SARAJEVO JOURNAL OF MATHEMATICS
Vol.5 (17) (2009), 109–118

ON THE PROXIMINALITY OF RIDGE FUNCTIONS

VUGAR E. ISMAILOV

ABSTRACT. Using two results of Garkavi, Medvedev and Khavinson [7], we give sufficient conditions for proximinality of sums of two ridge functions with bounded and continuous summands in the spaces of bounded and continuous multivariate functions respectively. In the first case, we give an example which shows that the corresponding sufficient condition cannot be made weaker for some subsets of $\mathbb{R}^n$. In the second case, we obtain also a necessary condition for proximinality. All the results are illuminated by numerous examples. The results, examples and following discussions naturally lead us to a conjecture on the proximinality of the considered class of ridge functions.

0. INTRODUCTION

In multivariate approximation theory, special functions called *ridge functions* are widely used. A ridge function is a multivariate function of the form $g(\mathbf{a} \cdot \mathbf{x})$, where g is a univariate function, $\mathbf{a} = (a_1, \ldots, a_n)$ is a vector (direction) different from zero, $\mathbf{x} = (x_1, \ldots, x_n)$ is the variable and $\mathbf{a} \cdot \mathbf{x}$ is the inner product. In other words, a ridge function is a composition of a univariate function with a linear functional over $\mathbb{R}^n$. These functions arise naturally in various fields. They arise in partial differential equations (where they are called *plane waves* [15]), in computerized tomography (see, e.g., [19,22]; the name ridge function was coined by Logan and Shepp[19] in one of the seminal papers on tomography), in statistics (especially, in the theory of projection pursuit and projection regression; see, e.g., [4,11]). Ridge functions are also the underpinnings of many central models in neural networks which has become increasing more popular in computer science, statistics, engineering, physics, etc. (see [24] and references therein). We refer the reader to Pinkus [23] for various motivations for the study of ridge functions and ridge function approximation.

2000 *Mathematics Subject Classification.* 41A30, 41A50, 41A63.
Key words and phrases. Ridge function, extremal element, proximinality, path, orbit.
This research was supported by INTAS Grant YSF-06-1000015-6283.

Let E be a normed linear space and F be a subspace. We say that F is proximinal in E if for any element $e \in E$ there exists at least one element $f_0 \in F$ such that

$$\|e - f_0\| = \inf_{f \in F} \|e - f\| .$$

In this case, the element f_0 is said to be extremal to e.

Although at present there are a great deal of interesting papers devoted to the approximation by ridge functions (see, e.g., [2,3,5,9,12,13,17,18,20,24, 25]), some problems of this approximation have not been solved completely yet. In the following, we are going to deal with one of such problems, namely with the problem of proximinality of the set of linear combinations of ridge functions in the spaces of bounded and continuous functions respectively. This problem will be considered in the simplest case when the class of approximating functions is the set

$$\mathcal{R} = \mathcal{R}\left(\mathbf{a}^1,\mathbf{a}^2\right) = \left\{g_1\left(\mathbf{a}^1 \cdot \mathbf{x}\right) + g_2\left(\mathbf{a}^2 \cdot \mathbf{x}\right) : g_i : \mathbb{R} \to \mathbb{R}, i = 1,2\right\}.$$

Here $\mathbf{a}^1$ and $\mathbf{a}^2$ are fixed directions and we vary over g_i. It is clear that this is a linear space. Consider the following three subspaces of $\mathcal{R}$. The first is obtained by taking only bounded sums $g_1\left(\mathbf{a}^1 \cdot \mathbf{x}\right) + g_2\left(\mathbf{a}^2 \cdot \mathbf{x}\right)$ over some set X in $\mathbb{R}^n$. We denote this subspace by $\mathcal{R}_a(X)$. The second and the third are subspaces of $\mathcal{R}$ with bounded and continuous summands $g_i\left(\mathbf{a}^i \cdot \mathbf{x}\right)$, $i = 1, 2$, on X respectively. These subspaces will be denoted by $\mathcal{R}_b(X)$ and $\mathcal{R}_c(X)$. In the case of $\mathcal{R}_c(X)$, the set X is considered to be compact.

Let $B(X)$ and $C(X)$ be the spaces of bounded and continuous multivariate functions over X respectively. What conditions must one impose on X in order that the sets $\mathcal{R}_a(X)$ and $\mathcal{R}_b(X)$ be proximinal in $B(X)$ and the set $\mathcal{R}_c(X)$ be proximinal in $C(X)$? We are also interested in necessary conditions for proximinality. It follows from a result of Garkavi, Medvedev and Khavinson (see Theorem 1 [7]) that $\mathcal{R}_a(X)$ is proximinal in $B(X)$ for all subsets X of $\mathbb{R}^n$. There is also an answer (see theorem 2 [7]) for proximinality of $\mathcal{R}_b(X)$ in $B(X)$. This will be discussed in Section 1. Is the set $\mathcal{R}_b(X)$ always proximinal in $B(X)$? There is an an example of a set $X \subset \mathbb{R}^n$ and a bounded function f on X for which there does not exist an extremal element in $\mathcal{R}_b(X)$.

In Section 2, we will obtain sufficient conditions for the existence of extremal elements from $\mathcal{R}_c(X)$ to an arbitrary function $f \in C(X)$. Based on a result of Marshall and O'Farrell [21], we will also give a necessary condition for proximinality of $\mathcal{R}_c(X)$ in $C(X)$. All the theorems, following discussions and examples of this paper will lead us naturally to a conjecture on the proximinality of the subspaces $\mathcal{R}_b(X)$ and $\mathcal{R}_c(X)$ in the spaces $B(X)$ and $C(X)$ respectively.

At the end of this section, we want to draw the readers attention to the more general case in which the number of directions is more than two. In this case, the set of approximating functions is

$$\mathcal{R}\left(\mathbf{a}^1,\dots,\mathbf{a}^r\right)=\left\{\sum_{i=1}^{r} g_i\left(\mathbf{a}^i\cdot\mathbf{x}\right) : g_i:\mathbb{R}\to\mathbb{R}, i=1,\dots,r\right\}.$$

In a similar way as above, one can define the sets $\mathcal{R}_a(X)$, $\mathcal{R}_b(X)$ and $\mathcal{R}_c(X)$. Using the results of [7], one can obtain sufficient (but not necessary) conditions for proximinality of these sets. This needs, besides paths (see Section 1), the consideration of some additional and more complicated relations between points of X. The case $r \geq 3$ will not be considered in the current paper, since our main purpose is to draw readers' attention to the given problems of proximinality in the simplest case of approximation. For the existing open problems connected with the set $\mathcal{R}\left(\mathbf{a}^1,\dots,\mathbf{a}^r\right)$, where $r \geq 3$, see [13] and [23].

1. Proximinality of $\mathcal{R}_b(X)$ in $B(X)$

We begin this section with the definition of a path with respect to two different directions $\mathbf{a}^1$ and $\mathbf{a}^2$. A path with respect to the directions $\mathbf{a}^1$ and $\mathbf{a}^2$ is a finite or infinite ordered set of points $(\mathbf{x}^1,\mathbf{x}^2,\dots)$ in $\mathbb{R}^n$ with the units $\mathbf{x}^{i+1}-\mathbf{x}^i$, $i=1,2,\dots$, in the directions perpendicular alternatively to $\mathbf{a}^1$ and $\mathbf{a}^2$. In the sequel, we simply use the term "path" instead of the long expression "path with respect to the directions $\mathbf{a}^1$ and $\mathbf{a}^2$". The length of a path is the number of its points and can be equal to ∞ if the path is infinite. A singleton is a path of the unit length. We say that a path $\left(\mathbf{x}^1,\dots,\mathbf{x}^m\right)$ belonging to some subset X of $\mathbb{R}^n$ is irreducible if there is not another path $\left(\mathbf{y}^1,\dots,\mathbf{y}^l\right)\subset X$ with $\mathbf{y}^1=\mathbf{x}^1$, $\mathbf{y}^l=\mathbf{x}^m$ and $l<m$. If in a path $\left(\mathbf{x}^1,\dots,\mathbf{x}^m\right)$ m is an even number and the set $\left(\mathbf{x}^1,\dots,\mathbf{x}^m,\mathbf{x}^1\right)$ is also a path, then the path $\left(\mathbf{x}^1,\dots,\mathbf{x}^m\right)$ is said to be closed. The notion of a path in the case when the directions $\mathbf{a}^1$ and $\mathbf{a}^2$ are basis vectors in $\mathbb{R}^2$ was first introduced by Diliberto and Straus [6] and exploited further in a number of works devoted to the approximation of bivariate functions by univariate functions (see, for example, [1,8,10,14,21]). Braess and Pinkus [2] used the notion in their solution to one problem of interpolation by ridge functions. It also appeared in characterization and construction of an extremal element from the set $\mathcal{R}_c(X)$ to a given continuous multivariate function (see [13]).

The following theorem follows from theorem 2 of [7]:

Theorem 1.1. *Let $X\subset\mathbb{R}^n$ and the lengths of all irreducible paths in X be uniformly bounded by some positive integer. Then each function in $B(X)$ has an extremal element in $\mathcal{R}_b(X)$.*

There are a large number of sets in $\mathbb{R}^n$ satisfying the hypothesis of this theorem. For example, if a set X has a cross section determined by one of the directions $\mathbf{a}^1$ or $\mathbf{a}^2$, then the set X satisfies the hypothesis of Theorem 1.1. By a cross section determined by the direction $\mathbf{a}^1$ we mean any set $X_{\mathbf{a}^1} = \{x \in X : \mathbf{a}^1 \cdot \mathbf{x} = c\}, c \in \mathbb{R}$, with the property: for any $\mathbf{y} \in X$ there exists a point $\mathbf{y}^1 \in X_{\mathbf{a}^1}$ such that $\mathbf{a}^2 \cdot \mathbf{y} = \mathbf{a}^2 \cdot \mathbf{y}^1$. By the similar way, one can define a cross section determined by the direction $\mathbf{a}^2$. Regarding Theorem 1.1, one may ask if the condition of the theorem is necessary for proximinality of $\mathcal{R}_b(X)$ in $B(X)$. While we do not know a complete answer to this question, we are going to give an example of a set X for which Theorem 1.1 fails. Let $\mathbf{a}^1 = (1;-1),\ \mathbf{a}^2 = (1;1)$. Consider the set

$$X = \{(2;\frac{2}{3}), (\frac{2}{3};-\frac{2}{3}), (0;0), (1;1), (1+\frac{1}{2};1-\frac{1}{2}), (1+\frac{1}{2}+\frac{1}{4};1-\frac{1}{2}+\frac{1}{4}), (1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8};1-\frac{1}{2}+\frac{1}{4}-\frac{1}{8}), \dots\}.$$

In what follows, the elements of X in the given order will be denoted by $\mathbf{x}^0, \mathbf{x}^1, \mathbf{x}^2, \dots$. It is clear that X is a path of the infinite length and $\mathbf{x}^n \to \mathbf{x}^0$, as $n \to \infty$. Let $\sum_{n=1}^{\infty} c_n$ be any divergent series with the terms $c_n > 0$ and $c_n \to 0$, as $n \to \infty$. In addition let f_0 be a function vanishing at the points $\mathbf{x}^0, \mathbf{x}^2, \mathbf{x}^4, \dots$, and taking values $c_1, c_2, c_3, \dots$ at the points $\mathbf{x}^1, \mathbf{x}^3, \mathbf{x}^5, \dots$ respectively. It is obvious that f_0 is continuous on X. The set X is compact and satisfies all the conditions of Proposition 2 of [21]. By this proposition, $\overline{\mathcal{R}_c(X)} = C(X)$. Therefore, for any continuous function on X, thus for f_0,

$$\inf_{g \in \mathcal{R}_c(X)} \|f_0 - g\|_{C(X)} = 0. \tag{1.1}$$

Since $\mathcal{R}_c(X) \subset \mathcal{R}_b(X)$, we obtain from (1.1) that

$$\inf_{g \in \mathcal{R}_b(X)} \|f_0 - g\|_{B(X)} = 0. \tag{1.2}$$

Suppose that f_0 has an extremal element $g_1^0 \left(\mathbf{a}^1 \cdot \mathbf{x}\right) + g_2^0 \left(\mathbf{a}^2 \cdot \mathbf{x}\right)$ in $\mathcal{R}_b(X)$. By the definition of $\mathcal{R}_b(X)$, the ridge functions $g_i^0, i = 1, 2$, are bounded on X. From (1.2) it follows that $f_0 = g_1^0 \left(\mathbf{a}^1 \cdot \mathbf{x}\right) + g_2^0 \left(\mathbf{a}^2 \cdot \mathbf{x}\right)$. Since $\mathbf{a}^1 \cdot \mathbf{x}^{2n} = \mathbf{a}^1 \cdot \mathbf{x}^{2n+1}$ and $\mathbf{a}^2 \cdot \mathbf{x}^{2n+1} = \mathbf{a}^2 \cdot \mathbf{x}^{2n+2}$, for $n = 0, 1, \dots$, we can write

$$\sum_{n=0}^{k} c_{n+1} = \sum_{n=0}^{k} \left[f_0(\mathbf{x}^{2n+1}) - f_0(\mathbf{x}^{2n})\right]$$

$$= \sum_{n=0}^{k} \left[g_2^0(\mathbf{a}^2 \cdot \mathbf{x}^{2n+1}) - g_2^0(\mathbf{a}^2 \cdot \mathbf{x}^{2n})\right] = g_2^0(\mathbf{a}^2 \cdot \mathbf{x}^{2k+1}) - g_2^0(\mathbf{a}^2 \cdot \mathbf{x}^0). \tag{1.3}$$

Since $\sum_{n=1}^{\infty} c_n = \infty$, we deduce from (1.3) that the function $g_2^0 \left(\mathbf{a}^2 \cdot \mathbf{x}\right)$ is not bounded on X. This contradiction means that the function f_0 does not have an extremal element in $\mathcal{R}_b(X)$. Therefore, the space $\mathcal{R}_b(X)$ is not proximinal in $B(X)$.

Remark. The above example is a slight generalization and an adaptation of Havinson's example (see [10]) to our case.

2. Proximinality of $\mathcal{R}_c(X)$ in $C(X)$

In this section, we are going to give sufficient conditions and also a necessary condition for proximinality of $\mathcal{R}_c(X)$ in $C(X)$.

Theorem 2.1. *Let the system of independent vectors $\mathbf{a}^1$ and $\mathbf{a}^2$ has a complement to a basis $\{\mathbf{a}^1, \ldots, \mathbf{a}^n\}$ in $\mathbb{R}^n$ with the property: for any point $\mathbf{x}^0 \in X$ and any positive real number δ there exist a number $\delta_0 \in (0, \delta]$ and a point $\mathbf{x}^\sigma$ in the set*

$$\sigma = \{\mathbf{x} \in X : \mathbf{a}^2 \cdot \mathbf{x}^0 - \delta_0 \leq \mathbf{a}^2 \cdot \mathbf{x} \leq \mathbf{a}^2 \cdot \mathbf{x}^0 + \delta_0\},$$

such that the system

$$\begin{cases} \mathbf{a}^2 \cdot \mathbf{x}' = \mathbf{a}^2 \cdot \mathbf{x}^\sigma \\ \mathbf{a}^1 \cdot \mathbf{x}' = \mathbf{a}^1 \cdot \mathbf{x} \\ \sum_{i=3}^n \left|\mathbf{a}^i \cdot \mathbf{x}' - \mathbf{a}^i \cdot \mathbf{x}\right| < \delta \end{cases} \tag{2.1}$$

has a solution $\mathbf{x}' \in \sigma$ for all points $\mathbf{x} \in \sigma$. Then the space $\mathcal{R}_c(X)$ is proximinal in $C(X)$.

Proof. Introduce the following mappings and sets:

$$\pi_i : X \to \mathbb{R},\ \pi_i(\mathbf{x}) = \mathbf{a}^i \cdot \mathbf{x},\ Y_i = \pi_i(X),\ i = 1, \ldots, n.$$

Since the system of vectors $\{\mathbf{a}^1, \ldots, \mathbf{a}^n\}$ is linearly independent, the mapping $\pi = (\pi_1, \ldots \pi_n)$ is an injection from X into the Cartesian product $Y_1 \times \cdots \times Y_n$. Besides, π is linear and continuous. By the open mapping theorem, the inverse mapping π^{-1} is continuous from $Y = \pi(X)$ onto X. Let f be a continuous function on X. Then the composition $f \circ \pi^{-1}(y_1, \ldots y_n)$ will be continuous on Y, where $y_i = \pi_i(\mathbf{x})$, $i = 1, \ldots, n$, are the coordinate functions. Consider the approximation of the function $f \circ \pi^{-1}$ by elements from

$$G_0 = \{g_1(y_1) + g_2(y_2) : \ g_i \in C(Y_i),\ i = 1, 2\}$$

over the compact set Y. Then one may observe that the function f has an extremal element in $\mathcal{R}_c(X)$ if and only if the function $f \circ \pi^{-1}$ has an extremal element in G_0. Thus the problem of proximinality of $\mathcal{R}_c(X)$ in $C(X)$ is reduced to the problem of proximinality of G_0 in $C(Y)$.

Let $T, T_1, \dots, T_{m+1}$ be metric compact spaces and $T \subset T_1 \times \cdots \times T_{m+1}$. For $i = 1, \dots, m$, let φ_i be the continuous mappings from T onto T_i. In [7], the authors obtained sufficient conditions for proximinality of the set

$$C_0 = \left\{ \sum_{i=1}^{n} g_i \circ \varphi_i : \ g_i \in C(T_i), \ i = 1, \dots m \right\}$$

in the space $C(T)$ of continuous functions on T. Since $Y \subset Y_1 \times Y_2 \times Z_3$, where $Z_3 = Y_3 \times \cdots \times Y_n$, we can use this result in our case for the approximation of the function $f \circ \pi^{-1}$ by elements from G_0. By this theorem, the set G_0 is proximinal in $C(Y)$ if for any $y_2^0 \in Y_2$ and $\delta > 0$ there exists a number $\delta_0 \in (0, \delta)$ such that the set $\sigma(y_2^0, \delta_0) = [y_2^0 - \delta_0, y_2^0 + \delta_0] \cap Y_2$ has $(2, \delta)$ maximal cross section. Hence there exists a point $y_2^\sigma \in \sigma(y_2^0, \delta_0)$ with the property: for any point $(y_1, y_2, z_3) \in Y$, with the second coordinate y_2 from the set $\sigma(y_2^0, \delta_0)$, there exists a point $(y_1', y_2^\sigma, z_3') \in Y$ such that $y_1 = y_1'$ and $\rho(z_3, z_3') < \delta$, where ρ is a metrics in Z_3. Since these conditions are equivalent to the conditions of Theorem 2.1, the space G_0 is proximinal in the space $C(Y)$. Then by the above conclusion, the space $\mathcal{R}_c(X)$ is proximinal in $C(X)$. □

Let us give some simple examples of compact sets satisfying the hypothesis of Theorem 2.1. For the sake of brevity, we restrict ourselves to the case $n = 3$.

(a) Let X be a closed ball in $\mathbb{R}^3$, a^1 and a^2 be two arbitrary orthogonal directions. Then Theorem 2.1 holds. Note that in this case, we can take $\delta_0 = \delta$ and a^3 as an orthogonal vector to both the vectors a^1 and a^2.

(b) Let X be the unit cube, $a^1 = (1; 1; 0)$, $a^2 = (1; -1; 0)$. Then Theorem 2.1 also holds. In this case, we can take $\delta_0 = \delta$ and $a^3 = (0; 0; 1)$. Note that the unit cube does not satisfy the hypothesis of the theorem for many directions (take, for example, $a^1 = (1; 2; 0)$ and $a^2 = (2; -1; 0)$).

In the following example, one can not always chose δ_0 to be equal to δ.

(c) Let $X = \{(x_1, x_2, x_3) : \ (x_1, x_2) \in Q, \ 0 \le x_3 \le 1\}$, where Q is the union of two triangles $A_1B_1C_1$ and $A_2B_2C_2$ with the vertices $A_1 = (0; 0)$, $B_1 = (1; 2)$, $C_1 = (2; 0)$, $A_2 = (1\frac{1}{2}; 1)$, $B_2 = (2\frac{1}{2}; -1)$, $C_2 = (3\frac{1}{2}; 1)$. Let $a^1 = (0; 1; 0)$ and $a^2 = (1; 0; 0)$. Then it is easy to see that Theorem 2.1 holds (the vector a^3 can be chosen as $(0; 0; 1)$). In this case, δ_0 can not be always chosen as equal to δ. Take, for example, $\mathbf{x}^0 = (1\frac{3}{4}; 0; 0)$ and $\delta = 1\frac{3}{4}$. If $\delta_0 = \delta$, then the second equation of the system (2.1) has not a solution for a point $(1; 2; 0)$ or a point $(2\frac{1}{2}; -1; 0)$. But if we take δ_0 not more than $\frac{1}{4}$, then for

$\mathbf{x}^{\sigma} = \mathbf{x}^0$ the system has a solution. Note that the last inequality $\left|\mathbf{a}^3 \cdot \mathbf{x}' - \mathbf{a}^3 \cdot \mathbf{x}\right| < \delta$ of the system can be satisfied with the equality $\mathbf{a}^3 \cdot \mathbf{x}' = \mathbf{a}^3 \cdot \mathbf{x}$ if $a^3 = (0; 0; 1)$.

It should be remarked that the results of [7] tell nothing about necessary conditions for proximinality of the spaces considered there. To fill this gap in our case, we want to give a necessary condition for proximinality of $\mathcal{R}_c(X)$ in $C(X)$. Our result will be based on the result of Marshall and O'Farrell given below. First, let us introduce some notation. By $\mathcal{R}_c^i$, $i = 1, 2$, we will denote the set of continuous ridge functions $g\left(\mathbf{a}^i \cdot \mathbf{x}\right)$ on the given compact set $X \subset \mathbb{R}^n$. Note that $\mathcal{R}_c = \mathcal{R}_c^1 + \mathcal{R}_c^2$. Besides, let $\mathcal{R}_c^3 = \mathcal{R}_c^1 \cap \mathcal{R}_c^2$. For $i = 1, 2, 3$, let X_i be the quotient space obtained by identifying points y_1 and y_2 in X whenever $f(y_1) = f(y_2)$ for each f in $\mathcal{R}_c^i$. By π_i denote the natural projection of X onto X_i, $i = 1, 2, 3$. Note that we have already dealt with the quotient spaces X_1, X_2 and the projections π_1, π_2 in the previous section (see the proof of Theorem 2.1). The relation on X, defined by setting $y_1 \approx y_2$ if y_1 and y_2 belong to some path, is an equivalence relation. Following Marshall and O'Farrell [21] the equivalence classes are called orbits. By $O(t)$ denote the orbit of X containing t. For $Y \subset X$, let $\omega_Y f$ be the oscillation of a function f on the set Y. That is,

$$\omega_Y f = \sup_{x, y \in Y} \left| f(x) - f(y) \right|.$$

Theorem 2.2. *Suppose that the space $\mathcal{R}_c(X)$ is proximinal in $C(X)$. Then there exists a positive real number c such that*

$$\sup_{t \in X} \ \omega_{O(t)} f \leq c \sup_{t \in X} \ \omega_{\pi_2^{-1}(\pi_2(t))} f \tag{2.2}$$

for all f in $\mathcal{R}_c^1$.

The proof is simple. In [21], Marshall and O'Farrell proved the following result (see Proposition 4 in [21]): Let A_1 and A_2 be closed subalgebras of $C(X)$ that contain the constants. Let (X_1, π_1), (X_2, π_2) and (X_3, π_3) be the quotient spaces and projections associated with the algebras A_1, A_2 and $A_3 = A_1 \cap A_2$ respectively. Then $A_1 + A_2$ is closed in $C(X)$ if and only if there exists a positive real number c such that

$$\sup_{z \in X_3} \ \omega_{\pi_3^{-1}(z)} f \leq c \sup_{y \in X_2} \ \omega_{\pi_2^{-1}(y)} f \tag{2.3}$$

for all f in A_1.

If $\mathcal{R}_c(X)$ is proximinal in $C(X)$, then it is necessarily closed and therefore, by the above proposition, (2.3) holds for the algebras $A_1^i = \mathcal{R}_c^i$, $i = 1, 2, 3$. The right-hand side of (2.3) is equal to the right-hand side of (2.2). Let t be some point in X and $z = \pi_3(t)$. Since each function $f \in \mathcal{R}_c^3$ is constant on the orbit of t (note that f is both of the form $g_1\left(\mathbf{a}^1\cdot\mathbf{x}\right)$ and of the form $g_2\left(\mathbf{a}^2\cdot\mathbf{x}\right)$), $O(t)\subset\pi_3^{-1}(z)$. Hence,

$$\sup_{t\in X}\ \omega_{O(t)}f\leq c\sup_{z\in X_3}\ \omega_{\pi_3^{-1}(z)}f. \tag{2.4}$$

From (2.3) and (2.4) we obtain (2.2).

Note that the inequality (2.3) provides not weaker but a less practicable necessary condition for proximinality than the inequality (2.2) does. On the other hand, there are many cases in which both the inequalities are equivalent. For example, let the lengths of irreducible paths of X be bounded by some positive integer n_0. In this case, it can be shown that the inequality (2.3), hence (2.2), holds with the constant $c=\frac{n_0}{2}$ and moreover $O(t)=\pi_3^{-1}(z)$ for all $t\in X$, where $z=\pi_3(t)$ (see the proof of Theorem 5 in [13]). Therefore, the inequalities (2.2) and (2.3) are equivalent for the considered class of sets X. The last argument shows that all the compact sets $X\subset\mathbb{R}^n$ over which $\mathcal{R}_c(X)$ is not proximinal in $C(X)$ should be sought in the class of sets having irreducible paths consisting sufficiently large number of points. For example, let $I=[0;1]^2$ be the unit square, $a^1=(1;1)$, $a^2=(1;\frac{1}{2})$. Consider the path

$$l_k=\{(1;0),(0;1),(\frac{1}{2};0),(0;\frac{1}{2}),(\frac{1}{4};0),\dots,(0;\frac{1}{2^k})\}.$$

It is clear that l_k is an irreducible path with the length $2k+2$, where k may be very large. Let g_k be a continuous univariate function on $\mathbb{R}$ satisfying the conditions: $g_k(\frac{1}{2^{k-i}})=i,\ i=0,\dots,k$, $g_k(t)=0$ if $t<\frac{1}{2^k},\ i-1\leq g_k(t)\leq i$ if $t\in(\frac{1}{2^{k-i+1}},\frac{1}{2^{k-i}}),\ i=1,\dots,k$, and $g_k(t)=k$ if $t>1$. Then it can be easily verified that

$$\sup_{t\in X}\ \omega_{\pi_2^{-1}(\pi_2(t))}g_k(\mathbf{a}^1\cdot\mathbf{x})\leq 1. \tag{2.5}$$

Since $\max_{\mathbf{x}\in I}g_k(\mathbf{a}^1\cdot\mathbf{x})=k$, $\min_{\mathbf{x}\in I}g_k(\mathbf{a}^1\cdot\mathbf{x})=0$ and $\omega_{\mathbf{x}\in O(t_1)}g_k(\mathbf{a}^1\cdot\mathbf{x})=k$ for $t_1=(1;0)$, we obtain that

$$\sup_{t\in X}\ \omega_{O(t)}g_k(\mathbf{a}^1\cdot\mathbf{x})=k. \tag{2.6}$$

Since k may be very large, from (2.5) and (2.6) it follows that the inequality (2.2) cannot hold for the function $g_k(\mathbf{a}^1\cdot\mathbf{x})\in\mathcal{R}_c^1$. Thus the space $\mathcal{R}_c(I)$ with the directions $a^1=(1;1)$ and $a^2=(1;\frac{1}{2})$ is not proximinal in $C(I)$.

It should be remarked that if a compact set $X\subset\mathbb{R}^n$ satisfies the hypothesis of Theorem 2.1, then the length of all irreducible paths are uniformly bounded (see the proof of Theorem 2.1 and lemma in [7]). We have already seen that if the last condition does not hold, then the proximinality of both $\mathcal{R}_c(X)$ in $C(X)$ and $\mathcal{R}_b(X)$ in $B(X)$ fail for some sets X. Besides the examples given above and in Section 1, one can easily construct many other examples of such sets. All these examples, Theorems 1.1, 2.1, 2.2 and the following remarks justify the statement of the following conjecture:

Conjecture. Let X be some subset of $\mathbb{R}^n$. The space $\mathcal{R}_b(X)$ is proximinal in $B(X)$ and the space $\mathcal{R}_c(X)$ is proximinal in $C(X)$ (in this case, X is considered to be compact) if and only if the lengths of all irreducible paths of X are uniformly bounded.

Remark 1. After completion of this work, Medvedev's result came to our attention (see [16, p.58]). His result, in particular, states that the set $R_c(X)$ is closed in $C(X)$ if and only if the lengths of all irreducible paths of X are uniformly bounded. Thus, in the case of $C(X)$, the necessity of the above conjecture was proved by Medvedev.

Remark 2. Note that there are situations in which a continuous function (a specially chosen function on a specially constructed set) has an extremal element in $\mathcal{R}_b(X)$, but not in $\mathcal{R}_c(X)$ (see [16, p.73]). One subsection of [16] (see p.68) is devoted to the proximinality of sums of two univariate functions with continuous and bounded summands in the spaces of continuous and bounded bivariate functions respectively. If $X \subset \mathbb{R}^2$ and $\mathbf{a}^1, \mathbf{a}^2$ be linearly independent directions in $\mathbb{R}^2$, then the linear transformation $y_1 = \mathbf{a}^1 \cdot \mathbf{x}$, $y_2 = \mathbf{a}^2 \cdot \mathbf{x}$ reduces the problems of proximinality of $\mathcal{R}_b(X)$ in $B(X)$ and $\mathcal{R}_c(X)$ in $C(X)$ to the problems considered in that subsection. But in general, when $X \subset \mathbb{R}^n$, $n > 2$, our case cannot be obtained from that of [16].

Acknowledgment. I learned about the monograph by Khavinson [16] from Allan Pinkus at the Technion. Using this opportunity, I would like to express my sincere gratitude to him.

References

[1] M–B. A. Babaev, *Estimates and ways for determining the exact value of the best approximation of functions of several variables by superpositions of functions of a smaller number of variables*, (Russian), *Special questions in the theory of functions*, (Russian), Izdat. "Elm", Baku, 1977, 3–23.

[2] D. Braess and A. Pinkus, *Interpolation by ridge functions*, J. Approx. Theory, 73 (1993), 218–236.

[3] M. D. Buhmann and A. Pinkus, *Identifying linear combinations of ridge functions*, Adv. Appl. Math., 22 (1999), 103–118.

[4] E. J. Candes, *Ridgelets: estimating with ridge functions* Ann. Stat., 31 (2003), 1561–1599.

[5] C. K. Chui and X. Li, *Approximation by ridge functions and neural networks with one hidden layer*, J. Approx. Theory, 70 (1992), 131–141.

[6] S. P. Diliberto and E. G. Straus, *On the approximation of a function of several variables by the sum of functions of fewer variables*, Pacific J. Math., 1 (1951), 195–210.

[7] A. L. Garkavi, V. A. Medvedev and S. Ya. Khavinson, *On the existence of a best uniform approximation of a function of several variables by the sum of functions of fewer variables*, Matematicheski Sbornik, 187 (1996), 3–14; English transl. in Sbornik Mathematics, 187 (1996), 623–634.
[8] M. V. Golitschek and W. A. Light, *Approximation by solutions of the planar wave equation*, SIAM J. Numer. Anal., 29 (1992), 816–830.
[9] Y. Gordon, V. Maiorov, M. Meyer and S. Reisner, *On the best approximation by ridge functions in the uniform norm*, Constr. Approx., 18 (2002), 61–85.
[10] S. Ja. Havinson, *A Chebyshev theorem for the approximation of a function of two variables by sums of the type* $\varphi(x) + \psi(y)$, Izv. Acad. Nauk. SSSR Ser. Mat., 33 (1969), 650–666; English tarnsl. in Math. USSR Izv., 3 (1969), 617–632.
[11] P. J. Huber, *Projection pursuit*, Ann. Stat., 13 (1985), 435–475.
[12] V. E. Ismailov, *A note on the best* L_2 *approximation by ridge functions*, Appl. Math. E-Notes, 7 (2007), 71–76.
[13] V. E. Ismailov, *Characterization of an extremal sum of ridge functions*, J. Comp. Appl. Math., 205 (2007), 105–115.
[14] V. E. Ismailov, *Methods for computing the least deviation from the sums of functions of one variable*, Sibirski Matematicheski Zhurnal, 47 (2006), 1076–1082; English transl. in Sib. Math. J., 47 (2006), 883–888.
[15] F. John, *Plane Waves and Spherical Means Applied to Partial Differential Equations*, Interscience, New York, 1955.
[16] S. Ya. Khavinson, *Best approximation by linear superpositions (approximate nomography)*, Translated from the Russian manuscript by D. Khavinson, Translations of Mathematical Monographs, 159. American Mathematical Society, Providence, RI, 1997, 175 pp.
[17] A. Kroo, *On approximation by ridge functions*, Constr. Approx., 13 (1997), 447–460.
[18] V. Ya Lin and A. Pinkus, *Fundamentality of ridge functions*, J. Approx. Theory, 75 (1993), 295–311.
[19] B. F. Logan and L. A. Shepp, *Optimal reconstruction of a function from its projections*, Duke Math. J., 42 (1975), 645–659.
[20] V. E. Maiorov, *On best approximation by ridge functions*, J. Approx. Theory, 99 (1999), 68–94.
[21] D. E. Marshall and A. G. O'Farrell, *Uniform approximation by real functions*, Fundam. Math., 104 (1979), 203–211.
[22] F. Natterer, *The Mathematics of Computerized Tomography*, Wiley, New York, 1986.
[23] A. Pinkus, *Approximating by ridge functions, in: Surface Fitting and Multiresolution Methods*, (A.Le Méhauté, C.Rabut and L.L.Schumaker, eds), Vanderbilt Univ. Press (Nashville), 1997, 279–292.
[24] A. Pinkus, *Approximation theory of the MLP model in neural networks*, Acta Numerica, 8 (1999), 143–195.
[25] Y. Xu, W. A. Light and E. W. Cheney, *Constructive methods of approximation by ridge functions and radial functions*, Numer. Algorithms, 4 (1993), 205–223.

(Received: March 12, 2008)
(Revised: January 13, 2009)

Mathematics and Mechanics Institute
Azerbaijan National Academy of Sciences
9, F. Agaev str., Az-1141, Baku, Azerbaijan
E–mail: vugaris@mail.ru