Strong unicity criterion in some space of operators

Grzegorz Lewicki

Abstract. Let $X$ be a finite dimensional Banach space and let $Y \subset X$ be a hyperplane. Let $\mathcal{L}_Y = \{ L \in \mathcal{L}(X,Y) : L|_Y = 0 \}$. In this note, we present sufficient and necessary conditions on $L_0 \in \mathcal{L}_Y$ being a strongly unique best approximation for given $L \in \mathcal{L}(X)$. Next we apply this characterization to the case of $X = l^p_\infty$ and to generalization of Theorem I.1.3 from [12] (see also [13]).

Keywords: best approximation, strongly unique best approximation, approximation in spaces of linear operators

Classification: Primary 41A65, 41A52, 41A35

0. Introduction.

Let $W$ be a normed linear space and let $V \subset W$ be its nonempty subset. An element $v \in V$ is called a best approximation to $w \in W$ iff

\begin{equation}
\|w - v\| = \text{dist}(w,V) = \inf\{\|w - y\| ; y \in V\}.
\end{equation}

If an element $v$ satisfies additionally

\begin{equation}
\|w - y\| \geq \|w - v\| + r \cdot \|y - v\| \quad \text{with a constant } r > 0
\end{equation}

independent of $y \in V$,

then $v$ is said to be a strongly unique best approximation (briefly SUBA) to $w \in V$.

The theory of strong uniqueness has its origin in the following result of Newman and Shapiro [11]. Given a compact Hausdorff space $T$, let $C(T, \mathbb{K})$ denote the Banach space of either complex ($\mathbb{K} = \mathbb{C}$) or real-valued ($\mathbb{K} = \mathbb{R}$) continuous functions on $V$. If $V$ is a Haar subspace of $C(T, \mathbb{K})$, then for every $w \in C(T, \mathbb{K})$, one can find a constant $r > 0$ such that the best approximation $v \in V$ satisfies one of the following inequalities:

\begin{equation}
\|w - y\| \geq \|w - v\| + r \cdot \|y - v\|, \quad \text{for } y \in V
\end{equation}

if $\mathbb{K} = \mathbb{R}$, and

\begin{equation}
\|w - y\|^2 \geq \|w - v\|^2 + r \cdot \|y - v\|^2, \quad \text{for } y \in V
\end{equation}

in the complex case.
The significance of this notion can be illustrated by Cheney’s observation that strong unicity of an optimal element yields the continuity of metric projection (see [6]). One can see that the proof of the convergence of the Remez algorithm depends, in fact, on strong unicity (for an extended version see [9]). For more precise information about strong unicity the reader is referred to [3], [4], [8], [11], [14], [15].

In this note we will investigate strong unicity in the case \( W = \mathcal{L}(X) \), the space of all linear operators going from a finite dimensional real Banach space \( X \) itself, and \( V = \mathcal{L}_Y(X, Y) = \{ L \in \mathcal{L}(X, Y) : L \, |_Y = 0 \} \) (we will write \( \mathcal{L}_Y \) for brevity), where \( Y \subset X \) is a hyperplane. We characterize strong unicity in terms of the functionals from \( X^* \) or \( Y^* \), which is more convenient for applications. Next we apply this characterization to the case of \( X \) being an arbitrary three dimensional Banach space and to \( X = l_1^\infty \). In particular, we generalize Theorem I.1.3 from [12] (see also [13]) and Theorem 2.5b) from [10].

Now we introduce some notations which will be used in this note. By \( S_X \) we will denote the unit sphere in a Banach space \( X \). The symbol \( \text{ext} \) stands for the set of all extremal points of \( S_X \). Given \( L \in \mathcal{L}(X) \), we write \( \mathcal{P}_Y(L) = \{ L_0 \in \mathcal{L}_Y : \| L - L_0 \| = \text{dist} (L, \mathcal{L}_Y) \} \). In this note, if nothing special is assumed, \( X \) will denote a finite dimensional real Banach space and \( f \) a functional from \( X^* \setminus \{0\} \). If \( Y \subset X \) is a linear subspace and \( A \subset X^* \) then \( A \, |_Y \) stands for a set of all restrictions of functionals from \( A \). In the sequel we will use the following

**Theorem 0.1** (see [10, Theorem 2.3]). Assume \( X \) is a reflexive space and \( Y \) is a Banach space both over the same field \( \mathbb{K} \) (\( \mathbb{K} = \mathbb{R} \) or \( \mathbb{K} = \mathbb{C} \)). Denote by \( \mathcal{K}(X, Y) \) the space of all compact operators going from \( X \) into \( Y \) and let \( \mathcal{V} \subset \mathcal{K}(X, Y) \) be a convex set. For given \( K \in \mathcal{K}(X, Y) \) and \( V \in \mathcal{V} \) put

\[
(0.5) \quad \text{crit}_V^*(K - V) = \{ h \in \text{ext} \, S_{Y^*} : \| h \circ (K - V) \| = \| K - V \| \}
\]

and for every \( h \in \text{crit}_V^*(K - V) \) define

\[
(0.6) \quad A_h = \{ x \in \text{ext} \, S_X : h(K - V)x = \| K - V \| \}.
\]

Then we have:

(a) \( V \in \mathcal{P}_Y(K) \) (the set of all best approximants to \( K \) in \( \mathcal{V} \)) if and only if for every \( U \in \mathcal{V} \) there exists \( h \in \text{crit}_V^*(K - V) \) with \( \inf \{ r e(h(U - V)x) : x \in A_h \} \leq 0 \).

(b) \( V \) is a SUBA to \( K \) in \( \mathcal{V} \) with a constant \( r > 0 \) if and only if for every \( U \in \mathcal{V} \) there exists \( h \in \text{crit}_V^*(K - V) \) with \( \inf \{ r e(h(U - V)x) : x \in A_h \} \leq r \cdot \| U - V \|. \)

I. The main result.

We start with two preliminary remarks.

**Remark 1.1.** For \( L \in \mathcal{L}(X) \) let us set

\[
(1.1) \quad \text{crit} (L) = \{ x \in S_X : \| Lx \| = \| L \| \}.
\]

Assume \( L_0 \in \mathcal{P}_Y(L) \), \( Y = \ker f \), \( \| f \| = 1 \) and \( \| L - L_0 \| > \| L \, |_Y \| \). Put

\[
(1.2) \quad C_{L-L_0} = \{ x \in \text{crit} (L - L_0) : f(x) > 0 \}.
\]
Then $C_{L-L_0}$ is a nonempty closed set, $C_{L-L_0} \cap -C_{L-L_0} = \emptyset$ and $C_{L-L_0} \cup -C_{L-L_0} = \text{crit}(L-L_0)$.

**Proof:** It is clear that the set $A = \{x \in \text{crit}(L-L_0) : f(x) \geq 0\}$ is closed. Since $\text{dist}(L, L_Y) > \|L \mid_Y \|$ and $Y$ is a hyperplane,

$$A = \{x \in \text{crit}(L-L_0) : f(x) > 0\},$$

which proves that $C_{L-L_0}$ is closed. The fact $\text{crit}(L-L_0) \cap Y = \emptyset$ implies immediately that $C_{L-L_0} \cup -C_{L-L_0} = \text{crit}(L)$. By (1.2) $C_{L-L_0} \cap -C_{L-L_0} = \emptyset$.

**Remark 1.2.** Let $L \in \mathcal{L}(X)$, $\text{dist}(L, \mathcal{L}_Y) > \|L \mid_Y \|$, $L_0 \in \mathcal{L}_Y$. Define

$$D_{L-L_0} = \{h \in \text{crit}^*_X(L-L_0) : C_{L-L_0} \cap A_h \neq \emptyset\}$$

(see (0.6)) and if $L \in \mathcal{L}(X,Y)$,

$$D^Y_{L-L_0} = \{h \in \text{crit}^*_Y(L-L_0) : C_{L-L_0} \cap A_h \neq \emptyset\}.$$

Then $D_{L-L_0}$ ($D^Y_{L-L_0}$ resp.) is a compact set, $D_{L-L_0} \cap -D_{L-L_0} = \emptyset$ ($D^Y_{L-L_0} \cap -D^Y_{L-L_0} = \emptyset$ resp.).

**Proof:** Assume $h \in \text{cl}(D_{L-L_0})$ and let $\{h_n\} \subset D_{L-L_0}$, $h_n \rightarrow h$. By (1.3), for every $n \in \mathbb{N}$ there exists $x_n \in C_{L-L_0} \cap A_{h_n}$, i.e. $h_n(L-L_0)x_n = \|L-L_0\|$. Passing to the subsequence, if necessary, we may assume $x_n \rightarrow x$. By Remark 1.1, $x \in C_{L-L_0}$. Note that $h(L-L_0)x = h_n(L-L_0)x + (h-h_n)(L-L_0)x = h_n(L-L_0)x_n + h_n(L-L_0)(x-x_n) + (h-h_n)(L-L_0)x$. Since the last two terms tend to 0 as $n \rightarrow \infty$, $h(L-L_0)x = \|L-L_0\|$ and consequently $x \in A_h$. Since $x \in C_{L-L_0}$, by (1.3) $h \in D_{L-L_0}$. Note that $\text{dist}(L, \mathcal{L}_Y) > \|L \mid_Y \|$ implies $D_{L-L_0} \cap -D_{L-L_0} = \emptyset$. The proof for the set $D^Y_{L-L_0}$ goes on in the same manner, so we omit it.

Now we state the main result of this note.

**Theorem 1.3.** Assume $L \in \mathcal{L}(X)$ and let $Y = \ker f$, $\|f\| = 1$. Assume furthermore that $\text{dist}(L, \mathcal{L}_Y) > \|L \mid_Y \|$ and let $L_0 \in \mathcal{L}_Y$. Then the following conditions are equivalent:

(a) $L_0$ is a SUBA to $L$ in $\mathcal{L}_Y$ ($L_0 \in \mathcal{P}_Y$ resp.),
(b) $0 \in \text{int conv } D_{L-L_0} \mid_Y$ ($0 \in \text{conv } D_{L-L_0} \mid_Y$ resp.).

**Proof:** Assume $L_0$ is a SUBA to $L$ in $\mathcal{L}_Y$ and let $0 \notin \text{int conv } D_{L-L_0} \mid_Y$. It means that there exists $\psi \in Y^{**}$ with $\psi(h) \geq 0$ for every $h \in D_{L-L_0} \mid_Y$ (we may assume $\|\psi\| = 1$). Since $Y$ is finite dimensional, $\psi = y$ for some $y \in S_Y$. Define $L_1 = f(\cdot) \cdot y$ and note that $L_1 \in \mathcal{L}_Y$. By (1.3) and Remark 1.2, for every $h \in D_{L-L_0}$ we have

$$\inf\{h(L_1x) : x \in A_h\} = \inf\{f(x) \cdot h(y) : x \in A_h\} = h(y) \cdot \inf\{f(x) : x \in A_h\} \geq 0 > -r \cdot \|L_1\| \text{ for every } r > 0.$$
By Theorem 0.1 $L_0$ is not a SUBA to $L$ in $\mathcal{L} |_Y$; a contradiction. Since by Remark 1.2 the set $D_{L-L_0}$ is compact and consequently $\text{conv} \, D_{L-L_0}$ is also compact, the proof of the second case goes on the same line.

To prove the converse, let us define a function $g : S_Y \to \mathbb{R}$ by

$$g(y) = \inf \{g_h(y) : h \in D_{L-L_0} \} \text{ for } y \in S_Y;$$

where $g_h(y) = \inf \{f(x) \cdot h(y) : x \in A_h \}$. Note that the function $S_Y \ni y \mapsto f(x) \cdot h(y)$ is continuous and consequently the functions $g_h$ and $g$ are upper-semicontinuous.

Now assume $0 \in \text{int } \text{conv} \, D_{L-L_0} |_Y$. It means that for every $y \in S_Y$ there exists $h \in D_{L-L_0}$ with $h(y) < 0$. (If no, then $D_{L-L_0} |_Y \subset \{ h \in Y^* : h(y) \geq 0 \}$ for some $y \in S_Y$ and consequently $\text{int } \text{conv} \, D_{L-L_0} |_Y \subset \{ h \in Y^* : h(y) > 0 \}$. But $0 \in \text{int } \text{conv} \, D_{L-L_0}$; a contradiction.) Since $\text{dist} \, (L, \mathcal{L}_Y) > \|L |_Y \|$ and $Y$ is a hyperplane, for every $y \in S_Y$, $g(y) < 0$. Since $g$ is upper-semicontinuous, the value $\gamma = \max \{ g(y) : y \in S_Y \}$ is attained in some point $y_0 \in S_Y$ and consequently $\gamma < 0$. We show that $L_0$ is a SUBA to $L$ in $\mathcal{L}_Y$ with $r = -\gamma$. To do this, fix $L_1 \in \mathcal{L}_Y \setminus \{ 0 \}$. It is clear that $L_1 = f(\cdot) \cdot y_1$ for some $y_1 \in Y \setminus \{ 0 \}$. Put $y_2 = y_1/\|y_1\|$, fix $\varepsilon > 0$ and take $h \in D_{L-L_0}$ with $g_h(y_2) < g(y_2) + \varepsilon$. Note that

$$g_h(y_2) = \inf \{ f(x) \cdot h(y_2) : x \in A_h \} = \inf \{ h(L_1 x)/\|y_1\| : x \in A_h \} \leq g(y_2) + \varepsilon \leq -r + \varepsilon, \quad \text{which gives } \inf \{ h(L_1 x) : x \in A_h \} \leq -(r - \varepsilon) \cdot \|L_1\|.$$

Following Theorem 0.1 and Remark 1.2, $L_0$ is SUBA to $L$ in $\mathcal{L}_Y$ with $r - \varepsilon$ for every $\varepsilon > 0$ and consequently with $r$. The proof is complete.

**Remark 1.4.** If $L \in \mathcal{L}(X,Y)$ then the set $D_{L-L_0}$ in Theorem 1.3 can be replaced by $D^Y_{L-L_0}$ (see (1.4)).

As an immediate consequence of Theorem 1.3 we get

**Corollary 1.5.** Assume $L \in \mathcal{L}(X), L_0 \in \mathcal{P}_Y(L), \|L - L_0\| > \|L |_Y \|$. Then the set $D_{L-L_0} |_Y$ is linearly dependent. If $L \in \mathcal{L}(X,Y)$, the same holds for $D^Y_{L-L_0}$.

Reasoning as in [10, Theorem 2.5] we may show

**Remark 1.6.** The constant $r$ defined in Theorem 1.3 is the best possible.

Now we will point out when the assumption $\text{dist} \, (L, \mathcal{L}_Y) > \|L |_Y \|$ is fulfilled.

**Remark 1.7.** Assume $X$ is a Banach space and let $Y \subset X$ be its complemented subspace. Let $\mathcal{P}(X,Y) = \{ P \in \mathcal{L}(X,Y) : P |_Y = \text{id} \}$. Take $P_0 \in \mathcal{P}(X,Y)$ and note that $\text{dist} \, (P_0, \mathcal{L}_Y) = \inf \{ \| P \| : P \in \mathcal{P}(X,Y) \} = \lambda(X,Y)$.

In many cases of hyperplanes, $\text{dist} \, (P_0, \mathcal{L}_Y) > \|P_0 |_Y \| = 1$ (see e.g. [2], [5]).

It is well known (see e.g. [12]) that if $X$ is not a Hilbert space then there exists a hyperplane $Y$ in $X$ satisfying $\lambda(X,Y) > 1$. If $\text{dist} \, (P_0, \mathcal{L}_Y) > 1$ then it is easy to show that $\text{dist} \, (L, \mathcal{L}_Y) > \|L |_Y \|$ if $\|L - P_0\| < \text{dist} \, (P_0, \mathcal{L}_Y) - 1$.

Now we show an estimation from above of the number $\text{dist} \, (L, \mathcal{L}_Y)$. 

Proposition 1.8. Assume $X$ is a Banach space and let $Y$ be its complemented subspace. Then for every $L \in \mathcal{L}(X,Y)$

$$\|L |_Y \| \leq \text{dist} (L, \mathcal{L}_Y) \leq \lambda(X,Y) \cdot \|L |_Y \|.$$

Proof: Fix $L \in \mathcal{L}(X,Y)$ and $\varepsilon > 0$. Take $P_\varepsilon \in \mathcal{P}(X,Y)$ with $\|P_\varepsilon\| < \lambda(X,Y) + \varepsilon$ and put $L_\varepsilon = L \circ (I - P_\varepsilon)$. It is clear that $L_\varepsilon \in \mathcal{L}_Y$. Compute,

$$\|L - L_\varepsilon\| = \|L - L \circ (I - P_\varepsilon)\| = \|L \circ P_\varepsilon\| \leq \|L |_Y \| \cdot \|P_\varepsilon\|,$$

which gives the desired result. \qed

Corollary 1.9. Assume that $\lambda(X,Y) = 1$. Then

$$\text{dist} (L, \mathcal{L}_Y) = \|L |_Y \| \quad \text{for every} \quad L \in \mathcal{L}(X,Y).$$

In particular if there exists $P_0 \in \mathcal{P}(X,Y)$ with $\|P_0\| = 1$, then the operator $L_0 = L \circ (I - P_0) \in P_Y(L)$.

Since in the case when $Y$ is a hyperplane we have $\lambda(X,Y) \leq 2$ (for more precise results see [1], [7], [12, p. 84], we immediately get

Corollary 1.10. Assume $Y \subset X$ is a hyperplane. Then

$$\|L |_Y \| \leq \text{dist} (L, \mathcal{L}_Y) \leq 2 \cdot \|L |_Y \| \quad \text{for every} \quad L \in \mathcal{L}(X,Y).$$

II. Applications.

Now we apply Theorem 1.3 to generalize Theorem I.1.3 from [12] (see also [13]).

Theorem 2.1. Assume $X$ is a three dimensional Banach space and let $Y \subset X$ be a hyperplane. Assume furthermore $L \in \mathcal{L}(X,Y)$, $\text{dist} (L, \mathcal{L}_Y) > \|L |_Y \|$. Then there exists $L_0 \in \mathcal{L}_Y$ which is a SUBA to $L$ in $\mathcal{L}_Y$.

Proof: Since $\mathcal{L}_Y$ is a finitely dimensional linear space, the set $\mathcal{P}_Y(L)$ is nonempty. Take an arbitrary $L_0 \in \mathcal{P}_Y(L)$. By Theorem 1.3 and Remark 1.4 it is sufficient to show that $0 \in \text{int \ convex \ } D_{L - L_0}\ appointment (1.4)$. Assume on the contrary that it is not true. Following Theorem 1.3, $0 \in \text{conv \ } D_{L - L_0}^\bigcap Y$. Since $\dim Y = 2$, $0 = \alpha \cdot h_1 + (1 - \alpha) \cdot h_2$, where $h_1, h_2 \in D_{L - L_0}^\bigcap Y$ and $\alpha \in (0,1)$. Since $\|h_1\| = \|h_2\| = 1$, we easily get $\alpha = 1/2$. Consequently $h_1 = -h_2$ which gives $h_1 \in D_{L - L_0}^\bigcap Y \cap -D_{L - L_0}^\bigcap Y$; a contradiction with Remark 1.2. \qed

Remark 2.2. The assumption $\text{dist} (L, \mathcal{L}_Y) > \|L |_Y \|$ in Theorem 2.1 is essential. Take e.g. $X = l_\infty^2$, $Y = \text{ker} f$, $f = (1/2, 1/2, 0)$. It is easy to check that the operators $P_1 = Id - f(\cdot) \cdot (2, 0, 0)$ and $P_2 = Id - f(\cdot) \cdot (0, 2, 0) \in \mathcal{P}(X,Y)$, $P_1 \neq P_2$, $\|P_1\| = \|P_2\| = 1$. Consequently the set $P_Y(P_1) \supset \{0, P_1 - P_2\}$ and strong unicity does not hold.
Remark 2.3. The assumption $\dim X = 3$ in Theorem 2.1 is essential. Take e.g. $X = l^1_\infty$, $Y = \ker f$, $f = (1/3, 1/3, 1/3, 0)$. It is well known (see [5] or [12]) that $\lambda(X,Y) = 4/3$. Take $P_0 \in \mathcal{P}(X,Y)$ with $\|P_0\| = 4/3$ (the formula for such a projection is given for example in [12, p. 104]). Then $\text{dist}(P_0, \mathcal{L}_Y) = \lambda(X,Y) = 4/3 > 1 = \|P_0\|_Y$. By Theorem 2.5b) of [10], $0$ is not a SUBA to $P_0$ in $\mathcal{L}_Y$.

Now we use Theorem 1.3 to extend Theorem 2.5b) from [10].

**Theorem 2.4.** Assume $X = l^n_\infty$, $Y = \ker f$, $\|f\|_1 = 1$. Let $L \in \mathcal{L}(X)$ and let $\text{dist}(L, \mathcal{L}_Y) > \|L\|_Y$ (by Theorem 1.3 from [5] and Remark 1.7 such operators exist if and only if $|f_i| < 1/2$ for $i = 1, \ldots, n$). If $|f_i| > 0$ for $i = 1, \ldots, n$ then there exists $L_0 \in \mathcal{L}_Y$ which is a SUBA to $L$ in $\mathcal{L}_Y$.

**Proof:** Since $\mathcal{L}_Y$ is a finitely dimensional space the set $\mathcal{P}_Y(L)$ is nonempty. Fix $L_0 \in \mathcal{P}_Y(L)$ and let $D_{L-L_0} = \{\phi_1, \ldots, \phi_k\}$ where $\phi_i = \pm e_{j(i)}$ for $i = 1, \ldots, k$. By Theorem 1.3, $0 \in \text{conv} D_{L-L_0}$. Hence $0 = \sum_{i=1}^l \lambda_i \cdot \phi_i$, $l \leq k$, $\lambda_i > 0$, $\sum_{i=1}^l \lambda_i = 1$. Since $\dim Y = n - 1$, by Carathéodory’s Theorem we can assume $l \leq n$. We will show that $l = n$. To do this, by Corollary 1.5, it is sufficient to show that for each $i \in \{1, \ldots, n\}$ the set $E_i = \{e_1, \ldots, e_{i-1}, e_{i+1}, \ldots, e_n\}$ is total over $Y$. So assume $\sum_{i=1}^n f_i \cdot y_i = 0$ and $e_j(y) = 0$ for $j \neq i$. It means that $y_j = 0$ for $j \neq i$ and $f_i \cdot y_i = 0$. Since $f_i \neq 0$, $y_i = 0$. Consequently $l = n$ and $0 \in \text{int conv} D_{L-L_0}$. By Theorem 1.3, $L_0$ is a SUBA to $L$ in $\mathcal{L}_Y$. The proof is complete.

Note that Remark 2.3 shows that the assumption $f_i \neq 0$ for $i = 1, \ldots, n$ in Theorem 2.4 is essential.

**References**


**Department of Mathematics, Jagiellonian University, Reymonta 4, 30–059 Kraków, Poland**

*(Received November 7, 1991)*