Proceedings of the Institute of Mathematics and Mechanics,
National Academy of Sciences of Azerbaijan
Volume 00, Number 0, XXXX, Pages 000–000

ON THE REPRESENTABILITY OF A SMOOTH FUNCTION BY SUMS OF GENERALIZED RIDGE FUNCTIONS

ASIM A. AKBAROV AND FIDAN M. ISGANDARLI

Abstract. In this paper, for the case $d = n-1$, we give a criterion under which a smooth function of n variables can be represented as a sum of generalized ridge functions of d variables and provide a partial solution to the smoothness problem in generalized ridge function representation.

1. Introduction

A multivariate function $F : \mathbb{R}^n \to \mathbb{R}$ of the form

$$F(\mathbf{x}) = f(\mathbf{a}^1 \cdot \mathbf{x}, ..., \mathbf{a}^d \cdot \mathbf{x})$$

is called a *generalized ridge function*, where $\mathbf{x} = (x_1, ..., x_n) \in \mathbb{R}^n$, $f : \mathbb{R}^d \to \mathbb{R}$ is a real-valued function of d variables $(1 \leq d < n)$ and $\mathbf{a}^j = (a_1^j, ..., a_n^j) \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, $j = 1, ..., d$ are fixed vectors (directions). For $d = 1$ a generalized ridge function is called a *ridge function*. Ridge functions and generalized ridge functions arise naturally in various fields. They arise in computerized tomography (see, e.g., [22, 23, 26, 29, 30]), statistics (see, e.g., [9, 13, 14, 18]), large-scale data analysis (see, e.g., [10, 12, 27, 34]) and neural networks (see, e.g., [19, 21, 28, 31]). These functions are also used in modern approximation theory as an effective tool for approximating complicated multivariate functions (see, e.g., [15, 16, 17, 24]). We refer the reader to the monographs of A.Pinkus [33] and V.Ismailov [20] for a detailed and systematic study of ridge functions.

One of the basic problems concerning the approximation by sums of ridge functions and generalized ridge functions is the problem of verifying the representability of a given multivariate function F as a sum of ridge functions and generalized ridge functions. Assume we are given a function $F : \mathbb{R}^n \to \mathbb{R}$, and fixed pairwise linearly independent directions $\mathbf{a}^{k,j} \in \mathbb{R}^n$, $k = 1, ..., m$, $j = 1, ..., d$. It is required to find a condition under which the function F can be represented as

$$F(\mathbf{x}) = \sum_{k=1}^{m} f_k(\mathbf{a}^{k,1} \cdot \mathbf{x}, ..., \mathbf{a}^{k,d} \cdot \mathbf{x}),$$

2010 *Mathematics Subject Classification.* 26B40; 39B22.
Key words and phrases. ridge function; generalized ridge function; Cauchy functional equation; function increment.

where $f_k : \mathbb{R}^d \to \mathbb{R}$, $k = 1, ..., m$ are arbitrarily behaved real-valued functions of d variables.

For ridge functions this problem was solved by P.Diaconis and M.Shahshahani [11].

Theorem 1.1 (P.Diaconis, M.Shahshahani [11]). *Let $\mathbf{a}^k \in \mathbb{R}^n$, $k = 1, ..., m$, be pairwise linearly independent vectors in $\mathbb{R}^n$. Denote by H^k, $k = 1, ..., m$, the hyperplane $\{\mathbf{c} \in \mathbb{R}^n : \mathbf{c}{\cdot}\mathbf{a}^k = 0\}$. Then a function $F \in C^m(\mathbb{R}^n)$ can be represented in the form*

$$F(\mathbf{x}) = \sum_{k=1}^{m} f_k(\mathbf{a}^k \cdot \mathbf{x}) + P(\mathbf{x}),$$

where $f_k \in C^m(\mathbb{R})$, $k = 1, ..., m$ and $P(\mathbf{x})$ is a polynomial of degree less than m, if and only if

$$\prod_{k=1}^{m} \sum_{s=1}^{n} c_s^k \frac{\partial F}{\partial x_s} = 0$$

for all vectors $\mathbf{c}^k = (c_1^k, ..., c_n^k) \in H^k$, $k = 1, ..., m$.

Remark 1.1. There are examples showing that one cannot simply dispense with the polynomial $P(\mathbf{x})$ in the above theorem.

In this paper, for the case $d = n - 1$, we give a criterion under which a smooth function of n variables can be represented as a sum of generalized ridge functions of d variables and provide a partial solution to the smoothness problem in generalized ridge function representation for this case.

2. On the representability of a smooth function by a sum of generalized ridge functions

Definition 2.1. Let $\{\mathbf{a}^1, ..., \mathbf{a}^d\}$ and $\{\mathbf{b}^1, ..., \mathbf{b}^d\}$ be linear independent vector systems in R^n $(1 \leq d < n)$. If

$$\mathbf{span}\{\mathbf{a^1}, ..., \mathbf{a^d}\} = \mathbf{span}\{\mathbf{b^1}, ..., \mathbf{b^d}\},$$

then the systems $\{\mathbf{a}^1, ..., \mathbf{a}^d\}$ and $\{\mathbf{b}^1, ..., \mathbf{b}^d\}$ are called equivalent, otherwise, if

$$\mathbf{span}\{\mathbf{a^1}, ..., \mathbf{a^d}\} \neq \mathbf{span}\{\mathbf{b^1}, ..., \mathbf{b^d}\},$$

then the systems $\{\mathbf{a}^1, ..., \mathbf{a}^d\}$ and $\{\mathbf{b}^1, ..., \mathbf{b}^d\}$ are called non-equivalent.

Remark 2.1. Obviously, if the systems $\{\mathbf{a}^1, ..., \mathbf{a}^d\}$ and $\{\mathbf{b}^1, ..., \mathbf{b}^d\}$ are equivalent, then any generalized ridge function of the form $F(\mathbf{x}) = f(\mathbf{a}^1 \cdot \mathbf{x}, ..., \mathbf{a}^d \cdot \mathbf{x})$ also has the form $F(\mathbf{x}) = g(\mathbf{b}^1 \cdot \mathbf{x}, ..., \mathbf{b}^d \cdot \mathbf{x})$. Therefore, when defining a generalized ridge function, without loss of generality, we can assume that the vectors $\mathbf{a}^1, ..., \mathbf{a}^d$ are unit and mutually perpendicular.

Let's consider the following problem: assume we are given a function $F : \mathbb{R}^n \to \mathbb{R}$, and fixed pairwise non-equivalent vector systems $\{\mathbf{a}^{1,1}, ..., \mathbf{a}^{1,d}\}$,...,$\{\mathbf{a}^{m,1}, ..., \mathbf{a}^{m,d}\}$

in $\mathbb{R}^n$. It is required to find a condition under which the function F can be represented as

$$F(\mathbf{x}) = \sum_{k=1}^{m} f_k(\mathbf{a}^{k,1} \cdot \mathbf{x}, ..., \mathbf{a}^{k,d} \cdot \mathbf{x}),$$

where $f_k : \mathbb{R}^d \to \mathbb{R}$, $k = 1, ..., m$ are arbitrarily behaved real-valued functions of d variables.

In this section we give a solution of this problem in case $d = n - 1$.

Theorem 2.1. *Let* $\{\mathbf{a}^{1,1}, ..., \mathbf{a}^{1,n-1}\}$,...,$\{\mathbf{a}^{m,1}, ..., \mathbf{a}^{m,n-1}\}$ *pairwise non-equivalent vector systems in* $\mathbb{R}^n$*. Then the function* $F \in C^m(\mathbb{R}^n)$ *can be represented in the form*

$$F(\boldsymbol{x}) = \sum_{k=1}^{m} f_k(\mathbf{a}^{k,1} \cdot \mathbf{x}, ..., \mathbf{a}^{k,n-1} \cdot \mathbf{x}) \tag{2.1}$$

if and only if

$$\frac{\partial^m F}{\partial l_1 ... \partial l_m}(\mathbf{x}) = 0 \tag{2.2}$$

for any $\mathbf{x} \in \mathbb{R}^n$*, where* $l_k \in \mathbb{R}^n$ *is a unit vector, perpendicular to the vectors* $\mathbf{a}^{k,1}, ..., \mathbf{a}^{k,n-1}$*,* $k = 1, ..., m$*.*

At first, we prove the auxiliary lemma.

Lemma 2.1. *Let* $\mathbf{a}^1, ..., \mathbf{a}^{n-1}$ *be any linearly independent vectors in* $\mathbb{R}^n$ *and the vector* $l \in \mathbb{R}^n$ *is not perpendicular to the vector space* $\mathbf{span}\{\mathbf{a^1}, ..., \mathbf{a^{n-1}}\}$*. Then for any function* $\phi(\mathbf{u}) = \phi(u_1, ..., u_{n-1}) \in C^1(\mathbb{R}^{n-1})$ *there exist a continuously differentiable generalized ridge function of the form* $\Phi(\mathbf{x}) = f(\mathbf{a}^1 \cdot \mathbf{x}, ..., \mathbf{a}^{n-1} \cdot \mathbf{x})$ *such that*

$$\frac{\partial \Phi}{\partial l}(\mathbf{x}) = \phi(\mathbf{a}^1 \cdot \mathbf{x}, ..., \mathbf{a}^{n-1} \cdot \mathbf{x}) \tag{2.3}$$

for any $\mathbf{x} \in \mathbb{R}^n$*.*

Proof of Lemma 2.1. It follows from Remark 2.1 that without loss of generality, we can assume that the vectors $\mathbf{a}^1, ..., \mathbf{a}^{n-1}$ are unit and mutually perpendicular. Denote by $\mathbf{a}^0$ the unit vector, perpendicular to the vectors $\mathbf{a}^1, ..., \mathbf{a}^{n-1}$. Let

$$l = \sum_{p=0}^{n-1} \alpha_p \cdot \mathbf{a}^p.$$

As the vector $l \in \mathbb{R}^n$ is not perpendicular to the vector space $\mathbf{span}\{\mathbf{a^1}, ..., \mathbf{a^{n-1}}\}$, then

$$\eta_0 = \sum_{p=1}^{n-1} \alpha_p^2 > 0.$$

Denote

$$\Phi(\mathbf{x}) = \frac{1}{\eta_0} \int_0^{\sum_{p=1}^{n-1} \alpha_p \mathbf{a}^p \cdot \mathbf{x}} \phi(s_1(t, \mathbf{x}), ..., s_{n-1}(t, \mathbf{x})) dt,$$

where

$$s_k(t, \mathbf{x}) = \frac{1}{\eta_0} \left(\alpha_k t + \sum_{p=1,\ p \neq k}^{n-1} (\alpha_p^2 \cdot \mathbf{a^k} - \alpha_p \alpha_k \cdot \mathbf{a^p}) \cdot \mathbf{x} \right), \quad k = 1, ..., n-1.$$

It follows from equations

$$\frac{\partial\Phi}{\partial\mathbf{a^k}}(\mathbf{x})=\frac{\alpha_k}{\eta_0}\cdot\phi(\mathbf{a}^1\cdot\mathbf{x},...,\mathbf{a}^{n-1}\cdot\mathbf{x})$$

$$+\frac{1}{\eta_0^2}\int_0^{\sum_{p=1}^{n-1}\alpha_p\mathbf{a}^p\cdot\mathbf{x}}\left(\sum_{j=1,\ j\neq k}^{n-1}\frac{\partial\phi}{\partial u_j}(s_1(t,\mathbf{x}),...,s_{n-1}(t,\mathbf{x}))\cdot(-\alpha_k\alpha_j)\right)dt$$

$$+\frac{1}{\eta_0^2}\int_0^{\sum_{p=1}^{n-1}\alpha_p\mathbf{a}^p\cdot\mathbf{x}}\left(\frac{\partial\phi}{\partial u_k}(s_1(t,\mathbf{x}),...,s_{n-1}(t,\mathbf{x}))\cdot\sum_{p=1,\ p\neq k}^{n-1}\alpha_p^2\right)dt,\ k=1,...,n-1,$$

$$\frac{\partial\Phi}{\partial\mathbf{a^0}}(\mathbf{x})=0,$$

that

$$\frac{\partial\Phi}{\partial l}(\mathbf{x})=\sum_{k=0}^{n-1}\alpha_k\frac{\partial\Phi}{\partial\mathbf{a^k}}(\mathbf{x})=\phi(\mathbf{a}^1\cdot\mathbf{x},...,\mathbf{a}^{n-1}\cdot\mathbf{x}).$$

On the other side, it follows from $\frac{\partial\Phi}{\partial\mathbf{a^0}}(\mathbf{x})=0$ that the function Φ is of the form $\Phi(\mathbf{x})=f(\mathbf{a}^1\cdot\mathbf{x},...,\mathbf{a}^{n-1}\cdot\mathbf{x})$. This completes the proof of the lemma.

Proof of Theorem 2.1. Necessity. Let the function $F\in C^m(\mathbb{R}^n)$ be of the form (2.1). For any $\mathbf{x}\in\mathbb{R}^n$ and $\mathbf{h}\in\mathbb{R}^n$ we denote by $\Delta_{\mathbf{h}}F(\mathbf{x})$ the increment

$$\Delta_{\mathbf{h}}F(\mathbf{x})=F(\mathbf{x}+\mathbf{h})-F(\mathbf{x})$$

of the function F. Then it follows from (2.1) that for any $\mathbf{x}\in\mathbb{R}^n$ and for any $t_1,...,t_m\in\mathbb{R}$

$$\Delta_{t_1l_1}...\Delta_{t_ml_m}F(\mathbf{x})=0,\tag{2.4}$$

where l_k is a unit vector, perpendicular to the vectors $\mathbf{a}^{k,1},...,\mathbf{a}^{k,n-1}$, $k=1,...,m$. It follows from (2.4) that for any $\mathbf{x}\in\mathbb{R}^n$

$$\frac{\partial^mF}{\partial l_1...\partial l_m}(\mathbf{x})=\lim_{t_1\to0+,...,t_m\to0+}\frac{\Delta_{t_1l_1}...\Delta_{t_ml_m}F(\mathbf{x})}{t_1\cdot...\cdot t_m}=0.$$

Sufficiency. Let the function $F\in C^m(\mathbb{R}^n)$ satisfy condition (2.2) for any $\mathbf{x}\in\mathbb{R}^n$. Let us write equation (2.2) in the form

$$\frac{\partial}{\partial l_1}\left[\frac{\partial^{m-1}F}{\partial l_2...\partial l_m}\right](\mathbf{x})=0.\tag{2.5}$$

It follows from (2.5) that the partial derivative $\frac{\partial^{m-1}F}{\partial l_2...\partial l_m}$ of the function F is independent of the direction l_1. Therefore there exists a function $\phi_1:\mathbb{R}^{n-1}\to\mathbb{R}$ such that

$$\frac{\partial^{m-1}F}{\partial l_2...\partial l_m}(\mathbf{x})=\phi_1(\mathbf{a}^{1,1}\cdot\mathbf{x},...,\mathbf{a}^{1,n-1}\cdot\mathbf{x}),\quad\mathbf{x}\in\mathbb{R}^n.\tag{2.6}$$

From condition $F\in C^m(\mathbb{R}^n)$ we obtain that $\phi_1\in C^1(\mathbb{R}^{n-1})$. Now let us write equation (2.6) in the form

$$\frac{\partial}{\partial l_2}\left[\frac{\partial^{m-2}F}{\partial l_3...\partial l_m}\right](\mathbf{x})=\phi_1(\mathbf{a}^{1,1}\cdot\mathbf{x},...,\mathbf{a}^{1,n-1}\cdot\mathbf{x}),\quad\mathbf{x}\in\mathbb{R}^n.\tag{2.7}$$

It follows from Lemma 2.1 that there exists a continuously differentiable generalized ridge function of the form

$$\Phi_1(\mathbf{x})=g_1(\mathbf{a}^{1,1}\cdot\mathbf{x},...,\mathbf{a}^{1,n-1}\cdot\mathbf{x}),\quad\mathbf{x}\in\mathbb{R}^n\tag{2.8}$$

such that

$$\frac{\partial \Phi_1}{\partial l_2}(\mathbf{x}) = \phi_1(\mathbf{a}^{1,1} \cdot \mathbf{x}, ..., \mathbf{a}^{1,n-1} \cdot \mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^n. \tag{2.9}$$

It follows from (2.7) and (2.9) that for any $\mathbf{x} \in \mathbb{R}^n$

$$\frac{\partial}{\partial l_2}\left[\frac{\partial^{m-2} F}{\partial l_3 ... \partial l_m} - \Phi_1\right](\mathbf{x}) = 0.$$

Then the function $\frac{\partial^{m-2} F}{\partial l_3 ... \partial l_m} - \Phi_1$ is independent of the direction l_2. Therefore there exist a function $\phi_2 : \mathbb{R}^{n-1} \to \mathbb{R}$ such that

$$\frac{\partial^{m-2} F}{\partial l_3 ... \partial l_m}(\mathbf{x}) - \Phi_1(\mathbf{x}) = \phi_2(\mathbf{a}^{2,1} \cdot \mathbf{x}, ..., \mathbf{a}^{2,n-1} \cdot \mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^n. \tag{2.10}$$

Since the functions $\frac{\partial^{m-2} F}{\partial l_3 ... \partial l_m}$ and Φ_1 are continuously differentiable, then we get that the function ϕ_2 is also continuously differentiable in $\mathbb{R}^{n-1}$. It follows from (2.8) and (2.10) that

$$\frac{\partial^{m-2} F}{\partial l_3 ... \partial l_m}(\mathbf{x}) = g_1(\mathbf{a}^{1,1} \cdot \mathbf{x}, ..., \mathbf{a}^{1,n-1} \cdot \mathbf{x}) + \phi_2(\mathbf{a}^{2,1} \cdot \mathbf{x}, ..., \mathbf{a}^{2,n-1} \cdot \mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^n.$$

Continuing the above process, until it reaches the function F, we obtain the desired result. This completes the proof of the theorem.

3. The smoothness problem in generalized ridge function representation

Another problem in the ridge function representation is the smoothness problem. Assume we are given a function $F : \mathbb{R}^n \to \mathbb{R}$ of the form

$$F(\mathbf{x}) = \sum_{k=1}^{m} f_k(\mathbf{a}^{k,1} \cdot \mathbf{x}, ..., \mathbf{a}^{k,d} \cdot \mathbf{x}), \tag{3.1}$$

where $\{\mathbf{a}^{1,1}, ..., \mathbf{a}^{1,d}\}$,...,$\{\mathbf{a}^{m,1}, ..., \mathbf{a}^{m,d}\}$ pairwise non-equivalent vector systems in $\mathbb{R}^n$, $1 \leq d < n-1$, $f_1, ..., f_m$ are arbitrarily behaved real-valued functions of d variables. Assume, in addition, that F is of a certain smoothness class, that is, $F \in C^s(\mathbb{R}^n)$, where $s \geq 0$ (with the convention that $C^0(\mathbb{R}^n) = C(\mathbb{R}^n)$). What can we say about the smoothness of the functions f_k? Do the functions f_k necessarily inherit all the smoothness properties of the F?

If $d = 1$ and $m = 1$ or $m = 2$ the answer to the above question is yes (see [8]). If $d = 1$ and $m \geq 3$ the picture drastically changes. For $d = 1$, $m = 3$, there are smooth functions which decompose into sums of very badly behaved ridge functions. This phenomena comes from the classical Cauchy Functional Equation. This equations,

$$h(x + y) = h(x) + h(y), \quad h : \mathbb{R} \to \mathbb{R},$$

looks very simple and has a class of simple solutions $h(x) = cx$, $c \in \mathbb{R}$. However, it easily follows from Hamel basis theory that Cauchy Functional Equation also has a large class of "wild" solutions. These solutions are called "wild" because they are extremely pathological. For example, they are not continuous at a point, not monotone on an interval, not bounded on any set of positive measure (see [1]).

Let h_1 be any "wild" solution of the Cauchy Functional Equation. Then the zero function can be represented as

$$0 = h_1(x) + h_1(y) - h_1(x + y). \tag{3.2}$$

Note that the functions involved in (3.2) are bivariate ridge functions with the directions $\mathbf{a}^1 = (1, 0)$, $\mathbf{a}^2 = (0, 1)$ and $\mathbf{a}^3 = (1, 1)$, respectively. This example shows that for smoothness of the representation (3.1) one must impose additional conditions on the representing functions f_k, $k = 1, ..., m$.

In case $d = 1$ it was first proved by M.Buhmann and A.Pinkus [8] that if in (3.1) $F \in C^s(\mathbb{R}^n)$, $s \geq m-1$ and $f_k \in L^1_{loc}(\mathbb{R})$ for each $k = 1, ..., m$, then $f_k \in C^s(\mathbb{R}^n)$, $k = 1, ..., m$. Later, A.Pinkus [32] extensively generalized this result. He solved this problem for any $s \in Z_+$, while imposing weaker conditions on the functions f_k.

In case $d \geq 2$ the situation is slightly more problematic. Consider, for example, the case $d = 2$, $n = 3$, $m = 2$, $\mathbf{a}^{1,1} = (1, 0, 0)$, $\mathbf{a}^{1,2} = (0, 1, 0)$, $\mathbf{a}^{2,1} = (0, 1, 0)$, $\mathbf{a}^{2,2} = (0, 0, 1)$. Thus

$$F(x_1, x_2, x_3) = f_1(x_1, x_2) + f_2(x_2, x_3).$$

Setting $f_1(x_1, x_2) = g(x_2)$ and $f_2(x_2, x_3) = -g(x_2)$ for any arbitrary univariate function g, we have

$$0 = f_1(x_1, x_2) + f_2(x_2, x_3),$$

and yet f_1 and f_2 do not exhibit any of the smoothness properties of the left-hand side of this equation.

Now consider the following natural and interesting question. Assume we are given a function $F \in C^s(\mathbb{R}^n)$ of the form (3.1). Is it true that there will always exist $g_k \in C^s(\mathbb{R}^d)$, $k = 1, ..., m$ such that

$$F(\mathbf{x}) = \sum_{k=1}^{m} g_k(\mathbf{a}^{k,1} \cdot \mathbf{x}, ..., \mathbf{a}^{k,d} \cdot \mathbf{x})?$$

This question was posed in M.Buhmann and A.Pinkus [8] for ridge function representation and Pinkus [33] for generalized ridge function representation. In [2, 3, 4, 6, 25], the authors gave a partial solution to the above representation problem for ridge function representation. In [7], this problem for ridge function representation was solved up to a multivariate polynomial:

Theorem 3.1 (R.Aliev, V.Ismailov [7])**.** *Assume a function $F \in C(\mathbb{R}^n)$ is of the form*

$$F(\boldsymbol{x}) = \sum_{k=1}^{m} f_k(\mathbf{a}^k \cdot \mathbf{x}), \tag{3.3}$$

where $\mathbf{a}^1$,..., $\mathbf{a}^m$ are given pairwise linearly independent directions in $\mathbb{R}^n$, $f_1, ..., f_m$ are arbitrarily behaved univariate functions. Then there exist continuous functions $g_k : \mathbb{R} \to \mathbb{R}$, $k = 1, ..., m$, and a polynomial P_{m-1} of degree at most $m - 1$ such that

$$F(\mathbf{x}) = \sum_{k=1}^{m} g_k(\mathbf{a}^k \cdot \mathbf{x}) + P_{m-1}(\mathbf{x}). \tag{3.4}$$

Corollary 3.1 (R.Aliev, V.Ismailov [7]). *Assume a function $F \in C^s(\mathbb{R}^n)$, $s \in N$ is of the form* (3.3). *Then there exist functions $g_k \in C^s(\mathbb{R})$, $k = 1, ..., m$, and a polynomial P_{m-1} of degree at most $m-1$ such that* (3.4) *holds.*

Corollary 3.2 (R.Aliev, V.Ismailov [7]). *Assume a function $F \in C^s(\mathbb{R}^2)$, $s \in Z_+ = N \cup \{0\}$ is of the form* (3.3). *Then there exist functions $g_k \in C^s(\mathbb{R})$, $k = 1, ..., m$, such that*

$$F(\mathbf{x}) = \sum_{k=1}^{m} g_k(\mathbf{a}^k \cdot \mathbf{x}).$$

In [5] a new proof of Theorem 3.1 is given. In this section, we give a partial solution to the posed problem for generalized ridge function representation.

Theorem 3.2. *Assume a function $F \in C^m(\mathbb{R}^n)$ is of the form*

$$F(\boldsymbol{x}) = \sum_{k=1}^{m} f_k(\mathbf{a}^{k,1} \cdot \mathbf{x}, ..., \mathbf{a}^{k,n-1} \cdot \mathbf{x}), \tag{3.5}$$

where $\{\mathbf{a}^{1,1}, ..., \mathbf{a}^{1,n-1}\}$,...,$\{\mathbf{a}^{m,1}, ..., \mathbf{a}^{m,n-1}\}$ pairwise non-equivalent vector systems in $\mathbb{R}^n$, $f_1, ..., f_m$ are arbitrarily behaved real-valued functions of $n-1$ variables. Then there exist functions $g_k \in C^1(\mathbb{R}^{n-1})$, $k = 1, ..., m$, such that

$$F(\boldsymbol{x}) = \sum_{k=1}^{m} g_k(\mathbf{a}^{k,1} \cdot \mathbf{x}, ..., \mathbf{a}^{k,n-1} \cdot \mathbf{x}). \tag{3.6}$$

Proof. Let the function $F \in C^m(\mathbb{R}^n)$ is of the form (3.5). Then it follows from Theorem 2.1 that

$$\frac{\partial^m F}{\partial l_1 ... \partial l_m}(\mathbf{x}) = 0$$

for any $\mathbf{x} \in \mathbb{R}^n$, where $l_k \in \mathbb{R}^n$ is a unit vector, perpendicular to the vectors $\mathbf{a}^{k,1}, ..., \mathbf{a}^{k,n-1}$, $k = 1, ..., m$. Then from the proof of the sufficiency of Theorem 2.1 it is clear that there exist continuously differentiable functions $g_k : \mathbb{R}^{n-1} \to \mathbb{R}$, $k = 1, ..., m$, such that (3.6) is satisfied. This completes the proof of the theorem.

References

[1] J. Aczel, *Lectures on Functional Equations and their Applications*, Academic Press, New York, 1966, 509 pp.

[2] R. A. Aliev, A. A. Asgarova, V. E. Ismailov, A note on continuous sums of ridge functions. *J. Approx. Theory* **237** (2019), 210-221.

[3] R. A. Aliev, A. A. Asgarova, V. E. Ismailov, On the Hölder continuity in ridge function representation. *Proc. Inst. Math. Mech. Natl. Acad. Sci. Azerb.* **45** (2019), no. 1, 31-40.

[4] R. A. Aliev, A. A. Asgarova, V. E. Ismailov, On the representation by bivariate ridge functions. *Ukrainian Math. J.* **73** (2021), no. 5, 675-685.

[5] R. A. Aliev, F. M. Isgandarli, On the representability of a continuous multivariate function by sums of ridge functions. *J. Approx. Theory* **304** (2024), Art. no. 106105.

[6] R. A. Aliev, V. E. Ismailov, On a smoothness problem in ridge function representation. *Adv. Appl. Math.* **73** (2016), 154-169.

[7] R. A. Aliev, V. E. Ismailov, A representation problem for smooth sums of ridge functions. *J. Approx. Theory* **257** (2020), Art. no. 105448.

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

[9] E. J. Candes, Ridgelets: estimating with ridge functions. *Ann. Statist.* **31** (2003), no. 5, 1561-1599.

[10] A. Cohen, I. Daubechies, R. DeVore, G. Kerkyacharian, D. Picard, Capturing ridge functions in high dimensions from point queries. *Constr. Approx.* **35** (2012), no. 2, 225-243.

[11] P. Diaconis, M. Shahshahani, On nonlinear functions of linear combinations. *SIAM J. Sci. Stat. Comput.* **5** (1984), no. 1, 175-191.

[12] B. Doerr, S. Mayer, The recovery of ridge functions on the hypercube suffers from the curse of dimensionality. *J. Complexity* **63** (2021), Art. no. 101521.

[13] D. L. Donoho, I. M. Johnstone, Projection-based approximation and a duality method with kernel methods. *Ann. Statist.* **17** (1989), no. 1, 58-106.

[14] J. H. Friedman, W. Stuetzle, Projection pursuit regression. *J. Amer. Statist. Assoc.* **76** (1981), no. 376, 817-823.

[15] Y. Gordon, V. Maiorov, M. Meyer, S. Reisner, On the best approximation by ridge functions in the uniform norm.*Constr. Approx.* **18** (2002), no. 1, 61-85.

[16] C. Greif, P. Junk, K. Urban, Linear/Ridge expansions: enhancing linear approximations by ridge functions. *Adv. Comput. Math.* **48** (2022), Art. no. 15.

[17] J. C. Gross, G. T. Parks, Optimization by moving ridge functions: derivative-free optimization for computationally intensive functions. *Engineering Optimization* **54** (2022), no. 4, 553-575.

[18] P. J. Huber, Projection pursuit. *Ann. Statist.* **13** (1985), no. 2, 435-475.

[19] V. E. Ismailov, Computing the approximation error for neural networks with weights varying on fixed directions. *Numer. Funct. Anal. Optim.* **40** (2019), no. 12, 1395-1409.

[20] V. E. Ismailov, *Ridge Functions and Applications in Neural Networks*, Mathematical Surveys and Monographs, 263. American Mathematical Society, 2021, 186 pp.

[21] V. E. Ismailov, E. Savas, Measure theoretic results for approximation by neural networks with limited weights. *Numer. Funct. Anal. Optim.* **38** (2017), no. 7, 819-830.

[22] I. Kazantsev, Tomographic reconstruction from arbitrary directions using ridge functions. *Inverse Problems* **14** (1998), no. 3, 635-645.

[23] I. Kazantsev, I. Lemahieu, Reconstruction of elongated structures using ridge functions and natural pixels. *Inverse Problems* **16**(2000), no.2, 505-517.

[24] I. G. Kazantsev, R. Z. Turebekov, M. A. Sultanov, Inpainting of regular textures using ridge functions. *J. Inverse and Ill-posed Problems* **30** (2022), no. 5, 759-766.

[25] A. A. Kuleshov, On some properties of smooth sums of ridge functions. *Proc. Steklov Inst. Math.* **294** (2016), 89–94.

[26] B. F. Logan, L. A. Shepp, Optimal reconstruction of a function from its projections. *Duke Math. J.* **42** (1975), no. 4, 645-659.

[27] Y. Malykhin, K. Ryutin, T. Zaitseva, Recovery of regular ridge functions on the ball. *Constr. Approx.* **56** (2022), 687–708.

[28] V. E. Maiorov, R. Meir, On the near optimality of the stochastic approximation of smooth functions by neural networks. *Adv. Comput. Math.* **13** (2000), no. 1, 79-103.

[29] R. B. Marr,On the reconstruction of a function on a circular domain from a sampling of its line integrals. *J. Math. Anal. Appl.* **45** (1974), no. 2, 357-374.

[30] F. Natterer, *The Mathematics of Computerized Tomography*, SIAM, Classics in Applied Mathematics, Series No. 32, 2001, 222 pp.

[31] A. Pinkus, Approximation theory of the MLP model in neural networks. *Acta Numerica* **8** (1999), 143-195.

[32] A. Pinkus, Smoothness and uniqueness in ridge function representation. *Indag. Math.* **24** (2013), no. 4, 725-738.

[33] A. Pinkus, *Ridge functions*, Cambridge Tracts in Mathematics 205. Cambridge University Press, 2015, 207 pp.

[34] H. Tyagi, V. Cevher, Learning nonparametric basis independent models from point queries via low-rank methods. *Appl. Comput. Harmonic Anal.* **37** (2014), no. 3, 389-412.

Asim A. Akbarov
Baku State University, Z. Khalilov str. 23, AZ1148, Baku, Azerbaijan
E-mail address: akbarovasim1980@gmail.com

Fidan M. Isgandarli
Baku State University, Z. Khalilov str. 23, AZ1148, Baku, Azerbaijan
E-mail address: fidanisgandarli100@gmail.com

Received: November 12, 2024; Revised: February 5, 2025; Accepted: February 22, 2025