Extensions of linear operators from hyperplanes of $l_\infty^{(n)}$

Marco Baronti, Vito Fragnelli, Grzegorz Lewicki

Abstract. Let $Y \subset l_\infty^{(n)}$ be a hyperplane and let $A \in \mathcal{L}(Y)$ be given. Denote

$$\begin{aligned} \mathcal{A} &= \{L \in \mathcal{L}(l_\infty^{(n)}, Y) : L\,|\,Y = A\} \text{ and} \\ \lambda_A &= \inf\{\|L\| : L \in \mathcal{A}\}. \end{aligned}$$

In this paper the problem of calculating of the constant λ_A is studied. We present a complete characterization of those $A \in \mathcal{L}(Y)$ for which $\lambda_A = \|A\|$. Next we consider the case $\lambda_A > \|A\|$. Finally some computer examples will be presented.

Keywords: linear operator, extension of minimal norm, element of best approximation, strongly unique best approximation

Classification: 41A35, 41A52, 41A65, 41A55

1. Introduction

Let X be a normed space and let $Y \subset X$ be a linear subspace. For given $A \in \mathcal{L}(Y)$ set

$$\mathcal{A}(X,Y) = \{L \in \mathcal{L}(X,Y) : L\,|_Y = A\} \tag{1.1}$$

and if $\mathcal{A}(X,Y)$ is nonempty,

$$\lambda_A(X,Y) = \inf\{\|L\| : L \in \mathcal{A}(X,Y)\}; \tag{1.2}$$

$$\mathcal{A}_0(X,Y) = \{L \in \mathcal{A}(X,Y) : \|L\| = \lambda_A(X,Y)\}. \tag{1.3}$$

In the case of $A = id_Y$, the set $\mathcal{A}(X,Y)$ corresponding to A (which may be empty) consists of all linear, continuous projections from X onto Y. We will denote it by $\mathcal{P}(X,Y)$. Note that the constant $\lambda_{id_Y}(X,Y)$ plays an essential role in the estimate of $\lambda_A(X,Y)$ because of the following inequality:

$$\|A\| \leq \lambda_A(X,Y) \leq \|A\| \cdot \lambda_{id_Y}(X,Y). \tag{1.4}$$

Moreover, if the set $\mathcal{P}(X,Y)$ is nonempty and $P \in \mathcal{P}(X,Y)$, then $A \circ P$ belongs to $\mathcal{A}(X,Y)$ for every $A \in \mathcal{L}(Y)$. By (1.4), if $\lambda_{id_Y}(X,Y) = 1$ then $\lambda_A(X,Y) = \|A\|$. It is worth saying that the case $\lambda_A(X,Y) > \|A\|$ is much more complicated (for examples of couples (X,Y) with $\lambda_{id_Y}(X,Y) > 1$ see [1], [2] and references in [4]).

In this paper the problem of calculating $\lambda_A(X,Y)$ and the determination of the set $\mathcal{A}_0(X,Y)$ is investigated. We consider the case $X = l_\infty^{(n)}$ (the space R^n with the maximum norm) and $Y \subset X$ being a hyperplane.

In Section 2 we will be concerned with the case $\lambda_{A(X,Y)} = \|A\|$. We prove a complete characterization of those $A \in \mathcal{L}(Y)$ for which $\lambda_A(X,Y) = \|A\|$. This characterization leads to an effective method of a determination of an element from the set $\mathcal{A}_0(X,Y)$.

In Section 3 we deal with the case $\lambda_A(X,Y) > \|A\|$. We prove, under some nonrestrictive assumptions on Y (see 2.5) that $\mathcal{A}_0(X,Y)$ consists of exactly one element, which means that there exists exactly one extension L_0 of A of minimal norm. (Since X is finite dimensional, the set $\mathcal{A}_0(X,Y)$ is nonempty.) Moreover, this extension is strongly unique, i.e.

$$\|L\| \;\geq\; \|L_0\| \;+\; r\|L - L_0\|$$

for every $L \in \mathcal{A}(X,Y)$ with a constant $r > 0$ depending only on $A \in \mathcal{L}(Y)$. Next, we present as in Section 2 an effective method of calculating L_0 and $\lambda_A(X,Y)$ in this case.

Section 4 deals with some computer experiments.

Now we present notations and a terminology which will be frequently used. In this paper, unless otherwise stated, X will stand for the space $l_\infty^{(n)}$ and Y be a hyperplane in X. By $\mathcal{L}(Y)$ ($\mathcal{L}(X,Y)$ resp.) we denote the space of all linear, continuous operators from Y into Y (from X into Y resp.). We will write $S_X(a,r)$ ($S_{X^*}(a,r)$ resp.) for the sphere with a center $a \in X$ and a radius r ($a \in X^*$ resp.). If $a = 0$ and $r = 1$ we abbreviate $S_X(0,1)$ ($S_{X^*}(0,1)$ resp.) to S_X (S_{X^*} resp.). For the same reason we will write $\mathcal{A}$, $\mathcal{A}_0$, λ_A instead of $\mathcal{A}(X,Y)$, $\mathcal{A}_0(X,Y)$, $\lambda_A(X,Y)$. The symbol $\mathrm{ext}\,(A)$ will stand for the set of all extremal points of a set A.

In this paper we assume that $\lambda_{id} > 1$, since, by (1.4), in the opposite case $\lambda_A = \|A\|$ for every $A \in \mathcal{L}(Y)$. Moreover, if P is a projection from X onto Y of norm 1, then $A \circ P \in \mathcal{A}_0$. So the problem of calculating an extension of minimal norm reduces to finding a projection of norm 1 which is well known in this case (see [2]). By [2], $\lambda_{id} > 1$ if and only if

$$|\,f_i\,| \;<\; 1/2 \tag{1.5}$$

for every $f = (f_1, \ldots, f_n) \in S_{X^*}$ with $Y = \ker f$.

Now we present some results which will be frequently used in this paper. Denote for $x \in X$, $x = (x_1, \ldots, x_n)$ and $i \in \{1, \ldots, n\}$

$$e_i(x) \;=\; x_i. \tag{1.6}$$

Following [3, Theorem 2.2.a]

$$\mathrm{ext}\,(S_{(\mathcal{L}(X,Y))^*}) \;=\; \{e_i \otimes x : x \in \mathrm{ext}\,(S_X),\; i = 1, \ldots, n\}, \tag{1.7}$$

where

$$(e_i \otimes x)(L) \;=\; e_i(Lx) \tag{1.8}$$

for every $L \in \mathcal{L}(X, Y)$.

Now assume that X is a normed space (real or complex) and let $Y \subset X$ be an n-dimensional linear subspace. Given $x \in X$ set

$$E(x) \;=\; \{f \in \operatorname{ext}(S_{X^*}) : f(x) = \|x\|\}. \tag{1.9}$$

A set $U = \{f_1, \dots, f_k\} \subset E(x)$ is called an I-set if and only if there exist positive numbers $\lambda_1, \dots, \lambda_k$ with

$$0 \;=\; \sum_{i=1}^{k} \lambda_i f_i\,|_Y \tag{1.10}$$

and any essential subset of U does not have this property. If $k = n+1$, the I-set U is called regular. The notion of I-set was introduced in [7]. The role of regular I-sets illustrates

Theorem 1.1 (see [7, Theorem 5.8]). *Assume X is a normed space and let $Y \subset X$ be an n-dimensional subspace. Let $x \in X \setminus Y$ and let $y_0 \in Y$ be the best approximation to x from Y. If $E(x - y_0)$ contains a regular I-set, then y_0 is the strongly unique best approximation to x from Y, i.e.*

$$\|x - y\| \;\geq\; \|x - y_0\| \;+\; r\|y - y_0\|$$

for any $y \in Y$, where the constant $r > 0$ is independent of $y \in Y$.

2

We start with the following

Proposition 2.1. *Let $Y = \ker(f)$ for some $f \in S_{X^*}$ satisfying $1/2 > |\,f_i\,| > 0$ for $i = 1, \dots, n$. Assume $A \in \mathcal{L}(Y)$ and $\lambda_A = \|A\|$. Denote for each $i_0 \in \{1, \dots, n\}$, $i \in \{1, \dots, n\} \setminus \{i_0\}$ $y_i^{i_0} = (y_i^{i_0}(1), \dots, y_i^{i_0}(n)) \in Y$ by*

$$y_i^{i_0}(j) \;=\; \begin{cases} 0 & \textit{if } \ j \neq i_0, i \\ 1 & \textit{if } \ j = i \\ -f_i/f_{i_0} & \textit{if } \ j = i_0. \end{cases} \tag{2.1}$$

If $L \in \mathcal{A}_0$, then for each $i_0 \in \{1, \dots, n\}$, $i \in \{1, \dots, n\} \setminus \{i_0\}$ there exists $g_i \in X^*$ with

$$g_i\,|_Y = e_i \circ A, \|g_i\| \leq \|A\| \tag{2.2}$$

such that

$$Lx = \sum_{i \neq i_0} g_i(x) y_i^{i_0} \quad \textit{for} \quad x \in X. \tag{2.3}$$

Conversely, if L has property (2.3) *for some $i_0 \in \{1, \ldots, n\}$ with g_i satisfying* (2.2) *and $\|e_{i_0} \circ L\| \leq \|A\|$, then $L \in \mathcal{A}_0$.*

PROOF: Fix $i_0 \in \{1, \ldots, n\}$, $L \in \mathcal{A}_0$ and let $U_{i_0} = \sum_{i \neq i_0} (e_i \circ L)(\cdot) y_i^{i_0}$. We show that $U_{i_0} = L$. Note that for $j \neq i_0$ and $x \in X$

$$(e_j \circ U_{i_0})(x) = \sum_{k \neq i_0} e_k(Lx) e_j(y_k^{i_0}) = e_j(Lx).$$

Since $f_{i_0} \neq 0$ and $U_{i_0}x, Lx \in Y$, $U_{i_0}x = Lx$. Put, for $i \neq i_0$, $g_i = e_i \circ L$. Since $L \in \mathcal{A}_0$ and $\lambda_A = \|A\|$, $g_i\,|\,_Y = e_i \circ A$ and $\|g_i\| = \|e_i \circ L\| \leq \|A\|$ for $i \neq i_0$. Now assume that L has property (2.3) with g_i satisfying (2.2) and $\|e_{i_0} \circ L\| \leq \|A\|$. Hence for $i, j \neq i_0$

$$e_j(Ly_i^{i_0}) = \sum_{k \neq i_0} g_k(y_i^{i_0}) e_j(y_k^{i_0}) = e_j(Ay_i^{i_0})$$

and consequently $L \in \mathcal{A}$. Note that $e_i \circ L = g_i$ for $i \neq i_0$. Since $\|L\| = \max_{i=1,\ldots,n} \|e_i \circ L\|$, we immediately get that $L \in \mathcal{A}_0$.

Note that $\mathcal{A}_0$ is a compact convex set. Hence, by the Krein-Milman Theorem, the set $\operatorname{ext}(\mathcal{A}_0)$ is nonempty. Moreover, we have

Proposition 2.2. *Let $A \in \mathcal{L}(Y)$ and let $\lambda_A = \|A\|$. If $L \in \operatorname{ext}(\mathcal{A}_0)$, then*

$$\operatorname{card}\{i : \|e_i \circ L\| = \|L\|\} \geq n - 1.$$

PROOF: Suppose that there exists $L \in \operatorname{ext}(\mathcal{A}_0)$ such that

$$\operatorname{card}\{i : \|e_i \circ L\| = \|L\|\} < n - 1.$$

Let $\|e_{i_1} \circ L\| < \|L\| = \|A\|$ and $\|e_{i_2} \circ L\| < \|A\|$ for $i_1, i_2 \in \{1, \ldots, n\}$, $i_1 \neq i_2$. It is easy to check that $L = \sum_{i \neq i_1} (e_i \circ L)(\cdot) y_i^{i_1}$. Define for $\lambda \in R$ $L_\lambda = \sum_{i \neq i_1} g_i(\ ) y_i^{i_1}$, where

$$g_i = \begin{cases} e_i \circ L & \text{if } i \neq i_2 \\ e_i \circ L + \lambda f & \text{if } i = i_2. \end{cases} \tag{2.4}$$

Note that $L_\lambda \in \mathcal{A}$, $L_\lambda \neq L$ for $\lambda \neq 0$ and $L = (L_{-\lambda} + L_\lambda)/2$ for every $\lambda \in R$. We show that $L_\lambda \in \mathcal{A}_0$ for $|\,\lambda\,|$ sufficiently small. It is clear that for $j = i_1, i_2$,

$$\begin{aligned} \|e_j \circ L_\lambda\| &= \|e_j \circ (L + \lambda f(\cdot) y_{i_2}^{i_1})\| \\ &\leq \|e_j \circ L\| + |\,\lambda\,|\ \|y_{i_2}^{i_1}\|. \end{aligned}$$

For $j \neq i_1, i_2$,

$$\|e_j \circ L_\lambda\| = \|e_j \circ L\| \leq \|A\|.$$

Since $\|e_{i_1} \circ L\| < \|A\|$ and $\|e_{i_2} \circ L\| < \|A\|$, the proof is complete. $\square$

Proposition 2.3. *Assume* $Y = \ker f$, *where* f *satisfies* (1.5) $f_i \neq 0$ *for* $i = 1, \dots, n$ *and*

$$f(x) \neq 0 \quad \textit{for every} \quad x \in \operatorname{ext}(S_{X^*}). \tag{2.5}$$

Let $A \in \mathcal{L}(Y)$. *If* $\|e_i \circ A\| = \|A\|$, *then there exists exactly one* $g \in S_{X^*}(0, \|A\|)$ *with* $g\,|_Y = e_i \circ A$. *If* $\|e_i \circ A\| < \|A\|$, *then there exist exactly two functionals* $g_1, g_2 \in S_{X^*}(0, \|A\|)$ *with* $g_j\,|_Y = e_i \circ A$ *for* $j = 1, 2$.

Proof: Without loss of generality we can assume that $\|A\| = 1$. First we consider the case $\|e_i \circ A\| = \|A\|$. Note that by (1.5) $\|e_i\,|_Y\| = 1$ (an element $y = (y_1, \dots, y_n)$, where

$$y_j = \begin{cases} (-\operatorname{sgn} f_j / \sum_{k \neq i} |\, f_k\,|) f_i & \text{if } j \neq i \\ 1 & \text{if } j = i \end{cases}$$

belongs to S_Y). By [5], $\operatorname{ext}(S_{Y^*}) \subset \{\pm e_j\,|_Y\}_{j=1,\dots,n}$. Now take $y^0 = (y_1^0, \dots, y_n^0) \in \operatorname{ext}(S_Y)$ with $(e_i \circ A) y^0 = \|e_i \circ A\| = 1$. By (2.5), there exists exactly one $i_0 \in 1, \dots, n$ with $|\, y_{i_0}^0\,| < 1$. Following [6, Lemma 1.1, p. 166]

$$e_i \circ A = \sum_{j \in J_i \subset \{1,\dots,n\} \setminus i_0} \lambda_j y_j^0 e_j\,|_Y. \tag{2.6}$$

where $\lambda_j > 0$ and $\sum_{j \in J_i} \lambda_j = 1$. We show that (2.6) is the unique expression of $e_i \circ A$ as a convex combination of points from the set $\operatorname{ext}(S_{Y^*})$ (with strictly positive coefficients). Indeed, let $e_i \circ A = \sum_{j \in J_1} \gamma_j y_j^0 e_j\,|_Y$, where $0 < \gamma_j$, $\sum_{j \in J_1} \gamma_j = 1$. Since $|\, y_{i_0}^0\,| < 1$, $J_1 \subset \{1, \dots, n\} \setminus \{i_0\}$. Hence, because $\{e_j\,|_Y\}_{j \neq i_0}$ is a basis of Y^*, $J_1 = J_i$ and $\gamma_j = \lambda_j$. Now define

$$g = \sum_{j \in J_i} \lambda_j y_j^0 e_j.$$

It is evident that $\|g\| = 1$ and $g\,|_Y = e_i \circ A$. We show that g is the unique extension of $e_i \circ A$ which preserves the norm. To do this, take $h \in S_{X^*}$, $h\,|_Y = e_i \circ A$. By [6, Lemma 1.1, p. 166]

$$h = \sum_{j \in Z} \gamma_j y_j^0 e_j,$$

where $\gamma_j > 0$, $\sum_{j \in Z} \gamma_j = 1$. Since $h(y^0) = (e_i \circ A)(y^0)$, $Z \subset \{1, \dots, n\} \setminus \{i_0\}$. Consequently, reasoning as above, we get $Z = J_i$ and $\lambda_j = \gamma_j$ for $j \in J_i$. Now assume $\|e_i \circ A\| < \|A\| = 1$. Applying the first part of the proof, we can show that there exists exactly one $h_i \in X^*$, $\|h_i\| = \|e_i \circ A\|$ and $h_i\,|_Y = e_i \circ A$. Note that if $g \in X^*$ and $g\,|_Y = e_i \circ A$, then $g = h_i + \lambda f$ for some $\lambda \in R$. Since $\|h_i\| < \|A\| = 1$, the line $h_i + \lambda f$ intersects S_{X^*} in exactly two points g_1, g_2. The proof is complete. □

Now for given $A \in \mathcal{L}(Y)$ and $i \in \{1,\dots,n\}$ denote

$$\operatorname{crit}_A = \{i \in \{1,\dots,n\} : \|e_i \circ A\| = \|A\|\} \tag{2.7}$$

and

$$\mathcal{E}_i = \{g \in S_{X^*}(0,\|A\|) : g\,|_Y = e_i \circ A\}. \tag{2.8}$$

Following Proposition 2.3 $\operatorname{card}(\mathcal{E}_i) = 1$ if $i \in \operatorname{crit}_A$ and $\operatorname{card}(\mathcal{E}_i) = 2$ in the opposite case. Let us set

$$\begin{aligned} \mathcal{D} = \{L \in \mathcal{L}(X,Y) : L = \sum_{i\neq i_0} g_i(\cdot)y_i^{i_0} \\ \text{for some } i_0 \in \{1,\dots,n\}, g_i \in \mathcal{E}_i\} \end{aligned} \tag{2.9}$$

($y_i^{i_0}$ is defined by (2.1)). Now we can state the main result of this section.

Theorem 2.4. *Suppose* $Y = \ker f$, $f = (f_1,\dots,f_n)$, *where* f *satisfies* (1.5), (2.5) *and* $f_i \neq 0$ *for* $i = 1,\dots,n$. *Let* $A \in \mathcal{L}(Y)$. *Then* $\lambda_A = \|A\|$ *if and only if there exists* $L \in \mathcal{D}$, $\|L\| = \|A\|$.

Proof: It is easy to check that $\mathcal{D} \subset \mathcal{A}$. Hence if $\|L\| = \|A\|$ for some $L \in \mathcal{D}$, then $\lambda_A = \|A\|$. If $\lambda_A = \|A\|$ take any $L \in \operatorname{ext}(\mathcal{A}_0)$. By Proposition 2.2, there exists $i_0 \in \{1,\dots,n\}$ such that $\|e_i \circ L\| = \|A\|$ for $i \neq i_0$. It is clear that $L = \sum_{i\neq i_0}(e_i \circ L(\cdot))y_i^{i_0}$. Hence $L \in \mathcal{D}$. The proof is complete. □

Propositions 2.2, 2.3 and Theorem 2.4 provide a method which permits to check if $\lambda_A = \|A\|$ or $\lambda_A > \|A\|$ for any $A \in \mathcal{L}(Y)$. This method consists of the following steps:

(a) calculating the set $\operatorname{ext}(S_Y)$;
(b) calculating the norm of $e_i \circ A$ for $i = 1,\dots,n$ using the set $\operatorname{ext}(S_Y)$;
(c) choosing for each $i \in \{1,\dots,n\}$ $y_i \in \operatorname{ext}(S_Y)$ satisfying $(e_i \circ A)y_i = \|A\|$;
(d) finding for $i = 1,\dots,n$ the unique functional $h_i \in X^*$ such that $h_i\,|_Y = e_i \circ A$ and $\|e_i \circ A\| = \|h_i\|$;
(e) finding the set $\mathcal{E}_i$ for each $i \in \{1,\dots,n\} \setminus \operatorname{crit}_A$;
(f) checking the norms of operators from the set $\mathcal{D}$.

Of course the method presented above is complicated and the point (f) needs a "good" algorithmic solution. But there exist operators $A \in \mathcal{L}(Y)$ for which we can check a simpler way, if $\lambda_A = \|A\|$.

Example 2.5. Assume $\|e_i \circ A\| = \|A\|$ for each $i \in \{1,\dots,n\}$. Then the set $\mathcal{D}$ consists of exactly one element.

Example 2.6. Assume $L \in \mathcal{L}(X,Y)$ is represented by a matrix $[l(i,j)]_{i,j=1,\ldots,n}$. Put $A = L\,|_Y$ and assume that there exists $i_0 \in \{1,\ldots,n\}$ such that for each $j \in \{1,\ldots,n\}$

$$\sum_{i=1}^{n} |\,(-f_i/f_{i_0})l(j,i_0) + l(j,i)\,| \;\leq\; \|A\|. \tag{2.10}$$

Then $\lambda_A = \|A\|$.

PROOF: Fix $i_0 \in \{1,\ldots,n\}$ satisfying (2.10). Define $L_1 = \sum_{i\neq i_0} e_i(\cdot)Ly_i^{i_0}$. It is clear that $L_1\,|_Y = L\,|_Y = A$. Moreover, it is easy to check that

$$\|L_1\| \;=\; \max_{j=1,\ldots,n} \sum_{i\neq i_0} |\,e_j(Ly_i^{i_0})\,|\,.$$

Observe that $|\,e_j(Ly_i^{i_0})\,| \;=\; |\,(-f_i/f_{i_0})l(j,i_0) + l(j,i)\,|$. Following (2.10), the proof is complete. □

3

We start with the following

Theorem 3.1. *Assume that $f \in S_{X^*}$, $f = (f_1,\ldots,f_n)$ satisfies* (1.5), (2.5) *and let $f_i \neq 0$ for $i = 1,\ldots,n$. Assume furthermore that $A \in \mathcal{L}(Y)$ and let $\lambda_A > \|A\|$. Define*

$$\mathcal{L}_Y \;=\; \{L \in \mathcal{L}(X,Y) : L \;=\; f(\cdot)\cdot y,\; y \in Y\}. \tag{3.1}$$

If $L_0 \in \mathcal{A}_0$ then there exists in $E(L_0)$ (*see* 1.9) *a regular I-set* (*see* 1.10) *with respect to $\mathcal{L}_Y$.*

PROOF: Let $L_0 \in \mathcal{A}_0$. It is easy to verify that

$$\|L_0\| \;=\; \operatorname{dist}(L_0, \mathcal{L}_Y).$$

Hence, by [6, Theorem 1.1, p. 170] $0 \in \operatorname{conv}(E(L_0))\,|\,_{\mathcal{L}_Y}$, i.e.

$$0 \;=\; \sum_{i=1}^{k} \lambda_i\varphi_i\,|\,\mathcal{L}_Y,$$

where $\lambda_i > 0$ and $\sum_{i=1}^{k} \lambda_i = 1$. Assume $k \in N$ is a minimal number for which the above equality is satisfied. If we show that $k = n$, then $\{\varphi_1,\ldots,\varphi_n\}$ will be the required regular I-set. By the Carathéodory Theorem, we may assume $k \leq n$ ($\dim \mathcal{L}_Y = n-1$). By (1.7), $\varphi_i = e_{j(i)} \otimes x_i$, where $j(i) \in \{1,\ldots,n\}$ and $x_i \in$

ext (S_{X^*}). There is no loss of generality in assuming $j(1) \le j(2) \le, \dots, \le j(k)$. First we show that $j(1) = 1$. Suppose on the contrary that $j(1) > 1$ and put

$$E_1 \;=\; \{i : j(i) \;=\; j(1)\}.$$

Then

$$0 \;=\; \sum_{i=1}^{k} \lambda_i (e_{j(i)} \otimes x_i) \,|\, \mathcal{L}_Y \;=\; \sum_{i \in E_1} \lambda_i (e_{j(1)} \otimes x_i) \,|\, \mathcal{L}_Y + \sum_{i \notin E_1} \lambda_i (e_{j(i)} \otimes x_i) \,|\, \mathcal{L}_Y .$$

Put

$$L_{j(1)} \;=\; f(\cdot) y^1_{j(1)}. \tag{3.2}$$

($y^1_{j_1}$ is defined by (2.1)). Note that if $i \notin E_1$, $j(1) < j(i)$. Hence for each $i \in E_1$

$$(e_{j(i)} \otimes x_i)(L_{j(1)}) \;=\; f(x_i) e_{j(i)}(y^1_{j(1)}) \;= 0.$$

Consequently,

$$0 \;=\; \sum_{i \in E_1} \lambda_i (e_{j(i)} \otimes x_i)(L_1) \;=\; \sum_{i \in E_1} \lambda_i f(x_i),$$

since $e_{j(1)}(y^1_{j(1)}) \;=\; 1$. To get a contradiction, we show that for every $i \in E_1$ $f(x_i) > 0$ or for every $i \in E_1$ $f(x_i) < 0$. By (2.5), for every $i \in E_1$ $f(x_i) \neq 0$. So suppose that there exist $i_1, i_2 \in E_1$ with $f(x_{i_1}) < 0$ and $f(x_{i_2}) > 0$. Hence it is easy to show that

$$(e_{j_1} \otimes y) \;=\; \|L_0\|$$

for some $y \in S_Y$. But this contradicts the assumption $\lambda_A > \|A\|$. So we have proved $j(1) = 1$. To end the proof of the theorem, we check that a map $i \to j(i)$ is surjective. If no, there exists $i_0 \in \{1, \dots, n\}$ with $j(i) \neq i_0$ for $i = 1, \dots, k$. Since $j(1) = 1$, $i_0 > 1$. Put $I_1 = \{i \in \{1, \dots, k\} : j(i) = 1\}$. An easy computation shows that

$$0 \;=\; \sum_{i=1}^{k} (e_{j(i)} \otimes x_i)(L_{i_0}) \;=\; (-f_{i_0}/f_1) \sum_{i \in I_1} \lambda_i f(x_i).$$

Reasoning as in the first part of the proof we get $f(x_i) > 0$ for each $i \in I_1$ or $f(x_i) < 0$ for each $i \in I_1$; a contradiction. Hence the map $i \to j(i)$ is surjective and consequently $k = n$. The proof is complete. $\square$

Reasoning as in Theorem 3.1 we can prove

Theorem 3.2. *Let $L \in \mathcal{L}(X)$ and let $L_0 \in \mathcal{P}_{\mathcal{L}_Y}(L)$ (the set of all best approximants to L from $\mathcal{L}_Y$). Assume* $\operatorname{dist}(L, \mathcal{L}_Y) > \|L\,|_Y\|$. *Then the set $E(L - L_0)$ contains a regular I-set.*

By Theorem 1.1 we get immediately

Corollary 3.3. *Let A, L_0, f be such as in Theorem* 3.1. *Then there exists $r > 0$ such that for every $L \in \mathcal{A}$*

$$\|L\| \;\geq\; \|L_0\| \;+\; r\|L - L_0\|.$$

In particular the set $\mathcal{A}_0$ consists of exactly one element.

Note that the assumption $\lambda_A > \|A\|$ in Theorem 3.1 and Corollary 3.3 is essential because of

Example 3.4. Let $n = 3$ and let $f = (1/3, 1/3, 1/3)$, $Y = \ker f$. Define $L \in \mathcal{L}(X, Y)$ as a matrix

$$L \;=\; \begin{pmatrix} a & -a & 0 \\ -a/2 & a/2 & 0 \\ -a/2 & a/2 & 0 \end{pmatrix},$$

where a is a fixed positive number. Put $A = L\,|_Y$. It is easy to verify that

$$\operatorname{ext}(S_Y) \;=\; \{\pm(1,-1,0), \pm(1,0,-1), \pm(0,1,-1)\}.$$

Hence $\|L\| = \|A\|$ and consequently $\lambda_A = \|A\|$. Consider for $\delta \in R$ an operator L_δ defined by a matrix

$$L_\delta \;=\; \begin{pmatrix} a & -a & 0 \\ -a/2+\delta & a/2+\delta & \delta \\ -a/2-\delta & a/2-\delta & -\delta \end{pmatrix}$$

Note that

$$L_\delta(-1,1,0) \;=\; (-2a, a, a) \;=\; L(-1,1,0)$$

and

$$L_\delta(-1,0,1) \;=\; (-a, a/2, a/2) \;=\; L(-1,0,1).$$

Hence $L_\delta\,|_Y = L\,|_Y = A$. It is easy to verify that $L_\delta \in \mathcal{L}(X, Y)$ and $\|L_\delta\| = \|A\|$ for $|\,\delta\,|$ sufficiently small. Hence the set $\mathcal{A}_0$ consists of more than one element.

Theorem 3.2 leads to an effective method of calculating $\operatorname{dist}(L, \mathcal{L}_Y)$ for given $L \in \mathcal{L}(X)$ if $\operatorname{dist}(L, \mathcal{L}_Y) > \|L\,|_Y\|$. To do this, consider for given $x_1, \ldots, x_n \in \operatorname{ext}(S_X)$ the following system of equations

$$\begin{aligned} (e_i \otimes x_i)(L - f(\cdot)(y_1, \ldots, y_n)) \;&=\; z \quad (i = 1, \ldots, n), \\ \sum_{i=1}^{n} f_i y_i \;&=\; 0, \end{aligned} \tag{3.3}$$

with unknown variables $y_1, \dots, y_n, z$. Assume additionally that

$$0 \in \operatorname{conv}((e_1 \otimes x_1)|_{\mathcal{L}_Y}, \dots, (e_n \otimes x_n)|_{\mathcal{L}_Y}). \tag{3.4}$$

Let $L_0 = f(\cdot)y^0 \in \mathcal{P}(\mathcal{L}_Y)(L)$. Then, in view of Theorem 3.2, if f satisfies (2.5) and $f_i \neq 0$ for $i = 1, \dots, n$, there exist $x_1, \dots, x_n \in \operatorname{ext}(S_X)$ such that $y_1^0, \dots, y_n^0, \operatorname{dist}(L, \mathcal{L}_Y)$ are a solution of (3.3) for the above $x_1, \dots, x_n$. So to find $L_0 \in \mathcal{P}_{\mathcal{L}_Y}$ and $\operatorname{dist}(L, \mathcal{L}_Y)$ it is sufficient to solve finite number of the equations (3.3) for $x_1, \dots, x_n$ satisfying (3.4). For verifying (3.4) we apply

Proposition 3.5. *Assume $x_1, \dots, x_n \in \operatorname{ext}(S_X)$. Let $f \in S_{X^*}$ satisfy (2.5) and let $f_i \neq 0$ for $i = 1, \dots, n$. Put $Y = \ker f$. Then*

$$\begin{aligned} 0 \in \operatorname{conv}((e_1 \otimes x_1)|_{\mathcal{L}_Y}, \dots, (e_n \otimes x_n)|_{\mathcal{L}_Y}) \quad &\textit{iff} \\ \operatorname{sgn}(f(x_j)f_1) = \operatorname{sgn}(f(x_1)f_j) \quad &\textit{for} \ \ j = 1, \dots, n. \end{aligned} \tag{3.5}$$

Proof: Fix $x_1, \dots, x_n \in \operatorname{ext}(S_X)$ and suppose

$$0 = \sum_{i=1}^{k} \lambda_i (e_i \otimes x_i)|_{\mathcal{L}_Y}.$$

Since $f_i \neq 0$ for $i = 1, \dots, n$ and f satisfies (2.5), $k = n$. Now take for $j = 2, \dots, n$ a map $L_j \in \mathcal{L}_Y$ defined by (3.2). Note that for $j = 2, \dots, n$

$$0 = \sum_{i=1}^{n} \lambda_i (e_i \otimes x_i)(L_j) = \lambda_1(-f_j/f_1)f(x_1) + \lambda_j f(x_j).$$

Consequently

$$\lambda_1/\lambda_j = f(x_j)f_1/f(x_1)f_j,$$

which completes the proof. □

Proposition 3.5 shows that for calculating $\operatorname{dist}(L, \mathcal{L}_Y)$ and $L_0 \in \mathcal{P}_{\mathcal{L}_Y}(L)$ it is sufficient to solve system (3.3) only for $x_1, \dots, x_n \in \operatorname{ext}(S_{X^*})$ satisfying (3.5). This fact leads to an algorithm for computing $\operatorname{dist}(L, \mathcal{L}_Y)$ which will be presented in the next section.

4

Referring to the previous theoretical results, we present some computer experiments. In particular, we implemented an algorithm for computing $\operatorname{dist}(L, \mathcal{L}_Y)$ or λ_A by solving a suitable linear system by two methods. First we present a method based on Proposition 3.5. Next we calculate $\operatorname{dist}(L, \mathcal{L}_Y)$ by a mathematical programming problem. Finally, some statistic concerning the situation $\lambda_A = \|A\|$ will be presented. The experiments were done for the case $n = 3$ on a personal computer Apple Macintosh.

First form of the extremum problem.

Referring to Theorem 3.2 and Proposition 3.5 we implemented the following program to calculate a vector y and a scalar z solving a set of linear systems.

Routines

Init

input from file: $L = (L_1, \ldots, L_n)$, f

(L_i denotes the i-th row of the corresponding to L matrix, f satisfies the assumptions of Theorem 3.1).

$x^i \leftarrow x^{oi}$ for $i = 1, \ldots, n$

($x^{o1}, \ldots, x^{on}$ have to satisfy (3.5)).

Solve

solve the system in y and z:

$\sum_{i=1}^n f_i \cdot y_i = \langle f, y\rangle = 0$

$\langle f, x^i\rangle \cdot y_i + z = \langle L_i, x^i\rangle$ for $i = 1, \ldots, n$

Norma

compute the norm:

norm $\leftarrow \max_{i\in\{1,\ldots,n\}} \sum_{j=1}^n |\, L_{ij} - f_j \cdot y_i \,|$

Newsys

if $z \neq$ norm define a new system:

$$\begin{cases} x^{1i} \leftarrow x^i & i = 1, \ldots, n \\ x^i_j \leftarrow \operatorname{sgn}(L_{ij} - f_j \cdot y_i)\ (1 \text{ if } \operatorname{sgn}(L_{ij} - f_j \cdot y_i) = 0) & i, j = 1, \ldots, n \\ E = \{i : \sum_{j=1}^n |\, L_{ij} - f_j \cdot y_i \,| = \text{norm}\} & \end{cases}$$

Case 1

$$\begin{cases} \operatorname{sgn}(\langle f, x^i\rangle) = \operatorname{sgn}(\langle f, x^{1i}\rangle) & \text{for } \ i = 1, \ldots, n \ \text{ or} \\ \operatorname{sgn}(\langle f, x^i\rangle) = -\operatorname{sgn}(\langle f, x^{1i}\rangle) & \text{for } \ i = 1, \ldots, n, \ \text{then} \\ \langle f, y\rangle = 0 & \\ \langle f, x^i\rangle \cdot y_i + z \ = \ \langle L_i, x^i\rangle & i = 1, \ldots, n \end{cases}$$

Case 2

$$\begin{cases} \operatorname{sgn}(\langle f, x^i\rangle) \ \neq \ \operatorname{sgn}(\langle f, x^{1i}\rangle) & \text{for some } \ i \in \{1, \ldots, n\} \\ \operatorname{sgn}(\langle f, x^i\rangle) \ = \ \operatorname{sgn}(\langle f, x^{1i}\rangle) & \text{for some } \ i \in E, \ \text{then} \\ \langle f, y\rangle = 0 & \\ \langle f, x^i\rangle \cdot y_i + z \ = \ \langle L_i, x^i\rangle & \text{if } \ \operatorname{sgn}(\langle f, x^i\rangle) \ = \ \operatorname{sgn}(\langle f, x^{1i}\rangle) \\ \langle f, x^{1i}\rangle \cdot y_i + z \ = \ \langle L_i, x^{1i}\rangle & \text{if } \ \operatorname{sgn}(\langle f, x^i\rangle) \ \neq \ \operatorname{sgn}(\langle f, x^{1i}\rangle) \end{cases}$$

Case 3

$$\begin{cases} \operatorname{sgn}(\langle f, x^i\rangle) \ \neq \ \operatorname{sgn}(\langle f, x^{1i}\rangle) & \text{for every } \ i \in E \\ \operatorname{sgn}(\langle f, x^i\rangle) \ = \ \operatorname{sgn}(\langle f, x^i_1\rangle) & \text{for some } \ i \notin E \end{cases}$$

Case 3A

$$\begin{cases} \operatorname{sgn}(\langle f, x^i\rangle) \;=\; \operatorname{sgn}(\langle f, x^{1i}\rangle), x^i \neq\; x^{1i} & \text{for some } \; i \notin E \\ \langle f, y\rangle \;=\; 0 & \\ \langle f, x^{1i}\rangle \cdot y_i + z \;=\; \langle L_i, x^{1i}\rangle & \text{if } \; i \in E \; \text{ or} \\ & \operatorname{sgn}(\langle f, x^i\rangle) \;\neq\; \operatorname{sgn}(\langle f, x^{1i}\rangle) \\ \langle f, x^i\rangle \cdot y_i + z \;=\; \langle L_i, x^i\rangle & \text{otherwise} \end{cases}$$

Case 3B

$$x^i = x^{1i} \;\text{ for every }\; x \notin E \;\text{ with}$$
$$\operatorname{sgn}(\langle f, x^i\rangle) \;=\; \operatorname{sgn}(\langle f, x^{1i}\rangle)$$

Stop

Main <u>Init</u>
repeat <u>Solve</u> ;
<u>Norma</u> ;
if $z \neq$ norm , then <u>Newsys</u> ;
until $z =$ norm or Case 3B.

Remark 4.1. The algorithm fails if $\operatorname{dist}(L, \mathcal{L}_Y) = \|L\,|_Y\|$ or Case 3B holds true. If $\operatorname{dist}(L, \mathcal{L}_Y) > \|L\,|_Y\|$ and Case 3B holds true, it is necessary to find $(x_1, \dots, x_n)$ satisfying (3.5) different from the previous ones and continue the procedure described above. (Here the classical Remez algorithm can be applied.) By Theorems 3.1, 3.2 and Proposition 3.5 we find a solution after a finite number of steps. Note that after every step the value z strictly increases. The proof of this fact is a simple consequence of (3.4) and the choice of new data-system. It is clear that after every step the value z estimates from below $\operatorname{dist}(L, \mathcal{L}\,|_Y)$.

Now we describe one particular situation in which Case 3B does not hold.

Remark 4.2. Let f be such as in Theorem 3.1. Let $z, y^0 = (y_1^0, \dots, y_n^0)$ be a solution of (3.3) with a data-system $L, f, x_1, \dots, x_n$ satisfying (3.4). If $z \geq \|L\,|_Y\|$, then for every $i \in E$ $\operatorname{sgn}(\langle f, x^i\rangle) = \operatorname{sgn}(\langle f, x^{1i}\rangle)$.

Proof: Suppose that there is $i \in E$ with $\operatorname{sgn}(\langle f, x^i\rangle) = -\operatorname{sgn}(\langle f, x^{1i}\rangle)$. Hence, there is $\alpha \in (0,1)$ such that $y = \alpha x^{1i} + (1-\alpha)x^i \in Y$. Put $L_0 = f(\cdot)y^0$. Then

$$\begin{aligned} \|L\,|_Y\| \;&\geq\; e_i(L - L_0)y = \alpha e_i(L - L_0)x^{1i} \\ &+(1-\alpha)e_i(L-L_0)x^i = \alpha z + (1-\alpha)\|L - L_0\| \\ &> \alpha\|L\,|_Y\| + (1-\alpha)\|L\,|_Y\| \;=\; \|L\,|_Y\|; \end{aligned}$$

a contradiction. □

Now we describe one class of such operators.

Example 4.3. Suppose that we have a matrix $L = [L_{ij}]_{i,j=1,\dots,n}$ and $x^1, \dots, x^n$ satisfying (3.4) such that $\|e_i \circ L\| = e_i(Lx^i)$ for $i = 1, \dots, n$. Put $\delta_1 = 1$, $\delta_i = f(x^1)f_i/f(x^i)f_1$ for $i = 2, \dots, n$. Let $\lambda_i = \delta_i / \sum_{j=1}^n \delta_j$ for $i = 1, \dots, n$. If

$$\sum_{i=1}^n \lambda_i \|e_i \circ L\| \;\geq\; \|L\,|_Y\|,$$

then the solution z, y^0 for a data-system $L, f, x^1, \dots, x^n$ satisfies $z \geq \|L\,|_Y\|$.

Proof: Let $L_0 = f(\cdot) \cdot y_0$. Note that, by (3.5),

$$\begin{aligned} z \;&=\; \sum_{i=1}^n \lambda_i e_i (L - L_0) x^i \;=\; \sum_{i=1}^n \lambda_i e_i (Lx^i) \\ &= \sum_{i=1}^n \lambda_i \|e_i \circ L\| \;\geq\; \|L\,|_Y\|, \end{aligned}$$

as required. □

Second form of the extremum problem.

We reformulate the problem of calculating a vector y and a scalar z as a mathematical programming problem:

$$\min z \text{ such that}$$

$$\sum_{j=1}^n |\, L_{ij} - f_j \cdot y_i \,| \;=\; z \text{ for } i = 1, \dots, n, \tag{4.1}$$

$$\sum_{j=1}^n f_j \cdot y_j = 0.$$

This problem is nonlinear by the absolute values in constraints in (4.1); to eliminate them the problem may be rewritten as a problem of calculating two matrices A^+, A^-, two vectors y^+, y^- and a scalar z by the following mathematical programming problem:

$$\min z \text{ such that}$$

$$\sum_{j=1}^n A_{ij}^+ + A_{ij}^- \;=\; z \text{ for i=1, \dots ,n} \tag{4.2}$$

$$A_{ij}^+ - A_{ij}^- \;=\; L_{ij} - f_j \cdot y_i \text{ for } i, j = 1, \dots, n \tag{4.3}$$

$$\sum_{j=1}^n f_j \cdot y_j \;=\; 0 \tag{4.4}$$

$$y_i^+ - y_i^- \;=\; y_i \text{ for } i = 1, \dots, n. \tag{4.5}$$

The constraints (4.3), (4.4) are equivalent to the constraints in (4.1) under suitable conditions:

$$A_{ij}^{+} + A_{ij}^{-} \;=\; |\, L_{ij} - f_j \cdot y_i \,|$$

if

$$A_{ij}^{+} - A_{ij}^{-} \;=\; 0. \tag{4.6}$$

This condition is also a nonlinear one; on the other side the complementary condition (4.6) may be considered in the solution of (4.2)–(4.5) by the simplex method (Dantzing, Hadley) adding the condition that at least one of the couple of variables (A_{ij}^{+}, A_{ij}^{-}) is not in the basis and consequently is equal to 0. If we denote by z^* the minimum of the problem (4.1), by z^0 the minimum of the problem (4.2)–(4.5) and by $\hat{z}$ the minimum of the problem (4.2)–(4.5) with the complementary condition (4.6), then the following relation holds:

$$\hat{z} \;\geq\; z^* \;\geq\; z^0. \tag{4.7}$$

The first step of this algorithm is to solve by the simplex method the problem (4.2)–(4.5) with the complementary condition (4.6) for obtaining $\hat{z}$; the second step is to solve by a simplex method the problem (4.2)–(4.5) for obtaining z^0. If $\hat{z} = z^0$, then $\hat{z} = z^* = z^0$ and we have solved the problem (4.1). If $\hat{z} \neq z^0$ we compute for $i = 1, \ldots, n$ $\sum_{j=1}^{n} A_{ij}^{+} + A_{ij}^{-}$; if they all are equal to z^0, then $z^* = z^0$, otherwise, because of (4.7) we have an approximation of z^*.

Routines
Init random input (between -3, 3) : L, f

Simplex simplex algorithm modified in order to check
the complementary condition;
print of current solution;

Main Init ;
constraints, matrix :

Simplex (with the complementary condition);
if $Sa \neq \emptyset$ then

Simplex (with the complementary condition);
print of the optimal solution and test;

Simplex (without the complementary condition);
print of the optimal solution and test.

At the end of this section we present a statistic experiment concerning a problem how often $\lambda_A = \|A\|$. We choose $n = 3$ and $f = (1, 1, 1)$. Then

$$Y = \ker(f) = \{(x, y, z) : x + y + z = 0\}.$$

Let $L \in \mathcal{L}(R^3, Y)$ be represented by a matrix $[L_{ij}]_{i,j=1,\ldots,n}$ and let $A = L\,|_Y$. Note that $F = \{(-1,1,0),(-1,0,1)\}$ is a basis of Y. It is easily seen that A has a following matrix representation with respect to F:

$$A = \begin{pmatrix} L_{22} - L_{21} & L_{23} - L_{21} \\ L_{32} - L_{31} & L_{33} - L_{31} \end{pmatrix}. \tag{4.8}$$

Note that

$$\|L\| \;=\; \sup\{\|Lx\| : \|x\| = 1\} \;=\; \max_{i=1,2,3} \sum_{j=1}^{3} |\,L_{ij}\,| \tag{4.9}$$

and

$$\|A\| \;=\; \sup\{\|Ay\| : y \in S_Y\} \;=\; \max\{\|A(1,0)\|, \|A(0,1)\|, \|A(-1,1)\|\}.$$

Routines

nxy compute the norm of the projected vector (x, y);

$$nxy(x,y) \;=\; \begin{cases} |\,x\,| + |\,y\,| & \text{if } x \cdot y \geq 0 \\ \max(|\,x\,|, |\,y\,|) & \text{if } x \cdot y < 0 \end{cases}$$

inrandom pseudorandom input with the following rules:
$f \leftarrow [1,1,1]$
L_{ij} random numbers (between -10 e 10), except that
for $j = n$ in order to have $L\,|_Y : Y \to Y$
compute the norm of L according to (4.9);
compute the projected matrix A according to (4.8);

Main

inrandom
$\|A\| = \max\{nxy(A_{11}, A_{22}), nxy(A_{12}, A_{22}), nxy(A_{12} - A_{11}, A_{22} - A_{21})\}$
$\|P_1\| = \max\{nxy(A_{12} + A_{11}, A_{22} + A_{21}), \|A\|\}$
$\|P_2\| = \max\{nxy(A_{12} - 2A_{11}, A_{22} - 2A_{21}), nxy(A_{12}, A_{22})\}$
$\|P_3\| = \max\{nxy(A_{11} - 2A_{12}, A_{21} - 2A_{22}), nxy(A_{11}, A_{21})\}$;
print the percentage of $\|A\| = \|P_1\|$ or $\|A\| = \|P_2\|$ or $\|A\| = \|P_3\|$.

The situation $\|A\| = \|P_1\|$ or $\|A\| = \|P_2\|$ or $\|A\| = \|P_3\|$ had a frequency of about 77%; in particular we had the following results:

N^0 of tests	At least one $\|P_i\| = \|A\|$	%
100	79	79.0
500	385	77.0
1000	784	78.4
2000	1572	78.6
3000	2328	77.6
5000	3845	76.9

This means that the assumption $\operatorname{dist}(L, \mathcal{L}\,|_Y) > \|L\,|_Y\|$ necessary in the first algorithm is satisfied in about 23% of problems.

References

[1] Baronti M., Papini P.L., *Norm one projections onto subspaces of l^p*, Ann. Mat. Pura Appl. **IV** (1988), 53–61.

[2] Blatter J., Cheney E.W., *Minimal projections onto hyperplanes in sequence spaces*, Ann. Mat. Pura Appl. **101** (1974), 215–227.

[3] Collins H.S., Ruess W., *Weak compactness in spaces of compact operators and vector valued functions*, Pacific J. Math. **106** (1983), 45–71.

[4] Odyniec Wl., Lewicki G., *Minimal Projections in Banach Spaces*, Lecture Notes in Math. **1449**, Springer-Verlag.

[5] Singer I., *On the extension of continuous linear functionals ...*, Math. Ann. **159** (1965), 344–355.

[6] ———, *Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces*, Springer-Verlag, Berlin, Heidelberg, New York, 1970.

[7] Sudolski J., Wojcik A., *Some remarks on strong uniqueness of best approximation*, Approximation Theory and its Applications **6** (1990), 44–78.

Istituto Matematico, Facoltà di Ingegneria, Piazzale Kennedy, pad. D, 16129 Genova, Italy

E-mail: baronti@igecuniv.cisi.unige.it

Università degli Studi di Genova, Dipartimento di Matematica, via Leon Battista Alberti, Genova, Italy

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

E-mail: lewicki@im.uj.edu.pl

(*Received* December 10, 1993, *revised* April 10, 1995)