V(\Re Y)\].

**Step 3.** Suppose now that \(B\) is self-adjoint. Then \(B - (V^* V - VV^*) = U_1 \{B - (V^* V - VV^*)\}\) is self-adjoint. Hence \(U_1\) is self-adjoint and commutes with \(|B - (V^* V - VV^*)|\), and hence \(Y = U_1 \{B - (V^* V - VV^*)|^{p-1}\}\) is self-adjoint. Therefore, (3.2) says that
\[(3.3) \quad YV = VY,\]
that is,
\[(3.4) \quad U_1 \{B - (V^* V - VV^*)|^{p-1} V = VU_1 \{B - (V^* V - VV^*)|^{p-1}.\]

**Step 4. Assertion:** \(V\) satisfies
\[(3.5) \quad BV - (V^* V - VV^*)V = VB - V(V^* V - VV^*).\]

To prove this assertion, note that equality (3.5) is equivalent to
\[(3.6) \quad U_1 \{B - (V^* V - VV^*)|V = VU_1 \{B - (V^* V - VV^*)|\].\]
Write \(Z = |B - (V^* V - VV^*)|^{p-1}\). Then (3.6) says that
\[(3.7) \quad U_1 Z^{\frac{1}{p-1}} V = VU_1 Z^{\frac{1}{p-1}}.\]
To prove (3.7) we approximate both sides of (3.7) by polynomials in \(Z\). The function \(f: t \rightarrow t^{1/(p-1)}\), where \(t \in \sigma(Z) \subseteq \mathbb{R}^+\), can be approximated uniformly by a sequence \((p_i)\) of polynomials without constant term (for \(f(0) = 0\)). Therefore, (3.7) will follow by the functional calculus (cf. [15] p. 998)) from \(U_1 p_i(Z)V = VU_1 p_i(Z)\) and this, in turn, will follow from
\[(3.8) \quad U_1 Z^n V = VU_1 Z^n.\]
To prove (3.8) note that in the polar decomposition of \(B - (V^* V - VV^*)\), \(\ker U_1 = \ker |B - (V^* V - VV^*)| = \ker Z\) (by the spectral theorem). Hence, \((\ker U_1)^\perp = \text{Ran } Z\) and \(U_1^* U_1\), the orthogonal projection onto \((\ker U_1)^\perp\), satisfies \(ZU_1^* U_1 Z = Z^2\). (It is simplest to write, where necessary, \(U_1^*\), even though \(U_1\) is self-adjoint.) Since \(Y = U_1 Z\), then (3.3) says that \(U_1 Z V = VU_1 Z\) and, as \(Y = Y^*\), \(ZU_1^* V = VZU_1^* Z = VZ^2\). Thus,
\[Z^2 V = ZU_1^* (U_1 Z V) = (ZU_1^* V)U_1 Z = VZU_1^* U_1 Z = VZ^2.\]
Taking positive square roots of $Z^2$ [15 Theorem 1.7.7 (vi)] we get $VZ = ZV$. Returning now to (3.3): for $n = 1$, (3.3) is the equality $U_1 Z V = V U_1 Z$ (which is (3.3)), and the inductive step follows from $Z V = Z V$. This proves the assertion.

**Step 5.** If, finally, $V^* V - V V^* = 0$ then (3.5) forces $B V = V B$. □

**Note.** The proof in Theorem 3.1 of the implication, $V$ is a critical point of $F_p$ $\Rightarrow$ $B V - (V^* V - V V^*) V = V B - V (V^* V - V V^*)$, only holds for $1 < p < \infty$ since the argument involving the function $f : t \to t^{1/(p-1)}$, where $0 \leq t < \infty$, only holds for $1 < p < \infty$.

Observe also that for non-self-adjoint $B$, in the case $p = 2$, it follows from equality (3.2) of the proof of Theorem 3.1 that if $V$ is a critical point of $F_2$ such that $V^* V - V V^* = 0$, then $(\Re B) V = V (\Re B)$. But this latter equality does not force $B V = V B$ even if $B$ is normal: witness: $B = \left[ \begin{array}{cc} 1 & 1 \\ 1 & 1 \end{array} \right]$.

**Theorem 3.2.** Let $B$ be self-adjoint, let $B X = X B$ and let $B$ be in $C_p$. Let $\mathcal{G} = \{ X : X^* X - X X^* \in C_p \}$ and let $F_p : \mathcal{G} \to \mathbb{R}^+$ be given by

$$F_p : X \to \| B - (X^* X - X X^*) \|_p^p.$$ 

Then:

(a) for $1 < p < \infty$, the map $F_p$ has a critical point at $V$ if and only if $V^* V - V V^* = 0$;

(b) for $0 < p \leq 1$, the map $F_p$ has a critical point at $V$ if $V^* V - V V^* = 0$ provided $\dim H < \infty$ and $B - (V^* V - V V^*)$ is invertible;

(c) for $p = 2$, the same result as in (a) holds if the condition on $B$ of self-adjointness is replaced by normality.

**Proof.** (a) Let $V$ be a critical point of $F_p$. Then equality (3.5) of the proof of Theorem 3.1 holds. As $B V = V B$ then $(V^* V - V V^*) V = V (V^* V - V V^*)$, whence, by Klinecke-Shirokov [11 Problem 323], $V^* V - V V^*$ is quasinilpotent and, hence, being self-adjoint, zero.

Conversely, let $V$ satisfy $V^* V - V V^* = 0$. Then the partial isometries $U_1$ and, say, $U$, occurring in the polar decompositions of $B - (V^* V - V V^*)$ and of $B$, coincide. Thus, $Y = U |B|^{p-1} \in C_1$ so that $Y^* = |B|^{p-1} U^{*p-1} \in C_1$.

We first prove that $Y^* V - V Y^* = 0$. Since $V^*$ and $V$ commute with $B$ they commute with $|B|$ (and hence with $|B|^{p-1}$). So, $|B| U^* V = |B| V U^*$. It follows that

$$\text{Ran}(U^* V - V U^*) \subseteq \text{Ker} |B| = \text{Ker} |B|^{p-1}.$$ 

Hence, since $|B|^{p-1} V = V |B|^{p-1}$, therefore $Y^* V - V Y^* = 0$.

Similarly, from the equality $|B| U^* V^* = |B| V U^* V^*$ it follows that $Y^* V^* - V Y^* = 0$. Hence, $V^* Y - V Y^* + V^* Y^* - Y^* V^* = 0$. Substitute into equality (3.1) of the proof of Theorem 3.1 (the expression for $D_V F_p$, the Fréchet derivative of $F_p$ at $V$). As $Y S \in C_1$ and $Y^* S \in C_1$, it follows by (3.1) that $(D_V F_p)(S) = 0$ for all $S$ in $\mathcal{L}(H)$.

(b) follows immediately from (a) as in [15 Theorem 3.2 (c)].

(c) Let $B$ be normal. If $V$ is a critical point of $F_2$, then equality (3.2) of the proof of Theorem 3.1 says that

$$[\Re B - (V^* V - V V^*)] V = V [\Re B - (V^* V - V V^*)].$$ 

Since $V$ commutes with $B$ then (by Fuglede) $V$ commutes with $B^*$ and hence with $\Re B$. The result now follows as in (a).
The proof of the converse implication (\(V \) satisfies \(V^*V - VV^* = 0 \Rightarrow V \) is a critical point of \(F_2\)) depends only on \(V \) and \(V^*\) commuting with \(B\) and is therefore the same as in (a). \(\square\)

Indeed, the proof in (a) of the implication, \(V^*V - VV^* = 0 \Rightarrow V \) is a critical point of \(F_p\), for \(1 < p < \infty\), holds (via Fuglede’s Theorem) for normal \(B\).

4. Global theory

**Theorem 4.1.** Let \(B \) be self-adjoint, let \(BX = XB\) and let \(B \) be in \(\mathcal{C}_p\). Let \(\mathcal{S} = \{X : X^*X - XX^* \in \mathcal{C}_p\}\). Then, if \(X \in \mathcal{S}\),

(a) for \(1 < p < \infty\),

\[
\|B - (X^*X - XX^*)\|_p \geq \|B\|_p
\]

with equality holding in (4.1) if and only if \(X^*X - XX^* = 0\);

(b) for \(p = 2\), the same result as in (a) holds if \(B \) is assumed normal rather than self-adjoint.

**Proof.** (a) First, suppose the operators \(X \) in \(\mathcal{S}\) are contractions, i.e., such that \(\|X\| \leq 1\). Suppose also that the underlying space \(H\) is finite dimensional. (The argument here is analogous to [14 Theorem 5.7].) The set of contractions is bounded and closed (for the condition \(X^*X - I \leq 0\) characterises the contractions, and the map \(X \rightarrow X^*X\) is continuous; cf. [11 Problem 129]). Hence, \(\mathcal{S}\) is compact since \(H\) is finite dimensional. Therefore, the continuous map \(F_p : X \rightarrow \|B - (X^*X - XX^*)\|_p\) is bounded, attains its bounds and thus has a global minimizer, and hence a critical point, at \(V\), say. Since, by Theorem [3,2 (a)], \(V^*V - VV^* = 0\), therefore

\[
\|B - (X^*X - XX^*)\|_p \geq \|B\|_p.
\]

Conversely, if equality holds in (4.2) for some point \(X\), then that \(X\) is a global minimizer, hence a critical point of \(F_p\), whence, by Theorem [3,2 (a)], \(X^*X - XX^* = 0\).

The extension to infinite-dimensional \(H\) is similar to [1 Theorem 3.5]. As the operator \(B\) is compact and normal there exists a basis \(\{\phi_i\}\) of \(H\) consisting of eigenvectors of \(B\) which may be ordered such that \(|\lambda_1| \geq |\lambda_2| \geq \ldots\) where \(B\phi_i = \lambda_i \phi_i\) (and where the eigenvalues are repeated according to multiplicity). Let

\[
H_k = \text{Span}\{\phi_i : B\phi_i = \lambda_i \phi_i, i = 1, \ldots, k\}.
\]

\(H_k\) is invariant under \(X\) and \(X^*\); for if \(\phi_i\) is an eigenvector of \(B\), then so are \(X\phi_i\) and \(X^*\phi_i\) with the same eigenvalues (since \(B\) commutes with \(X\) and \(X^*\)). Therefore, if \(E_k\) denotes the orthogonal projection onto \(H_k\), then \(E_kX = XE_k\). Hence \(E_kBE_k\) commutes with \(E_kXE_k\) (and with \(E_kX^*E_k\)) and hence, by the finite-dimensional inequality (4.2) applied to the contraction \(E_kXE_k\),

\[
\|(E_kBE_k) - [(E_kXE_k)^*(E_kXE_k) - (E_kXE_k)(E_kXE_k)^*]\|_p \geq \|E_kBE_k\|_p,
\]

that is, \(\|E_k[B - (X^*X - XX^*)]E_k\|_p \geq \|E_kBE_k\|_p\). Now let \(k \rightarrow \infty\). Then \(E_k \rightarrow I\) and from [3 Lemma 2] (cf. [1 Theorem 3.5]), it follows that inequality (4.2) holds for infinite-dimensional \(H\).

The condition that the operator \(X\) in \(\mathcal{S}\) is a contraction may now be lifted. Let \(X\) be arbitrary in \(\mathcal{S}\); then by applying the inequality (4.2) to the contraction \(X/\|X\|\), the result immediately follows.
Theorem 4.2. Let $B$ be normal, let $BX = XB$ and let $B$ be in $C_1$. Then, if $X^*X - XX^* \in C_1$,

$$\|B - (X^*X - XX^*)\|_1 \geq \|B\|_1.$$  

Proof. Let $B = U|B|$ be the polar decomposition of $B$. As $U$ is a partial isometry, so is $U^*$, and so $\|U^*\| = 1$. Since, by [18, Theorem 2.3.10], $\|U^*T\|_1 \leq \|U^*\|\|T\|_1 = \|T\|_1$ for arbitrary $T$ in $C_1$. Then, by [18 Lemma 2.3.3],

$$\|B - (X^*X - XX^*)\|_1 \geq \|B - U^*(X^*X - XX^*)\|_1$$

(4.3)

where

$$\tau[|B| - U^*(X^*X - XX^*)] = \sum_n \langle |B| - U^*(X^*X - XX^*) | \phi_n, \phi_n \rangle$$

for an arbitrary orthonormal basis $\{\phi_i\}$ of $H$.

Take $\{\phi_i\}$ as the orthonormal basis of $H$ consisting of eigenvectors of the compact normal operator $|B|$. Let $\{\psi_m\}$ be an orthonormal basis of $\text{Ker} |B|$ and let $\{\xi_k\}$ be an orthonormal basis of $(\text{Ker} |B|)^\perp$ consisting of eigenvectors of $|B|$. Thus, $\{\phi_n\} = \{\psi_m\} \cup \{\xi_k\}$. Then $\sum_m \langle \{ B - U^*(X^*X - XX^*) \} \psi_m, \psi_m \rangle = 0$ because $\psi_m \in \text{Ker} |B| = \text{Ker} U$; and $\sum_k \langle B \xi_k, \xi_k \rangle = \|B\|_1$. Further, since $BX = XB$ and $BX^* = X^*B$, it can be checked that $\langle U^*X^*X \xi_k, \xi_k \rangle = \langle X^*U^*X \xi_k, \xi_k \rangle$. Hence, by the invariance of trace [18 Theorem 2.2.4 (v)], $\tau[U^*(X^*X - XX^*)] = 0$. Therefore, by (4.3),

$$\|B - (X^*X - XX^*)\|_1 \geq \tau(|B|) = \|B\|_1.$$  

□

In the special case when $B$ is positive, the proof of the trace norm result is simple and does not require the commutativity condition.

Theorem 4.3. Let $B$ be positive and be in $C_1$. Then, if $X^*X - XX^* \in C_1$,

$$\|B - (X^*X - XX^*)\|_1 \geq \|B\|_1.$$  

Proof. Let $B$ be positive so that $B = |B|$. Then, by [18 Lemma 2.3.3] and the linearity of trace [18 Theorem 2.2.4],

$$\|B - (X^*X - XX^*)\|_1 \geq \tau[B - (X^*X - XX^*)]$$

$$= \tau[B] = \tau[|B|] = \|B\|_1.$$  

□

If $B$ and $X$ do not commute, then either $\|B - (X^*X - XX^*)\|_p \geq \|B\|_p$, for $1 \leq p < \infty$, is reversed (Example 4.1) or $\|B - (X^*X - XX^*)\|_p = \|B\|_p$ without $X^*X - XX^* = 0$ (Example 4.2).

Example 4.1. Take $B = \begin{bmatrix} 3 & 0 \\ 0 & -3 \end{bmatrix}$ and $X = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$ so that $B = B^*$ and $BX \neq XB$. For $1 \leq p < \infty$, as $\|T\|_p = \sum_i s_i^p(T)$, where $s_i(T)$ denotes the $i$th eigenvalue of $T$, we get, for $1 \leq p < \infty$,

$$\|B - (X^*X - XX^*)\|_p = 1^p + 1^p < 3^p + 3^p = \|B\|_p^p.$$
Example 4.2. Take $X = f \otimes g$ and $B = f \otimes f$, where $f \neq g$ and $\|f\| = \|g\| = 1$, so that $B = B^* (\geq 0)$ and $BX \neq XB$. Then $X^*X - XX^* = f \otimes f - g \otimes g \neq 0$ and, as $\|f \otimes g\|_p = \|f\| \|g\| \|f\|_p$ for $1 \leq p < \infty$, we have

$$\|B - (X^*X - XX^*)\|_p = \|g \otimes g\|_p = \|f \otimes f\|_p = \|B\|_p.$$ 

Finally, the inequality $\|B - (X^*X - XX^*)\|_p \geq \|B\|_p$ may be reversed for $0 < p < 1$.

Example 4.3. Take $B = [\begin{smallmatrix} 3 & 0 \\ 0 & 3 \end{smallmatrix}] (\geq 0)$ and $X = [\begin{smallmatrix} 3 & \sqrt{3} \\ -3 & 3 \end{smallmatrix}]$ so that $B = B^*$ and $BX = XB$. Then for $0 < p < 1$ we have the strict inequality

$$\|B - (X^*X - XX^*)\|_p^p = 6^p < 2 \cdot 3^p = \|B\|_p^p.$$ 

(This example also shows that even if the conditions of Theorem 4.2 are met, a minimizer of $\|B - (X^*X - XX^*)\|_1$ need not be normal: for here

$$\|B - (X^*X - XX^*)\|_1 = \|B\|_1,$$

yet $X^*X - XX^* \neq 0$.)

References


DEPARTMENT OF MATHEMATICS, MIDDLESEX UNIVERSITY, THE BURROUGHS, LONDON NW4 4BT, UNITED KINGDOM

E-mail address: p.maher@mdx.ac.uk