STUDIA MATHEMATICA 183 (2) (2007)

On the approximation by compositions of fixed multivariate functions with univariate functions

by

VUGAR E. ISMAILOV (Baku)

Abstract. The approximation in the uniform norm of a continuous function $f(\mathbf{x}) = f(x_1, \ldots, x_n)$ by continuous sums $g_1(h_1(\mathbf{x})) + g_2(h_2(\mathbf{x}))$, where the functions h_1 and h_2 are fixed, is considered. A Chebyshev type criterion for best approximation is established in terms of paths with respect to the functions h_1 and h_2.

1. Exposition of the problem. It is well known that in many problems of approximation of bivariate functions by sums of univariate functions the concept of a path is central. A *path* is a finite or infinite ordered set of points in the xy plane such that the line segments joining consecutive points are of positive length and are alternately parallel to the x and y axes. The idea of paths, in this context, was first introduced by Diliberto and Straus [4] and exploited further in a number of works, e.g. [5, 7, 8, 10, 13]. In connection with the problem of interpolation by linear combinations of ridge functions, Braess and Pinkus [1] introduced the notion of a path with respect to distinct directions $\mathbf{a}$ and $\mathbf{b}$. This is an ordered set of points $(\mathbf{v}^1, \ldots, \mathbf{v}^n) \subset \mathbb{R}^2$ with edges $\mathbf{v}^i\mathbf{v}^{i+1}$ in alternating directions $\mathbf{a}$ and $\mathbf{b}$. These objects give a geometric method for deciding if a set of points $\{\mathbf{x}^i\}_{i=1}^m$ has the NI-property (non-interpolation property) (see [1]).

Our aim is to bring into consideration more general objects: paths with respect to two continuous functions. We will show how these objects appear in the characterization of extremal elements in the approximation problem considered below.

Let Q be a compact set in $\mathbb{R}^n$. Consider the approximation of a function $f \in C(Q)$ by elements of the set

$$C_{h_1h_2} = C_{h_1h_2}(Q) = \{g \in C(Q) : g(\mathbf{x}) = g_1(h_1(\mathbf{x})) + g_2(h_2(\mathbf{x}))\},$$

where the functions $h_i \in C(Q)$ are prescribed and we vary over functions

2000 *Mathematics Subject Classification*: 41A30, 41A50, 41A65.
Key words and phrases: path, ridge function, best approximation.
Supported by INTAS Grant YSF-06-1000015-6283.

$g_i : h_i(Q) \to \mathbb{R}$, $i = 1, 2$. The continuity of g_i on $h_i(Q)$, $i = 1, 2$, is not necessary, but the sum $g_1(h_1(\mathbf{x})) + g_2(h_2(\mathbf{x}))$ should be continuous on Q. It is not difficult to see that linear combinations of functions from $C_{h_1h_2}$ belong to $C_{h_1h_2}$.

Our aim is to find necessary and sufficient geometrical conditions for a function $g_0 \in C_{h_1h_2}$ to be a best approximation to f, i.e. for

$$\|f - g_0\| = \max_{\mathbf{x}\in Q} |f(\mathbf{x}) - g_0(\mathbf{x})| = E(f, C_{h_1h_2}),$$

where

$$E(f) = E(f, C_{h_1h_2}) := \inf_{g\in C_{h_1h_2}} \|f - g\|$$

is the error in approximating from $C_{h_1h_2}(Q)$.

In multivariate approximation theory and in some applications such as computerized tomography, statistics, and neural networks, special functions called ridge functions are widely used (see, e.g., [1–3, 9, 11, 12, 14, 16–19]). A *ridge function* is a multivariate function of the form $g(\mathbf{a} \cdot \mathbf{x})$, where g is a univariate function, $\mathbf{a}$ is a fixed vector (direction) in $\mathbb{R}^n$ different from zero, $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{a} \cdot \mathbf{x}$ is the inner product of $\mathbf{a}$ and $\mathbf{x}$. Note that the problem of approximation by sums of two ridge functions with fixed directions is a special case of the problem considered here.

2. Main result. We begin with a definition. Let Q be a compact set in $\mathbb{R}^n$ and $h_i \in C(Q)$, $i = 1, 2$.

Definition 2.1. A finite or infinite set $p = (\mathbf{p}_1, \mathbf{p}_2, \dots) \subset Q$, where $\mathbf{p}_i \neq \mathbf{p}_{i+1}$, with either $h_1(\mathbf{p}_1) = h_1(\mathbf{p}_2)$, $h_2(\mathbf{p}_2) = h_2(\mathbf{p}_3)$, $h_1(\mathbf{p}_3) = h_1(\mathbf{p}_4), \dots$ or $h_2(\mathbf{p}_1) = h_2(\mathbf{p}_2)$, $h_1(\mathbf{p}_2) = h_1(\mathbf{p}_3)$, $h_2(\mathbf{p}_3) = h_2(\mathbf{p}_4), \dots$ is called a *path with respect to the functions h_1 and h_2.*

In the following, we will simply say "path" instead of "path with respect to the functions h_1 and h_2".

If in a finite path $(\mathbf{p}_1, \dots, \mathbf{p}_{n+1})$, $\mathbf{p}_{n+1} = \mathbf{p}_1$ and n is even, then the path $(\mathbf{p}_1, \dots, \mathbf{p}_n)$ is said to be *closed*. Note that a minimal closed path may consist of two distinct points $\mathbf{p}_1$ and $\mathbf{p}_2$. In this case, the equality $h_i(\mathbf{p}_1) = h_i(\mathbf{p}_2)$ must be satisfied for both $i = 1$ and $i = 2$.

To each closed path $p = (\mathbf{p}_1, \dots, \mathbf{p}_{2n})$ we associate the functional

$$G_p(f) = \frac{1}{2n} \sum_{k=1}^{2n} (-1)^{k+1} f(\mathbf{p}_k).$$

It has the following obvious properties:

(a) If $g \in C_{h_1h_2}$, then $G_p(g) = 0$.
(b) $\|G_p\| \le 1$ and if $\mathbf{p}_i \neq \mathbf{p}_j$ for all $i \neq j$, $1 \le i, j \le 2n$, then $\|G_p\| = 1$.

We need the following auxiliary lemmas.

LEMMA 2.2. *If a compact set Q contains closed paths, then*

$$\sup_{p\subset Q} |G_p(f)| \le E(f, C_{h_1h_2}), \tag{1}$$

where the sup is taken over all closed paths. Moreover, inequality (1) *is sharp, i.e., there exist functions for which* (1) *turns into equality.*

Proof. Let p be any path of Q and $g \in C_{h_1h_2}(Q)$. Then by the linearity of G_p and properties (a) and (b),

$$|G_p(f)| = |G_p(f-g)| \le \|f-g\|. \tag{2}$$

Since the left-hand and right-hand sides of (2) do not depend upon g and p, respectively, it follows from (2) that

$$\sup_{p\subset Q} |G_p(f)| \le \inf_{g\in C_{h_1h_2}} \|f-g\|. \tag{3}$$

To prove the sharpness of (1) note that if p is a closed path, then there is a closed path $p' = (\mathbf{p}'_1, \ldots, \mathbf{p}'_{2n})$ such that $p' \subset p$ and all points of p' are distinct. Indeed, p' can be obtained by the following simple algorithm: if the points of p are not all distinct, let i and $k > 0$ be the minimal indices such that $\mathbf{p}_i = \mathbf{p}_{i+2k}$; delete from p the subsequence $\mathbf{p}_{i+1}, \ldots, \mathbf{p}_{i+2k}$ and call the resulting path p; repeat the above step until all points of p are distinct; set $p' := p$. By Urysohn's lemma, there exists a continuous function f' such that $f'(\mathbf{p}'_i) = 1$ for $i = 1, 3, \ldots, 2n-1$, $f'(\mathbf{p}'_j) = -1$ for $j = 2, 4, \ldots, 2n$, and $-1 < f'(\mathbf{x}) < 1$ for all $\mathbf{x} \in Q \setminus p'$. Then

$$G_{p'}(f') = \|f'\| = 1. \tag{4}$$

On the other hand, it is obvious that

$$E(f', C_{h_1h_2}) \le \|f'\|. \tag{5}$$

From (3)–(5) it follows that

$$\sup_{p\subset Q} |G_p(f')| = E(f', C_{h_1h_2}),$$

and moreover sup is attained by the closed path p', so $0 \in C_{h_1h_2}$ is a best approximation to f'. ∎

For any $h \in C(Q)$, set

$$Q_t = \{\mathbf{x} \in Q : h(\mathbf{x}) = t\}, \qquad T_h = \{t \in \mathbb{R} : Q_t \ne \emptyset\}.$$

LEMMA 2.3. *Let Q be a convex compact set in $\mathbb{R}^n$ and $f, h \in C(Q)$. Then the functions*

$$g_1(t) = \max_{\substack{\mathbf{x}\in Q\\ h(\mathbf{x})=t}} f(\mathbf{x}), \qquad g_2(t) = \min_{\substack{\mathbf{x}\in Q\\ h(\mathbf{x})=t}} f(\mathbf{x}), \qquad t \in T_h,$$

are defined and continuous on T_h.

Proof. Observe that T_h is a closed interval or a point. The case of a point is trivial. So, assume that $T_h = [c_1, c_2]$, where $c_1 \neq c_2$. Suppose g_1 is not continuous on $[c_1, c_2]$ and t_0 is a point of discontinuity. First assume that $t_0 \in [c_1, c_2)$. Without loss of generality we may consider g_1 to be discontinuous from the right at t_0, i.e.

$$\exists \varepsilon > 0 \ \forall t' > t_0, t' \in [c_1, c_2] \ \exists t_1 \in (t_0, t'] : \ |g_1(t_1) - g_1(t_0)| > \varepsilon. \tag{6}$$

Fix some t'. Since f is continuous on the compact set Q, there exist $\mathbf{y}_0, \mathbf{y}_1 \in Q$ such that $g_1(t_0) = f(\mathbf{y}_0)$ and $g_1(t_1) = f(\mathbf{y}_1)$. Since Q is convex, it contains the line segment $[\mathbf{y}_0, \mathbf{y}_1]$. Set

$$Y_0 = \{\mathbf{y} \in Q : f(\mathbf{y}) = g_1(t_0)\}.$$

It is obvious that Y_0 is closed, $\mathbf{y}_0 \in Y_0$ and $\mathbf{y}_1 \notin Y_0$. Write

$$Y_0' = Y_0 \cap [\mathbf{y}_0, \mathbf{y}_1].$$

There is a point $\mathbf{y}_0' \in Y_0'$ such that

$$\varrho(\mathbf{y}_1, Y_0') = \varrho(\mathbf{y}_1, \mathbf{y}_0').$$

It is clear that $h(\mathbf{y}_0') = t_0$ and $h(\mathbf{y}_1) = t_1$. Since h is continuous on $[\mathbf{y}_0, \mathbf{y}_1]$, for any $t \in (t_0, t_1]$ there exists $\mathbf{y} \in (\mathbf{y}_0', \mathbf{y}_1]$ such that $h(\mathbf{y}) = t$. Then it is not difficult to see that there exist sequences $\{t_n\} \subset (t_0, t_1]$ and $\{\mathbf{y}_n\} \subset (\mathbf{y}_0', \mathbf{y}_1]$ such that $t_n \downarrow t_0$, $\mathbf{y}_n \to \mathbf{y}_0'$ and $h(\mathbf{y}_n) = t_n$. It follows from (6) that there exists a sequence $\{t_n'\}$ such that $t_0 < t_n' \leq t_n$ and at the same time

$$|g_1(t_n') - g_1(t_0)| > \varepsilon \quad \text{for all } n. \tag{7}$$

For each n there exist $\mathbf{y}_n' \in (\mathbf{y}_0', \mathbf{y}_n]$ and $\mathbf{y}_n'' \in Q$ such that $h(\mathbf{y}_n') = t_n'$ and $f(\mathbf{y}_n'') = g_1(t_n')$. Then (7) can be written in the following form:

$$|f(\mathbf{y}_n'') - f(\mathbf{y}_0')| > \varepsilon \quad \text{for all } n. \tag{8}$$

Since $h(\mathbf{y}_n'') = t_n'$ and $f(\mathbf{y}_n'')$ is the maximum of all $f(\mathbf{y})$, whereas $h(\mathbf{y}) = t_n'$, we find that

$$f(\mathbf{y}_n') \leq f(\mathbf{y}_n'') \quad \text{for all } n. \tag{9}$$

Since $\mathbf{y}_n \to \mathbf{y}_0'$, $\mathbf{y}_n'$ also tends to $\mathbf{y}_0'$. Then $f(\mathbf{y}_n') \to f(\mathbf{y}_0')$ and $h(\mathbf{y}_n') \to h(\mathbf{y}_0')$. The sequence $\{\mathbf{y}_n''\}$ contains a converging subsequence. Without loss of generality we may assume that $\{\mathbf{y}_n''\}$ itself converges to some point $\mathbf{y}'' \in Q$. Then we deduce from (8) and (9) that

$$\left| f(\mathbf{y}'') - f(\mathbf{y}_0') \right| \geq \varepsilon \tag{10}$$

and

$$f(\mathbf{y}_0') \leq f(\mathbf{y}''). \tag{11}$$

Let us prove that $f(\mathbf{y}'') = f(\mathbf{y}_0')$. Indeed, since $h(\mathbf{y}_n'') = t_n'$, $\mathbf{y}_n'' \to \mathbf{y}''$, $t_n' \to t_0$, it follows from the continuity of h that $h(\mathbf{y}'') = t_0$. Now, since

$h(\mathbf{y}_0') = h(\mathbf{y}'') = t_0$ and $f(\mathbf{y}_0')$ is the maximum of all $f(\mathbf{y})$, whereas $h(\mathbf{y}) = t_0$, it follows from (11) that

$$f(\mathbf{y}_0') = f(\mathbf{y}'').$$

The last equality together with (10) contradicts the choice of ε.

In the same way we can prove that g_1 is continuous at $t = c_2$ and g_2 is also continuous on T_h. ∎

DEFINITION 2.4. A finite or infinite path $(\mathbf{p}_1, \mathbf{p}_2, \ldots)$ is said to be *extremal* for a function $u \in C(Q)$ if $u(\mathbf{p}_i) = (-1)^i \|u\|$, $i = 1, 2, \ldots$, or $u(\mathbf{p}_i) = (-1)^{i+1}\|u\|$, $i = 1, 2, \ldots$

THEOREM 2.5. *Let Q be a convex compact set in $\mathbb{R}^n$. A necessary and sufficient condition for a function $g_0 \in C_{h_1h_2}$ to be a best approximation to the given function $f \in C(Q) \setminus C_{h_1h_2}$ is the existence of a closed or infinite path $l = (\mathbf{p}_1, \mathbf{p}_2, \ldots)$ extremal for the function $f_1 = f - g_0$.*

Proof. Necessity. Let $g_0(\mathbf{x}) = g_{1,0}(h_1(\mathbf{x})) + g_{2,0}(h_2(\mathbf{x})) \in C_{h_1h_2}(Q)$ be a best approximation. We must show that if there is no closed path extremal for f_1, then there exists a path extremal for f_1 with infinite length (number of points). Suppose that, on the contrary, there exists a positive integer N such that the length of each path extremal for f_1 is at most N and no path extremal for f_1 is closed. Define

$$f_0 = f, \qquad f_n = f_{n-1} - g_{1,n-1} - g_{2,n-1}, \qquad n = 2, 3, \ldots,$$

where

$$g_{1,n-1}(\mathbf{x}) = g_{1,n-1}(h_1(\mathbf{x})) = \frac{1}{2}\Big(\max_{\substack{\mathbf{y}\in Q\\ h_1(\mathbf{y})=h_1(\mathbf{x})}} f_{n-1}(\mathbf{y}) + \min_{\substack{\mathbf{y}\in Q\\ h_1(\mathbf{y})=h_1(\mathbf{x})}} f_{n-1}(\mathbf{y})\Big)$$

$$g_{2,n-1}(\mathbf{x}) = g_{2,n-1}(h_2(\mathbf{x})) = \frac{1}{2}\Big(\max_{\substack{\mathbf{y}\in Q\\ h_2(\mathbf{y})=h_2(\mathbf{x})}} (f_{n-1}(\mathbf{y}) - g_{1,n-1}(h_1(\mathbf{y}))) + \min_{\substack{\mathbf{y}\in Q\\ h_2(\mathbf{y})=h_2(\mathbf{x})}} (f_{n-1}(\mathbf{y}) - g_{1,n-1}(h_1(\mathbf{y})))\Big).$$

By Lemma 2.3, all the functions f_n, $n = 2, 3, \ldots$, are continuous on Q. By assumption g_0 is a best approximation to f. Hence $\|f_1\| = E(f)$. Now we show that $\|f_2\| = E(f)$. Indeed, for any $\mathbf{x} \in Q$,

$$f_1(\mathbf{x}) - g_{1,1}(h_1(\mathbf{x})) \le \frac{1}{2}\Big(\max_{\substack{\mathbf{y}\in Q\\ h_1(\mathbf{y})=h_1(\mathbf{x})}} f_1(\mathbf{y}) - \min_{\substack{\mathbf{y}\in Q\\ h_1(\mathbf{y})=h_1(\mathbf{x})}} f_1(\mathbf{y})\Big) \le E(f) \tag{12}$$

and

$$f_1(\mathbf{x}) - g_{1,1}(h_1(\mathbf{x})) \ge \frac{1}{2}\Big(\min_{\substack{\mathbf{y}\in Q\\ h_1(\mathbf{y})=h_1(\mathbf{x})}} f_1(\mathbf{y}) - \max_{\substack{\mathbf{y}\in Q\\ h_1(\mathbf{y})=h_1(\mathbf{x})}} f_1(\mathbf{y})\Big) \ge -E(f). \tag{13}$$

In the same way, using (12) and (13), it can be shown that for any $\mathbf{x} \in Q$,

$$-E(f) \le f_2(\mathbf{x}) = f_1(\mathbf{x}) - g_{1,1}(h_1(\mathbf{x})) - g_{2,1}(h_2(\mathbf{x})) \le E(f).$$

Therefore,

$$\|f_2\| \le E(f). \tag{14}$$

Since $f_2 - f$ belongs to $C_{h_1h_2}$, we deduce from (14) that $\|f_2\| = E(f)$. In the same way, one can show that $\|f_n\| = E(f)$ for any n.

We now prove that if $f_1(\mathbf{p}_0) < E(f)$ for some $\mathbf{p}_0 \in Q$, then $f_2(\mathbf{p}_0) < E(f)$. We first prove that if $f_1(\mathbf{p}_0) < E(f)$, then

$$f_1(\mathbf{p}_0) - g_{1,1}(h_1(\mathbf{p}_0)) < E(f). \tag{15}$$

Indeed, if

$$\max_{\substack{\mathbf{y}\in Q \\ h_1(\mathbf{y})=h_1(\mathbf{p}_0)}} f_1(\mathbf{y}) = E(f) \quad \text{and} \quad \min_{\substack{\mathbf{y}\in Q \\ h_1(\mathbf{y})=h_1(\mathbf{p}_0)}} f_1(\mathbf{y}) = -E(f),$$

then

$$f_1(\mathbf{p}_0) - g_{1,1}(h_1(\mathbf{p}_0)) = f_1(\mathbf{p}_0) < E(f).$$

And if

$$\max_{\substack{\mathbf{y}\in Q \\ h_1(\mathbf{y})=h_1(\mathbf{p}_0)}} f_1(\mathbf{y}) = E(f) - \varepsilon_1 \quad \text{and} \quad \min_{\substack{\mathbf{y}\in Q \\ h_1(\mathbf{y})=h_1(\mathbf{p}_0)}} f_1(\mathbf{y}) = -E(f) + \varepsilon_2,$$

where $\varepsilon_1, \varepsilon_2 \ge 0$, $\varepsilon_1 + \varepsilon_2 \ne 0$, then it is not difficult to verify that

$$f_1(\mathbf{p}_0) - g_{1,1}(h_1(\mathbf{p}_0)) \le E(f) - \frac{\varepsilon_1 + \varepsilon_2}{2} < E(f).$$

In the same way we can prove that if $f_1(\mathbf{p}_0) - g_{1,1}(h_1(\mathbf{p}_0)) < E(f)$, then

$$f_1(\mathbf{p}_0) - g_{1,1}(h_1(\mathbf{p}_0)) - g_{2,1}(h_2(\mathbf{p}_0)) < E(f). \tag{16}$$

So, if $f_1(\mathbf{p}_0) < E(f)$, then $f_2(\mathbf{p}_0) < E(f)$. Repeating the same techniques from (15) to (16), it can be shown that if $f_1(\mathbf{p}_0) > -E(f)$, then $f_2(\mathbf{p}_0) > -E(f)$. Therefore, if $f_2(\mathbf{p}_0) = E(f)$, then $f_1(\mathbf{p}_0) = E(f)$, and if $f_2(\mathbf{p}_0) = -E(f)$, then $f_1(\mathbf{p}_0) = -E(f)$. This simply means that each path extremal for f_2 will be extremal for f_1.

Now we show that if any path extremal for f_1 has length at most N, then any path extremal for f_2 has length at most $N-1$. Suppose that, on the contrary, there is a path extremal for f_2 of length N, say $q = (\mathbf{q}_1, \ldots, \mathbf{q}_N)$. We may assume that $h_2(\mathbf{q}_{N-1}) = h_2(\mathbf{q}_N)$. As shown above, q is also extremal for f_1. Assume that $f_1(\mathbf{q}_N) = E(f)$. Then there is no $\mathbf{q}_0 \in Q$ such that $\mathbf{q}_0 \ne \mathbf{q}_N$, $h_1(\mathbf{q}_0) = h_1(\mathbf{q}_N)$ and $f_1(\mathbf{q}_0) = -E(f)$. Indeed, if there were such a $\mathbf{q}_0$ and $\mathbf{q}_0 \notin q$, then the path $(\mathbf{q}_1, \ldots, \mathbf{q}_N, \mathbf{q}_0)$ would be extremal for f_1. But this would contradict our assumption that any path extremal for f_1 has length at most N. Also, if there were such a $\mathbf{q}_0$ with $\mathbf{q}_0 \in q$, we could form some closed path extremal for f_1, contrary to our assumption.

Hence

$$\max_{\substack{\mathbf{y}\in Q\\ h_1(\mathbf{y})=h_1(\mathbf{q}_N)}} f_1(\mathbf{y}) = E(f), \qquad \min_{\substack{\mathbf{y}\in Q\\ h_1(\mathbf{y})=h_1(\mathbf{q}_N)}} f_1(\mathbf{y}) > -E(f).$$

Therefore,

$$|f_1(\mathbf{q}_N) - g_{11}(h_1(\mathbf{q}_N))| < E(f).$$

From the last inequality it is not difficult to deduce that

$$|f_2(\mathbf{q}_N)| < E(f).$$

This means that, contrary to our assumption, the path $(\mathbf{q}_1, \ldots, \mathbf{q}_N)$ cannot be extremal for f_2. Hence any path extremal for f_2 has length at most $N-1$.

In the same way, it can be shown that any path extremal for f_3 has length at most $N-2$, any path extremal for f_4 has length at most $N-3$ and so on. Finally, we conclude that there is no path extremal for f_{N+1}. In this case, for any $\mathbf{x} \in Q$,

$$|f_{N+1}(\mathbf{x})| < E(f). \tag{17}$$

Since f_{N+1} is continuous on Q, it follows from (17) that

$$\|f_{N+1}\| < E(f).$$

Since the function $f_{N+1} - f$ belongs to $C_{h_1h_2}$, the last strict inequality contradicts the definition of $E(f)$. Therefore, our assumption that there does not exist an infinite path extremal for f_1 is not valid.

Sufficiency. Let $l = (\mathbf{p}_1, \ldots, \mathbf{p}_{2n})$ be a closed path extremal for f_1. It can be easily verified that

$$|G_l(f)| = \|f - g_0\|. \tag{18}$$

By Lemma 2.2,

$$|G_l(f)| \le E(f). \tag{19}$$

It follows from (18), (19) and the definition of $E(f)$ that g_0 is a best approximation.

Let now $l = (\mathbf{p}_1, \mathbf{p}_2 \ldots,)$ be an infinite path extremal for f_1. Without loss of generality we may assume that the points $\mathbf{p}_i$ are all distinct (in the other case, we could form a closed path and prove in a few lines as above that g_0 is a best approximation). Consider the sequence $l_n = (\mathbf{p}_1, \ldots, \mathbf{p}_n)$, $n = 1, 2, \ldots,$ of finite paths and the path functionals

$$F_{l_n}(f) = \frac{1}{n} \sum_{i=1}^{n} (-1)^{i-1} f(\mathbf{p}_i).$$

Unlike G_l, these functionals do not annihilate the set $C_{h_1h_2}$. But it can be easily verified that $\|F_{l_n}\| = 1$ for all $n \in \mathbb{N}$. Indeed, $\|F_{l_n}(w)\| \le \|w\|$ for all continuous functions w over Q and $\|F_{l_n}(w_0)\| = \|w_0\|$ for a continuous function w_0 taking value $+1$ at the points $\mathbf{p}_i \in l_n$ with i odd, -1 at $\mathbf{p}_i \in l_n$ with i even, and values from the interval $(-1;1)$ at all other points of Q. By the well-known result of functional analysis (any bounded set in E^*, the dual of a separable Banach space E, is precompact in the weak* topology), the sequence $\{F_{l_n}\}_{n=1}^{\infty}$ has a weak* cluster point. Denote it by F. Note that for any $n \in \mathbb{N}$,

$$|F_{l_n}(g_1 + g_2)| \leq \frac{2}{n}\,(\|g_1\| + \|g_2\|),$$

where $g_1(\mathbf{x}) = g_1(h_1(\mathbf{x}))$ and $g_2(\mathbf{x}) = g_2(h_2(\mathbf{x}))$. Therefore, $F(g) = 0$ for all $g \in C_{h_1h_2}$. Moreover, it is clear that $\|F\| \leq 1$. From the last two properties of F it follows that

$$|F(f)| = |F(f - g)| \leq \|f - g\| \tag{20}$$

for all $g \in C_{h_1h_2}$. Taking inf over g on the right-hand side of (20), we obtain

$$|F(f)| \leq E(f). \tag{21}$$

Since the paths l_n are extremal for $f_1 = f - g_0$,

$$|F_{l_n}(f - g_0)| = \|f - g_0\|.$$

Hence

$$|F(f)| = |F(f - g_0)| = \|f - g_0\|. \tag{22}$$

Now by (21) and (22), we conclude that g_0 is a best approximation. ∎

It is well known that characterization theorems of this type are essential in approximation theory. Chebyshev was the first to prove a similar result for polynomial approximation. Khavinson [10] characterized extremal elements in a special case of the problem considered. His case allows the approximation of a continuous bivariate function $f(x, y)$ by functions of the type $\varphi(x) + \psi(y)$. It should be noted that the techniques used in the proof of Theorem 2.5 are completely different from those in [10].

Remark. The question of existence of a best approximation from the set $C_{h_1h_2}$ to a function f in $C(Q)$ (or, in other words, the proximinality of this set in the space of all continuous functions) is far from trivial. Some geometrical conditions on Q sufficient for the existence of a best approximation may be found in [6]. These conditions a priori require that the mapping $h = (h_1, h_2) : Q \to h_1(Q) \times h_2(Q)$ should separate points of Q. Necessary conditions for the proximinality of $C_{h_1h_2}$ can be easily obtained from the known general result of Marshall and O'Farrell [15] established for the sum of two algebras (see Proposition 4 in [15]). Unfortunately, there is not yet a complete answer (necessary and sufficient conditions on Q) to the above question even in the simplest case when $Q \subset \mathbb{R}^2$ and $h_i(x_1, x_2) = x_i$ for $i = 1, 2$ (see, for example, [7]).

Acknowledgements. I would like to thank the referee, whose insightful comments helped a lot to improve the content and presentation of the paper.

References

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

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

[3] E. J. Candés, *Ridgelets*: *estimating with ridge functions*, Ann. Statist. 31 (2003), 1561–1599.

[4] 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.

[5] N. Dyn, W. A. Light and E. W. Cheney, *Interpolation by piecewise-linear radial basis functions*, J. Approx. Theory 59 (1989), 202–223.

[6] 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*, Mat. Sb. 187 (1996), 3–14 (in Russian); English transl.: Sb. Math. 187 (1996), 623–634.

[7] —, —, —, *On the existence of a best uniform approximation of functions of two variables by sums of the type* $\varphi(x) + \psi(y)$, Sibirsk. Mat. Zh. 36 (1995), 819–827 (in Russian); English transl.: Siberian Math. J. 36 (1995), 707–713.

[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 [S. Ya. Khavinson], *A Chebyshev theorem for the approximation of a function of two variables by sums of the type* $\varphi(x) + \psi(y)$, Izv. Akad. Nauk SSSR Ser. Mat. 33 (1969), 650–666 (in Russian); English transl.: Math. USSR- Izv. 3 (1969), 617–632.

[11] I. G. Kazantsev, *Tomographic reconstruction from arbitrary directions using ridge functions*, Inverse Problems 14 (1998), 635–645.

[12] A. Kroo, *On approximation by ridge functions*, Constr. Approx. 13 (1997), 447–460.

[13] W. A. Light and E. W. Cheney, *On the approximation of a bivariate function by the sum of univariate functions*, J. Approx. Theory 29 (1980), 305–323.

[14] V. E. Maiorov, *On best approximation by ridge functions*, ibid. 99 (1999), 68–94.

[15] D. E. Marshall and A. G. O'Farrell, *Uniform approximation by real functions*, Fund. Math. 104 (1979), 203–211.

[16] K. I. Oskolkov, *Ridge approximation*, *Chebyshev–Fourier analysis and optimal quadrature formulas*, Proc. Steklov Inst. Math. 219 (1997), 265–280.

[17] P. P. Petrushev, *Approximation by ridge functions and neural networks*, SIAM J. Math. Anal. 30 (1998), 155–189.

[18] 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.

[19] V. Temlyakov, *On approximation by ridge functions*, preprint, Dept. of Math., Univ. of South Carolina, 1996, 12 pp.

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

Received May 14, 2006
Revised version August 6, 2007 (5922)