EAST JOURNAL ON APPROXIMATIONS
Volume 16, Number 3 (2010), 193-207

OPTIMAL SUBSPACES IN NORMED SPACES*

H. H. Cuenya, F. E. Levis, M. D. Lorenzo,
and C. N. Rodriguez

In this paper, we prove existence of optimal subspaces in a normed space. We give properties of diameters of a subspace, and properties of optimal subspaces and of their deviations. Characterization and uniqueness of optimal subspaces in an Hilbert space are considered.

Mathematics Subject Classification (2000): Primary 41A65. Secondary 41A28.

Key words and phrases: n-dimensional diameter, optimal subspace, reflexive space.

1. Introduction

Let $(F, \| \ \|)$ be a normed space and let $m \in \mathbb{N}$. We consider a monotone norm ρ defined in $\mathbb{R}^m$, i.e., ρ is a norm such that $\rho(x_1, \dots, x_m) \leq \rho(y_1, \dots, y_m)$ if $|x_i| \leq |y_i|$, $1 \leq i \leq m$. In addition, we will say that ρ is strictly monotone if the strict inequality holds when we have strict inequality in some coordinate.

Let $Y = \{f_1, \dots, f_m\} \subset F$. For U a proximinal subset of F, we write $P_U(f_k)$, the metric projection of f_k on the set U, $1 \leq k \leq m$. If $d(f, U)$ is the distance from a point f to set U, we denote

$$E(Y, U) = \rho(d(f_1, U), \dots, d(f_m, U)), \tag{1}$$

the *deviation* of the set Y from the set U.

For $n \in \mathbb{N}$ we consider the set

$$\Pi_n(F) = \{V \text{ subspace of } F : \dim V \leq n\}.$$

*Totally supported by Universidad Nacional de Río Cuarto, CONICET and ANPCyT.

The value

$$E(Y) := \inf_{V \in \Pi_n(F)} E(Y, V), \tag{2}$$

is called the *n-dimensional diameter* of the set Y. We say that a linear subspace $V_0 \in \Pi_n(F)$ is *an n-optimal subspace* for Y if $E(Y) = E(Y, V_0)$. We shall omit Y in (1) - (2) in those sections where it remains fixed.

Given a finite set Y, throughout this paper we denote by X the linear space generated by the elements of Y and we write $X = span\, Y$. We observe that if X has dimension at most n, then X is an n-optimal subspace for Y. We will always assume that $n < dim\, X$.

The concepts n-dimensional diameter and n-optimal subspace were introduced by A. N. Kolmogorov in [4]. Other works about this concepts can be seen in [3] and [8]. Recently in [1] and [2] it was proved the existence of n-optimal subspaces in a Hilbert space. They give a constructive proof of existence and applications to problem of finding a model space that describes a given class of signals or images.

The present paper is organized as follows. In Section 2 we prove more general results on existence of n-optimal subspaces (Theorems 1, 2, and Remark 1). In Sections 3 and 4 we study properties of deviations, n-dimensional diameters and n-optimal subspaces. Finally, in Section 5 we give a characterization of n-optimal subspaces and prove a uniqueness result in Hilbert spaces.

2. Existence of optimal subspaces

The following Lemma was proved in [6, p. 273].

Lemma 1. *Let F be a Banach space of dimension n. Then there exist n linearly independent elements $e_1, \dots, e_n \in F$ and n functionals $g_1, \dots, g_n \in F^*$ such that $\|e_k\| = \|g_k\| = 1$, $g_i(e_k) = 1$ if $i = k$, and $g_i(e_k) = 0$ if $i \neq k$, $1 \leq i, k \leq n$.*
Consequently, for every $e = \sum_{i=1}^{n} \alpha_i e_i \in F$ we have then $|\alpha_i| \leq \|e\|$, $1 \leq i \leq n$.

In the next theorem if $\mathbb{N}_0 = \mathbb{N}$ we will denote by $\mathbb{N}_i$ a subsequence of $\mathbb{N}_{i-1}$ for $i = 1, 2$.

Theorem 1. *Suppose F is a reflexive space and let $Y = \{f_1, \dots, f_m\} \subset F$. Then there exists $V_0 \in \Pi_n(F)$ such that V_0 is an n-optimal subspace for Y.*

Proof. Let $\{V_s\}_{s \in \mathbb{N}} \subset \Pi_n(F)$ be such that

$$E = \lim_{s \to \infty} E(V_s).$$

Let $g_{sk} \in P_{V_s}(f_k)$. It is easy to see that $\|g_{sk}\| \leq 2\|f_k\|$. So, there exists a positive constant M satisfying

$$\|f_k - g_{sk}\| \leq M, \quad 1 \leq k \leq m, \quad s \in \mathbb{N}.$$

Therefore, there are a subsequence of $\{V_s\}_{s\in\mathbb{N}}$, say $\{V_s\}_{s\in\mathbb{N}_1,}$, and $r_k \in \mathbb{R}$, $1 \leq k \leq m$, such that

$$\lim_{s\in\mathbb{N}_1, s\to\infty} \|f_k - g_{sk}\| = r_k \tag{3}$$

and the dimension of V_s is constant, say l, for all $s \in \mathbb{N}_1$. The last fact is a consequence of that the dimension of V_s is at most n for all $s \in \mathbb{N}$. By Lemma 1, for each $s \in \mathbb{N}_1$, there exists a basis $\{e_{js}\}_{j=1}^{l}$ of V_s such that $\|e_{js}\| = 1$, and if $g_{sk} = \sum_{j=1}^{l} c_{js}^k e_{js}$, then $|c_{js}^k| \leq \|g_{sk}\|$. So, using the triangle inequality, we get

$$|c_{js}^k| \leq M + \sup_{1\leq i\leq m} \|f_i\|, \quad 1 \leq k \leq m, \quad 1 \leq j \leq l, \quad s \in \mathbb{N}_1. \tag{4}$$

Since F is reflexive, there are a subsequence of $\{V_s\}_{s\in\mathbb{N}_1}$, say $\{V_s\}_{s\in\mathbb{N}_2}$, and $e_j \in F$, $1 \leq j \leq l$, such that e_{js} weakly converges to e_j, $s \in \mathbb{N}_2$, $s \to \infty$. From (4), we can assume

$$\lim_{s\in\mathbb{N}_2, s\to\infty} c_{js}^k = c_j^k, \quad 1 \leq j \leq l, \quad 1 \leq k \leq m.$$

Thus g_{sk} weakly converges to $\sum_{j=1}^{l} c_j^k e_j =: b_k \in span\{e_1, ..., e_l\} =: V_0 \in \Pi_n(F)$. Now, using the weak lower semicontinuity of the norm $\|.\|$ and (3), we get

$$\|f_k - b_k\| \leq \liminf_{s\in\mathbb{N}_2} \|f_k - g_{sk}\| = r_k, \quad 1 \leq k \leq m.$$

So, the monotonicity of ρ implies

$$\begin{aligned} E(V_0) &\leq \rho\left(\|f_1 - b_1\|, \ldots, \|f_m - b_m\|\right) \leq \rho(r_1, \ldots, r_m) \\ &= \lim_{s\in\mathbb{N}_2, s\to\infty} E(V_s) = E. \end{aligned}$$

□

Remark 1. a) If F has finite dimension then F is a reflexive space, so by Theorem 1 there exists an n-optimal subspace for Y.

b) When F is the a space conjugate to some Banach space, then there exists an n-optimal subspace for Y. In fact, the proof follows by replacing in Theorem 1 the weak convergence by w^*-convergence.

We recall that a linear subspace U of F is a Chebyshev space if $P_U(f)$ is a one-point set for all $f \in F$ (see [6]).

Lemma 2. *Suppose X is a Chebyshev space and let $Y = \{f_1, \dots, f_m\} \subset F$. If P_X is a linear operator, then there exists a linear subspace $V_0 \in \Pi_n(F)$, such that $V_0 \subset X$ and*

$$E(V_0) \leq \|P_X\| E(V) \quad \text{for all } V \in \Pi_n(F).$$

Proof. Since X has finite dimension, replacing in Theorem 1 F by X, there is a linear subspace $V_0 \in \Pi_n(F)$, $V_0 \subset X$, such that

$$E(V_0) \leq E(V') \quad \text{for all } V' \in \Pi_n(F), \quad V' \subset X.$$

Let $V \in \Pi_n(F)$ and $V' = P_X(V) \subset X$. Since $Y \subset X$ and P_X is a linear operator, we have $P_X(f_k) = f_k$, $1 \leq k \leq m$, and $V' \in \Pi_n(F)$. We choose $g'_k \in P_{V'}(f_k)$ and $g_k \in P_V(f_k)$, so $P_X(g_k) \in V'$. Then

$$\|f_k - g'_k\| \leq \|f_k - P_X(g_k)\| = \|P_X(f_k - g_k)\| \leq \|P_X\| \|f_k - g_k\|, \tag{5}$$

and consequently $E(V_0) \leq E(V') \leq \|P_X\| E(V)$. □

The next theorem immediately follows from Lemma 2.

Theorem 2. *Under the same assumptions as in Lemma 2, if $\|P_X\| = 1$, then there exists $V_0 \in \Pi_n(F)$ such that V_0 is an n-optimal subspace for Y.*

Remark 2. If X is not a Chebyshev space, but there is a lineal metric selection of P_X (see [5, p. 25]) of norm 1, then the same proof of Lemma 2 shows the existence of an n-optimal subspace for Y.

Next, we give an example such that Remark 2 can be applied, but Remark 1, b) cannot.

Example 1. By Theorem 15.5 in [7, p. 454] the space $l_1(\mathbb{N})$ has a subspace F which is not isomorphic to any conjugate Banach space. Moreover, F has the following sequence $\{f_n\}_{n \in \mathbb{N}}$ as a monotone basis:

$$f_n = x_n - \frac{1}{2} x_{2n+1} - \frac{1}{2} x_{2n+2}, \quad n \in \mathbb{N},$$

where $x_n(m) = \delta_n(m)$ and δ_n is the Kronecker delta. Let $Y = \{f_1, f_2\}$. For $g = \sum_{n=1}^{\infty} \alpha_n f_n \in F$, and $X = \mathit{span}\, Y$, a straightforward computation shows that $P : F \to X$ defined by $P(g) = \alpha_1 f_1 + \alpha_2 f_2$ is a lineal metric selection of P_X with $\|P\| = 1$.

3. Properties of optimal subspaces

Lemma 3. *Suppose ρ is a strictly monotone norm. If $V_0 \in \Pi_n(F)$ is an n-optimal subspace for Y, then $dim V_0 = n$.*

Proof. Suppose $\dim V_0 = r$, $r < n$. Since $V_0 \neq X = span\, Y$, there exists $1 \leq j \leq m$ such that $f_j \notin V_0$. Let $W = V_0 \oplus span\{f_j\}$. Clearly $W \in \Pi_n(F)$. We choose $g_k \in P_W(f_k)$ and $h_k \in P_{V_0}(f_k)$, $1 \leq k \leq m$. Since $\|f_k - g_k\| \leq \|f_k - h_k\|$, $\|f_j - h_j\| > 0$, and $\|f_j - g_j\| = 0$, then $E(W) < E(V_0)$, a contradiction. □

Definition 1. *Let $Z \subset F$ be a Chebyshev subspace. Then we say that Z has property (P) if $\|P_Z\| = 1$ and $\|P_Z(f)\| = \|f\|$ implies $f \in Z$.*

Lemma 4. *Let F be a strictly convex space and let $Z \subset F$ be a Chebyshev subspace. If $\|P_Z\| = 1$, then Z has the property (P).*

Proof. Let $f \in F$ be such that $\|P_Z(f)\| = \|f\|$. Suppose $f \notin Z$, then $g = \frac{f}{\|f\|} \notin Z$. Let $u = \frac{g+P_Z(g)}{2}$. Since $\|u - P_Z(g)\| + \|u - g\| = \|g - P_Z(g)\|$, then $\|u - P_Z(g)\| \leq \|u - h\|$ for all $h \in Z$, and so

$$P_Z(g) = P_Z(u). \tag{6}$$

On the other hand, the operator P_Z is positive homogeneous, so we have $\|P_Z(g)\| = \|g\| = 1$. The strict convexity of F implies $\|u\| < 1$. Since $\|P_Z\| = 1$, (6) implies that $1 = \|P_Z(u)\| \leq \|P_Z\|\|u\| = \|u\|$, a contradiction. □

Remark 3. a) Every linear subspace of a Hilbert space has property (P).

b) The strict convexity is not a necessary condition for property (P) to occur (see Example 1).

In [1] the authors proved that if F is a Hilbert space, then the existence of an n-optimal subspace for Y implies the existence of an n-optimal subspace for Y contained in X. The following theorem shows that necessarily, all n-optimal subspaces for Y must be contained in X, even for more general normed spaces.

Theorem 3. *Let F be a strictly convex space and let ρ be a strictly monotone norm. Suppose X has property (P) and P_X is a linear operator. If $V_0 \in \Pi_n(F)$ is an n-optimal subspace for Y, then $V_0 \subset X$.*

Proof. Let $V_0 \in \Pi_n(F)$ be an n-optimal subspace for Y and $V = P_X(V_0) \subset X$. Since P_X is a linear operator, $V \in \Pi_n(F)$. So, (5) implies that V is an n-optimal subspace for Y. Moreover,

$$\|f_k - P_V(f_k)\| = \|f_k - P_X(P_{V_0}(f_k))\|, \quad 1 \le k \le m, \tag{7}$$

and

$$\|P_X(f_k - P_{V_0}(f_k))\| = \|f_k - P_{V_0}(f_k)\|, \quad 1 \le k \le m. \tag{8}$$

From (7) and the uniqueness of the best approximant,

$$P_V(f_k) = P_X(P_{V_0}(f_k)), \quad 1 \le k \le m. \tag{9}$$

As X has the property (P), by (8) we get $P_X(f_k - P_{V_0}(f_k)) = f_k - P_{V_0}(f_k)$, $1 \le k \le m$, and so

$$P_X(P_{V_0}(f_k)) = P_{V_0}(f_k), \quad 1 \le k \le m. \tag{10}$$

Let $\overline{X} = span\{P_{V_0}(f_1), \dots, P_{V_0}(f_m)\} \subset V_0$. From (9) and (10), we have

$$\overline{X} \subset V_0 \cap V. \tag{11}$$

In addition, $\|f_k - P_{\overline{X}}(f_k)\| \le \|f_k - P_{V_0}(f_k)\|$, $1 \le k \le m$. Then $E(\overline{X}) \le E(V_0)$, i.e., $\overline{X} \in \Pi_n(F)$ is an n-optimal subspace for Y. By Lemma 3 we know that $\overline{X}$, V and V_0 have dimension n, so (11) implies $V_0 = V \subset X$. □

4. Deviations and diameters

The proof of the next proposition follows the same patterns as the proof of [6, Theorem 6.10, p. 157].

Proposition 1. *Let $Y_1 = \{f_1, \dots, f_m\} \subset F$, $Y_2 = \{h_1, \dots, h_m\} \subset F$ and let $U \subset F$. The following statements holds true.*

a) $|E(Y_1, U) - E(Y_2, U)| \le \rho(\|f_1 - h_1\|, \dots, \|f_m - h_m\|)$;

b) If U is a linear subspace, then $E(\alpha Y_1, U) = |\alpha| E(Y_1, U)$ for all $\alpha \in \mathbb{R}$;

c) $E(Y_1 + Y_2, U) \le E(Y_1, U) + E(Y_2, U)$;

d) If $U_1 \subset U$, then $E(Y_1, U) \le E(Y_1, U_1)$.

We denote the supremum norm in $\mathbb{R}^m$ by $\|x\|_\infty$, i.e., $\|x\|_\infty = \max\limits_{1 \le k \le m} |x_k|$, $x = (x_1, \dots, x_m) \in \mathbb{R}^m$, and set $e = (1, \dots, 1) \in \mathbb{R}^m$.

Proposition 2. *Let $Y = \{f_1, ..., f_m\} \subset F$, and let $U \subset F$. Assume that $\|x\|_\infty \leq \rho(x)$ for all $x \in \mathbb{R}^m$. Then*

$$\inf_{\epsilon>0, Y\subset U+\epsilon S_1} \epsilon \leq E(Y,U) \leq \rho(e) \inf_{\epsilon>0,\ Y\subset U+\epsilon S_1} \epsilon, \tag{12}$$

where S_1 is the closed ball in F with center 0 and radius 1.

In addition, the two inequalities in (12) become equalities if only if ρ is the supremum norm.

Proof. Given $n \in \mathbb{N}$, let $g_k \in U$, $1 \leq k \leq m$, be such that $\|g_k - f_k\| \leq \frac{1}{n} + d(f_k, U)$. Then

$$f_k = g_k + (f_k - g_k) \in U + \left(\frac{1}{n} + d(f_k,U)\right) S_1 \subset U + \left(\frac{1}{n} + E(Y,U)\right) S_1.$$

Thus

$$Y \subset U + \left(\frac{1}{n} + E(Y,U)\right) S_1. \tag{13}$$

On the other hand, let $\epsilon > 0$ be such that $Y \subset U + \epsilon S_1$. For $f_k \in Y$, there exist $y \in S_1$, $g \in U$ such that $f_k = g + \epsilon y$. Then

$$d(f_k,U) \leq \|f_k - g_k\| \leq \|f_k - g\| + \frac{1}{n} \leq \epsilon + \frac{1}{n},$$

hence

$$E(Y,U) \leq \left(\epsilon + \frac{1}{n}\right) \rho(e). \tag{14}$$

Since n is arbitrary, from (13) and (14) we get (12). Finally, if ρ is the supremum norm, clearly all inequalities in (12) are equalities.

Conversely, the equalities in (12) imply $\rho(e) = 1$, and from monotonicity of ρ it follows that the closed ball in $\mathbb{R}^m$ of center 0 and radius 1 in the supremum norm is contained in the closed ball in $\mathbb{R}^m$ of center 0 and radius 1 in the ρ norm. Since the supremum norm is less than or equal to the ρ norm, the two balls coincide. So, $\rho = \|.\|_\infty$. This concludes the proof. □

Remark 4. Notice that Proposition 2 was proved in [6], when ρ is the supremum norm and U is a linear subspace of F.

Our next goal is to examine continuity of the deviation of the set $Y \subset F$ from a set U as function of the set U.

The one to one correspondence between proximinal sets and its associated metric projections enables us to devise a notion of distance between proximinal sets. Given two proximinal sets U_1, U_2, we define a distance by

$$d_*(U_1,U_2) = \sup\left\{\frac{\|g-h\|}{\|f\|} :\ f \neq 0, g \in P_{U_1}(f), h \in P_{U_2}(f)\right\}.$$

The next lemma immediately follows.

Lemma 5. *Let $Y = \{f_1, \dots, f_m\} \subset F$ and let U_1, U_2 be subsets of F. Then*

$$|E(Y, U_1) - E(Y, U_2)| \leq \rho\left(d(f_1, U_1) - d(f_1, U_2), \dots, d(f_m, U_1) - d(f_m, U_2)\right).$$

The following proposition establishes a Lipschitz property of the function $E(Y, \cdot)$. It is a direct consequence of Lemma 5.

Proposition 3. *Let $Y = \{f_1, \dots, f_m\} \subset F$ and let U_1, U_2 be proximinal subsets of F. Then*

$$|E(Y, U_1) - E(Y, U_2)| \leq \rho(e) \max_{1 \leq k \leq m} \|f_k\| d_*(U_1, U_2).$$

Now, we consider the Hausdorff space $(\mathcal{H}, h)$, where $\mathcal{H} = \{K \subset F : K \text{ is a non empty compact set}\}$, and $h : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ is the metric defined by $h(K_1, K_2) = \max\{d(K_1, K_2), d(K_2, K_1)\}$, with $d(K_1, K_2) = \max_{f \in K_1}\{\|f - g\| : g \in P_{K_2}(f)\}$.

Our next lemma gives a relation between the deviation over linear subspaces and the deviation over subsets in $\mathcal{H}$.

Lemma 6. *Let $Y = \{f_1, \dots, f_m\} \subset F$. If $r > 2 \max_{1 \leq k \leq m} \|f_k\|$, then*

$$E(Y, U) = E(Y, U \cap S_r)$$

for all $U \subset F$ with $0 \in U$, where S_r is the closed ball in F with 0 and radius $r > 0$.

Proof. Given $n \in \mathbb{N}$, let $g_k \in U$, $1 \leq k \leq m$, be such that $\|g_k - f_k\| \leq \frac{1}{n} + d(f_k, U)$. If $\|g_k\| > r$ for some k, we have

$$\frac{1}{n} + d(f_k, U) \geq \|f_k - g_k\| \geq \|g_k\| - \|f_k\| > r - \|f_k\|.$$

Hence, $d(f_k, U) \geq r - \|f_k\| > \|f_k\|$, which contradicts to $0 \in U$. Thus, we have $g_k \in U \cap S_r$, and consequently $d(f_k, U \cap S_r) \leq \frac{1}{n} + d(f_k, U)$. Since n is arbitrary, $d(f_k, U \cap S_r) = d(f_k, U)$, $1 \leq k \leq m$. The claim of Lemma 6 immediately follows. □

Proposition 4. *Let $Y = \{f_1, \dots, f_m\} \subset F$ and let $U_1, U_2 \in \Pi_n(F)$. Then*

$$|E(Y, U_1) - E(Y, U_2)| \leq \rho(e) h(U_1 \cap S_r, U_2 \cap S_r), \tag{15}$$

where S_r is defined as in Lemma 6.

Proof. For the sake of simplicity, $p_K(f)$ will denote an arbitrary element of the set $P_K(f)$. We shall show that if $f \in F$ and $\mathfrak{g} : \mathcal{H} \to \mathbb{R}$ is the function defined by $\mathfrak{g}(K) = d(\{f\}, K)$, then

$$|\mathfrak{g}(K_1) - \mathfrak{g}(K_2)| \leq h(K_1, K_2). \tag{16}$$

Indeed, if $K_1, K_2 \in \mathcal{H}$ and $y \in K_1$, we have

$$\begin{aligned}\|p_{K_1}(f) - p_{K_2}(p_{K_1}(f))\| &\leq \|p_{K_1}(f) - p_{K_2}(y)\| \leq \max_{x \in K_1} \|x - p_{K_2}(y)\| \\ &= d(K_1, K_2) \leq h(K_1, K_2).\end{aligned}$$

Hence,

$$\begin{aligned}\mathfrak{g}(K_2) = \|f - p_{K_2}(f)\| &\leq \|f - p_{K_2}(p_{K_1}(f))\| \leq \|f - p_{K_1}(f)\| \\ &+ \|p_{K_1}(f) - p_{K_2}(p_{K_1}(f))\| \leq \mathfrak{g}(K_1) + h(K_1, K_2).\end{aligned}$$

Analogously, we can get $\mathfrak{g}(K_1) \leq \mathfrak{g}(K_2) + h(K_1, K_2)$, which proves (16). Finally, by Lemma 6, Lemma 5 and (16), we obtain (15). □

The following proposition is an immediate consequence of Lemma 6.

Proposition 5. *Let $Y = \{f_1, \ldots, f_m\} \subset F$. If $r \geq 2 \max_{1 \leq k \leq m} \|f_k\|$, then*

$$E(Y) = \inf_{W \in \Pi_n^r(F)} E(Y, W),$$

where $E(Y)$ is defined by (2) and $\Pi_n^r(F) = \{V \cap S_r : V \in \Pi_n(F)\}$.

Next, we prove that the n-dimensional diameter of a set depends continuously on the set.

Proposition 6. *Let $Y_1 = \{f_1, \ldots, f_m\} \subset F, Y_2 = \{h_1, \ldots, h_m\} \subset F$. Then*

$$|E(Y_1) - E(Y_2)| \leq \rho(\|f_1 - h_1\|, \ldots, \|f_m - h_m\|).$$

Proof. Let $Y_1 = \{f_1, \ldots, f_m\} \subset F, Y_2 = \{h_1, \ldots, h_m\} \subset F$. By the definition of $E(Y_1)$, there exists $U_1 \in \Pi_n(F)$ such that

$$E(Y_1, U_1) < E(Y_1) + \epsilon.$$

Since $E(Y_2) \leq E(Y_2, U_1)$, using Lemma 1 a) we obtain

$$\begin{aligned}E(Y_2) - E(Y_1) &< E(Y_2, U_1) - E(Y_1, U_1) + \epsilon \\ &\leq \rho(\|f_1 - h_1\|, \ldots, \|f_m - h_m\|) + \epsilon\end{aligned}$$

for all $\epsilon > 0$. Then

$$E(Y_2) - E(Y_1) \leq \rho(\|f_1 - h_1\|, \ldots, \|f_m - h_m\|).$$

Analogously we can obtain $E(Y_1) - E(Y_2) \leq \rho(\|f_1 - h_1\|, \ldots, \|f_m - h_m\|)$. This completes the proof. □

5. Characterization and uniqueness of n-optimal subspaces in a Hilbert space

In this section, we characterize the n-optimal subspaces when F is a Hilbert space. We begin with the particular case when $F = \mathbb{R}^k$ and ρ is the Euclidean norm in $\mathbb{R}^m$.

Let $Y = \{f_1, \dots f_m\}$ be a set of vectors in $\mathbb{R}^k$, $X = span\{f_1, \dots, f_m\}$, and $r = \dim X > n$. We set

$$G = \begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix} \in \mathbb{R}^{m \times k}.$$

Since $G^tG = \left(\sum_{i=1}^{m} f_i^t f_i\right)$ is a symmetric matrix, then there exists an orthogonal matrix $Q \in \mathbb{R}^{k\times k}$ such that

$$Q^t \left(\sum_{i=1}^{m} f_i^t f_i\right) Q = diag(\lambda_1, \dots, \lambda_k), \quad \text{and} \quad \lambda_1 \geq \dots \geq \lambda_k \geq 0. \tag{17}$$

We observe that the range of G^tG is $r > n$, therefore $\lambda_n \neq 0$. We denote $p = \max\{j : 1 \leq j \leq n,\ \lambda_j > \lambda_n\}$ if $\lambda_1 > \lambda_n$, and $p = 0$, otherwise, and $s = \max\{j : 1 \leq j \leq k, \lambda_j = \lambda_n\}$.

Let $V \subset \mathbb{R}^k$, $\dim V = n$, and let $\{v_1, \dots, v_n\}$ be an orthonormal basis for V. Set $A = [v_1^t \ \dots \ v_n^t] \in \mathbb{R}^{k\times n}$ and $B = Q^tA \in \mathbb{R}^{k\times n}$.

We proceed with three lemmas.

Lemma 7. *Let $F = \mathbb{R}^k$, and let ρ be the Euclidean norm on $\mathbb{R}^m$. Then $E(Y, V) = \sum_{j=1}^{k} \lambda_j - \sum_{j=1}^{k} \lambda_j (BB^t)_{jj}$.*

Proof. It is easy to see that $E(Y, V) = \sum_{j=1}^{m} \left(\|f_j\|^2 - \sum_{i=1}^{n} (< f_j, v_i >)^2\right)$. Let $\Delta = diag(\lambda_1, \dots, \lambda_k)$. Since

$$\sum_{j=1}^{m} \|f_j\|^2 = trace(GG^t) = trace(G^tG) = \sum_{j=1}^{k} \lambda_j$$

and

$$\begin{aligned} \sum_{i=1}^{n}\sum_{j=1}^{m} (< f_j, v_i >)^2 &= \sum_{i=1}^{n} v_i G^t G v_i^t = \sum_{i=1}^{n} v_i Q \Delta Q^t v_i^t = \sum_{i=1}^{n}\sum_{j=1}^{k} \lambda_j (Q^t v_i^t)_j^2 \\ &= \sum_{j=1}^{k} \lambda_j \left(\sum_{i=1}^{n} (Q^t v_i^t)_j^2\right) = \sum_{j=1}^{k} \lambda_j (BB^t)_{jj}, \end{aligned}$$

we have $E(Y,V)=\sum_{j=1}^{k}\lambda_j-\sum_{j=1}^{k}\lambda_j(BB^t)_{jj}$. □

Lemma 8. *The following conditions are satisfied:*

$$trace(BB^t)=n \quad and \quad 0\le (BB^t)_{ii}\le 1, \quad 1\le i\le k. \tag{18}$$

Proof. Clearly, $trace(BB^t)=trace(B^tB)=trace(A^tA)=n$.
As $BB^t=(BB^t)^tBB^t$, we get $0\le \sum_{j=1}^{k}(BB^t)_{ij}^2=(BB^t)_{ii}$, $1\le i\le k$. Therefore

$$0=\sum_{j=1,j\neq i}^{k}(BB^t)_{ij}^2+(BB^t)_{ii}\left((BB^t)_{ii}-1\right), \quad 1\le i\le k. \tag{19}$$

Thus, $(BB^t)_{ii}\le 1$, $1\le i\le k$. □

Lemma 9. *Let $F=\mathbb{R}^k$, and let ρ be the Euclidean norm in $\mathbb{R}^m$. Suppose that V is n-optimal for Y. Then*

a) If $p>0$, we have $(BB^t)_{ii}=1$, $1\le i\le p$;

b) If $s<r$, we have $(BB^t)_{ii}=0$, $s+1\le i\le r$.

Proof. From Lemma 7, we have $E(Y,V)=\sum_{j=1}^{k}\lambda_j-\sum_{j=1}^{k}\lambda_j(BB^t)_{jj}$. Then Theorem 4.5 in [1] implies $0=E(Y,V)-E(Y)=\sum_{i=1}^{n}\lambda_i-\sum_{i=1}^{k}\lambda_i(BB^t)_{ii}$, i.e.,

$$\sum_{i=1}^{p}\lambda_i(1-(BB)_{ii}^t)+\lambda_n\Big((n-p)-\sum_{i=p+1}^{s}(BB^t)_{ii}\Big)=\sum_{i=s+1}^{k}\lambda_i(BB^t)_{ii}. \tag{20}$$

a) Suppose that there exists j, $1\le j\le p$, such that $0\le (BB^t)_{jj}<1$. Then $\lambda_n(1-(BB^t)_{jj})<\lambda_j(1-(BB^t)_{jj})$ and by (20) we have,

$$\begin{aligned}\lambda_n\Big(n-\sum_{i=1}^{s}(BB^t)_{ii}\Big)&=\lambda_n\sum_{i=1}^{p}(1-(BB^t)_{ii})+\lambda_n\Big((n-p)-\sum_{i=p+1}^{s}(BB^t)_{ii}\Big)\\&<\sum_{i=1}^{p}\lambda_i(1-(BB^t)_{ii})+\lambda_n\Big((n-p)-\sum_{i=p+1}^{s}(BB^t)_{ii}\Big)\\&=\sum_{i=s+1}^{k}\lambda_i(BB^t)_{ii}\le\sum_{i=s+1}^{k}\lambda_n(BB^t)_{ii}.\end{aligned}$$

Hence, $n-\sum_{i=1}^{s}(BB^t)_{ii} < \sum_{i=s+1}^{k}(BB^t)_{ii}$, i.e., $n < trace(BB^t)$, which contradicts (18). Therefore, $(BB^t)_{ii} = 1$, $1 \leq i \leq p$.

b) If there exists j, $s+1 \leq j \leq r$, such that $0 < (BB^t)_{jj} \leq 1$, then $\lambda_j(BB^t)_{jj} < \lambda_n(BB^t)_{jj}$, and by (20) we get,

$$\begin{aligned}\lambda_n\Big(n-\sum_{i=1}^{s}(BB^t)_{ii}\Big) &= \lambda_n\sum_{i=1}^{p}(1-(BB^t)_{ii}) + \lambda_n\Big((n-p)-\sum_{i=p+1}^{s}(BB^t)_{ii}\Big)\\ &\leq \sum_{i=1}^{p}\lambda_i(1-(BB^t)_{ii}) + \lambda_n\Big((n-p)-\sum_{i=p+1}^{s}(BB^t)_{ii}\Big)\\ &= \sum_{i=s+1}^{k}\lambda_i(BB^t)_{ii} < \lambda_n\sum_{i=s+1}^{k}(BB^t)_{ii}.\end{aligned}$$

Hence, $n - \sum_{i=1}^{s}(BB^t)_{ii} < \sum_{i=s+1}^{k}(BB^t)_{ii}$, i.e., $n < trace(BB^t)$, which again contradicts (18). Thus $(BB^t)_{ii} = 0$, $s+1 \leq i \leq r$. □

For a matrix H, we denote by $R(H)$ the range of H.

The following theorem characterizes the n-optimal subspaces in $\mathbb{R}^k$.

Theorem 4. *Let $F = \mathbb{R}^k$, and let ρ be the Euclidean norm on $\mathbb{R}^m$. Then V is n-optimal for Y if only if $V = span\{q_1^t, \ldots, q_p^t\} \oplus W$, where W is any subspace of $span\{q_{p+1}^t, \ldots, q_s^t\}$, $dimW = n-p$, and q_j is the j^{th} column of matrix Q.*

Proof. When the set of indices satisfying certain condition is empty, we shall mean that this condition must be omitted. Suppose V is n-optimal for Y. Next, our goal is to show that

$$n-p = \sum_{i=p+1}^{s}(BB^t)_{ii} \quad \text{for} \quad p \geq 0. \tag{21}$$

There are only four cases to be considered: $s < r$, $p > 0$; $s = r$, $p > 0$; $s < r$, $p = 0$, and $s = r$, $p = 0$. For the first three cases (21) is consequence of (20) and Lemma 9, while the last case directly follows from (20).

Next, we shall prove that

$$(BB^t)_{ii} = 0, \quad s+1 \leq i \leq k. \tag{22}$$

If $p = 0$, then using (21) and Lemma 8 we obtain (22). If $p > 0$, Lemma 9, (a) implies $\sum_{i=1}^{p}(BB^t)_{ii} = p$. From (21) it follows that $\sum_{i=1}^{s}(BB^t)_{ii} = n$. Now,

by Lemma 8 we again get (22).
From Lemma 9, (19) and (22), we get

$$0 = \sum_{j=1, j\neq i}^{k} (BB^t)_{ij}^2, \quad 1 \le i \le p \quad \text{or} \quad s+1 \le i \le k.$$

Since BB^t is a symmetric matrix, BB^t is the block matrix $\begin{bmatrix} I_p & 0 & 0 \\ 0 & \widetilde{B} & 0 \\ 0 & 0 & 0 \end{bmatrix}$,
where I_p is the identity matrix of order p, and $\widetilde{B}$ is certain square matrix of order $s-p$.
We put

$$Q = [[q_1 \ \dots \ q_p][q_{p+1} \ \dots \ q_s][q_{s+1} \ \dots \ q_k]] \quad \text{and} \quad B = \begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix},$$

with $B_1 \in \mathbb{R}^{p\times n}$, $B_2 \in \mathbb{R}^{(s-p)\times n}$ and $B_3 \in \mathbb{R}^{(k-s)\times n}$. Since $BB^tB = B$, we get $B_3 = 0$. As a consequence, $A = QB = [q_1 \ \dots \ q_p]B_1 + [q_{p+1} \ \dots \ q_s]B_2$. Since $R([q_1 \ \dots \ q_p]B_1) \subset R([q_1 \ \dots \ q_p])$, $R([q_{p+1} \ \dots \ q_s]B_2) \subset R([q_{p+1} \ \dots \ q_s])$, and $R([q_1 \ \dots \ q_p]) \cap R([q_{p+1} \ \dots \ q_s]) = \emptyset$, we get

$$R(A) \subset R([q_1 \ \dots \ q_p]) \oplus R([q_{p+1} \ \dots \ q_s]B_2).$$

As $B_2B_2^t = \widetilde{B}$ and $rank(BB^t) = n$, we have $rank(B_2) = rank(\widetilde{B}) = n-p$. Therefore, $rank([q_{p+1} \ \dots \ q_s]B_2) = n-p$, and thus,

$$R(A) = R([q_1 \ \dots \ q_p]) \oplus R([q_{p+1} \ \dots \ q_s]B_2).$$

We conclude that $V = span\{q_1^t, \dots, q_p^t\} \oplus W$, where W is a subspace of $span\{q_{p+1}^t, \dots, q_s^t\}$, $\dim W = n-p$. This completes the proof of the necessity.

Conversely, if $V = span\{q_1^t, \dots, q_p^t\} \oplus W$, where W is any subspace of $span\{q_{p+1}^t, \dots, q_s^t\}$, $\dim W = n-p$, we have $E(Y,V) = \sum_{j=n+1}^{k} \lambda_j$. Then V is n-optimal for Y. □

The following theorem is an immediate consequence of Theorems 11 and 4.

Theorem 5. *Let F be a Hilbert space, $Y = \{f_1, \dots, f_m\} \subset F$ and $X = span\{f_1, \dots, f_m\}$. Let $k = dimX$, and let $V \subset F$, $dimV = n < k$. Let $\tau : (X, \rho) \to (\mathbb{R}^k, Euclidean\ norm)$ be an isometric isomorphism, and let $Q \in \mathbb{R}^{k\times k}$ be an orthogonal matrix such that*

$$Q^t \left(\sum_{i=1}^{m} \tau(f_i)^t \tau(f_i) \right) Q = diag(\lambda_1, \dots, \lambda_k) \quad and \quad \lambda_1 \ge \dots \ge \lambda_k \ge 0. \tag{23}$$

Define $p = \max\{j : 1 \le j \le n,\ \lambda_j > \lambda_n\}$ *if* $\lambda_1 > \lambda_n$ *or* $p = 0$*, otherwise, and* $s = \max\{j : 1 \le j \le k, \lambda_j = \lambda_n\}$*. Then* V *is* n*-optimal for* Y *if only if* $V = span\{\tau^{-1}(q_1^t), \ldots, \tau^{-1}(q_p^t)\} \oplus W$*, where* W *is any subspace of* $span\{\tau^{-1}(q_{p+1}^t), \ldots, \tau^{-1}(q_s^t)\}$*,* $dimW = n - p$*, and* q_j *is the* j^{th} *column of matrix* Q*.*

Corollary 1. *Under the assumptions of Theorem 5, we have*

a) There is a unique n*-optimal subspace* V *for* Y *if and only if* $\lambda_n > \lambda_{n+1}$*. In this case,* $V = span\{q_1^t, \ldots, q_n^t\}$*.*

b) If $\lambda_1 = \lambda_k$*, then any subspace of dimension* n *is* n*-optimal for* Y*.*

Remark 5. The sufficiency in Corollary 1, *a*) was established in [1].

Acknowledgements. The authors thank to the referee for his suggestions to improve this paper.

References

[1] A. ALDROUBI, C. CABRELLI, D. HARDIN, AND U. MOLTER, Optimal shift invariant spaces and their parseval frame generators, *Appl. Comput. Harmon. Anal.* **23**(2007), 273–283.

[2] A. ALDROUBI, C. CABRELLI, AND U. MOLTER, Optimal non-linear models for sparsity and sampling, *J. Fourier Anal. Appl.* **14**, 5(2008), 793–812.

[3] A. L. GARKAVI, On the best net and best section of a set in a normed space, *Izv. Akad. Nauk SSSR* **26**, 1(1962), 87–106.

[4] A. N. KOLMOGOROV, On the best approximation of functions of a given class, *Ann. of Math.* **37**(1936), 107–110.

[5] A. PINKUS, *On L_1 Approximation*, Cambridge tracts in mathematics 93, Cambridge University Press, 1989.

[6] I. SINGER, *Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces*, New York, Springer Verlag, 1970.

[7] I. SINGER, *Bases in Banach Spaces I*, New York, Springer Verlag, 1970.

[8] V. M. TIKHOMIROV, Diameters of sets in function spaces and the theory of best approximation, *Russian Math. Surveys* **15**(1960), 75–111.

Received July 24, 2009

H. H. Cuenya, F. E. Levis, M. D. Lorenzo, C. N. Rodriguez
Departamento de Matemática
Facultad de Ciencias Exactas Físico Químicas y Naturales
Universidad Nacional de Río Cuarto
Río Cuarto, 5800, ARGENTINA

E-mails: hcuenya@exa.unrc.edu.ar
flevis@exa.unrc.edu.ar
mlorenzo@exa.unrc.edu.ar
crodriguez@exa.unrc.edu.ar