



A note on the representation of continuous functions by linear superpositions[☆]

Vugar E. Ismailov

Mathematics and Mechanics Institute, Azerbaijan National Academy of Sciences, Az-1141, Baku, Azerbaijan

ARTICLE INFO

Article history:

Received 12 March 2011

Received in revised form

17 April 2011

2000 Mathematics Subject Classification:

primary 26B40

41A45

41A63

Keywords:

Uniform separation of measures

Bolt of lightning

Path

ABSTRACT

We consider the problem of the representation of real continuous functions by linear superpositions $\sum_{i=1}^k g_i \circ p_i$ with continuous g_i and p_i . This problem has been considered by many authors. But complete and at the same time explicit and practical solutions to the problem were given only for the case $k = 2$. For $k > 2$, a fairly practical sufficient condition for the representation can be found in Sternfeld (1978) [17] and Sproston and Strauss (1992) [16]. In this short note, we give a necessary condition of such a kind for the representability of continuous functions.

© 2011 Elsevier GmbH. All rights reserved.

1. Introduction

Let X be a set and p_i , $i = 1, \dots, k$, be arbitrarily fixed functions over X . For a given set Y , let $T(Y)$, $B(Y)$ and $C(Y)$ stand for the space of all, bounded, and continuous real functions on Y , respectively. Consider the following three sets:

$$S(X) = S(p_1, \dots, p_k; X) = \left\{ \sum_{i=1}^k g_i(p_i(x)) : x \in X, g_i \in T(\mathbb{R}), i = 1, \dots, k \right\},$$

$$S_b(X) = S_b(p_1, \dots, p_k; X) = \left\{ \sum_{i=1}^k g_i(p_i(x)) : x \in X, g_i \in B(\mathbb{R}), i = 1, \dots, k \right\},$$

[☆] This research was supported by the Science Development Foundation under the President of the Republic of Azerbaijan – Grant EIF-2010-1(1)-40/07-1.

E-mail address: vugaris@mail.ru.

$$S_c(X) = S_c(p_1, \dots, p_k; X) = \left\{ \sum_{i=1}^k g_i(p_i(x)) : x \in X, g_i \in C(\mathbb{R}), i = 1, \dots, k \right\}.$$

For the second set, the functions p_i , $i = 1, \dots, k$, are considered to be bounded on X . For the third set, we assume that X is a compact Hausdorff space and that the functions p_i are continuous on X . Members of these sets will be called linear superpositions (see [20]).

At present, there are many works investigating possibilities of the equalities $S(X) = T(X)$, $S_b(X) = B(X)$, $S_c(X) = C(X)$ (see [9] and references therein). Here, we are interested in the last equality $S_c(X) = C(X)$.

The famous Kolmogorov superposition theorem states that for X being the unit cube in $\mathbb{R}^d$ there exist functions $p_i \in C(X)$, $i = 1, \dots, 2d+1$, such that $S_c(p_1, \dots, p_{2d+1}; X) = C(X)$ (see [11]). Further, the functions p_i can be chosen as sums of univariate functions. This deep result, which solved Hilbert's 13th problem, was generalized in many directions. One such direction was in choosing various sets X of $\mathbb{R}^d$, or even general metric spaces (see, e.g., [3,4,9,14]). In all of these works, the functions $p_1, \dots, p_k$ guaranteeing the equality $S_c(p_1, \dots, p_k; X) = C(X)$ were incalculable. Regarding the nature of these functions, for some sets X , they can be chosen to be at most from the class $Lip1$ (see [5]).

Appearing in the late 60's, the work of Vitushkin and Henkin [20] showed that for $p_1, \dots, p_k$ (k may be very large) having besides continuity also smoothness properties, even the density of $S_c(p_1, \dots, p_k; X)$ in $C(X)$ does not generally hold. Thus, the question about when $S_c(X) = C(X)$ and $\bar{S}_c(X) = C(X)$ was raised. Clearly, any answer depends on both the behavior of $p_1, \dots, p_k$ and the structure of X .

For the above problem of representation, the first crucial step was made by Sternfeld [17]. He showed that the problem with its nature is dual to the problem of uniform separation of measures of some certain class (see [17,19]). The duality approach enabled him to prove that the number of terms in the Kolmogorov superposition formula cannot be reduced (see [18]). Let S be a class of measures defined on some field of subsets of X and $F = \{p\}$ be a family of functions defined on X . F uniformly separates measures of the class S if there exists a number $0 < \lambda \leq 1$ such that for each μ in S the equality $\|\mu \circ p^{-1}\| \geq \lambda \|\mu\|$ holds for some $p \in F$. In this terminology, $S_b(p_1, \dots, p_k; X) = B(X)$ and $S_c(p_1, \dots, p_k; X) = C(X)$ if and only if the family $\{p_1, \dots, p_k\}$ uniformly separates measures of the classes $I_1(X)$ and $C(X)^*$ correspondingly (see [19]). Since $I_1(X) \subset C(X)^*$ (the set of all regular Borel measures includes, in particular, discrete measures), Sternfeld concluded that the equality $S_c(X) = C(X)$ implies $S_b(X) = B(X)$. In [7], we showed that each of these two equalities implies $S(X) = T(X)$. That is, if some representation by linear superpositions holds for continuous (or bounded) functions, then it holds for all functions.

Sproston and Strauss [16] gave a practically convenient sufficient condition for the space $S_c(X)$ to be the whole of $C(X)$ (in fact, their result was equivalently formulated in terms of sums of closed algebras). To describe the condition, define the set functions

$$\tau_i(Z) = \left\{ x \in Z : |p_i^{-1}(p_i(x)) \cap Z| \geq 2 \right\}, \quad Z \subset X, i = 1, \dots, k,$$

where $|Y|$ denotes the cardinality of a set Y considered. Define $\tau(Z)$ to be $\bigcap_{i=1}^k \tau_i(Z)$ and define $\tau^2(Z) = \tau(\tau(Z))$, $\tau^3(Z) = \tau(\tau^2(Z))$ and so on inductively. The result of [16] says that $S_c(p_1, \dots, p_k; X) = C(X)$ provided that $\tau^n(X) = \emptyset$ for some positive integer n . In fact, this condition first appeared in the work of Sternfeld [17], where the author proved that $\tau^n(X) = \emptyset$ (for some n) guarantees that the family $\{p_1, \dots, p_k\}$ uniformly separates measures of the class $I_1(X)$ and also regular Borel measures if X is a compact metric space. Sproston and Strauss proved the last statement for X being a compact Hausdorff space. For $k = 2$, the condition is also necessary for the representation, but not in general if $k > 2$ (see the counterexample in [16]).

For $k = 2$, the above condition $\tau^n(X) = \emptyset$ can be expressed in terms of sets of points in X which were introduced in the literature under different names such as “bolts of lightning” [8,9,12], “trips” [13], “paths” [6], “loops” [2], etc. These objects are geometrically explicit. A path with respect to two functions p_1 and p_2 can be described as a trace of some point in X traveling (more precisely, jumping) in alternating level sets of the functions p_1 and p_2 . If the point returns to its primary position after such traveling, the set obtained is called a closed path. It is not difficult to prove that $\tau^n(X) = \emptyset$ if

and only if there are no closed paths in X and the lengths (number of points) of all paths are uniformly bounded (see [9]).

Paths with respect to two coordinate functions (and two algebras) have been extensively implemented by Marshall and O'Farrell [13,12] to solve the problem concerning the density of $S_c(p_1, p_2; X)$ in $C(X)$. Their work [12] showed the essence of such paths by explaining that every regular Borel measure orthogonal to $S_c(p_1, p_2; X)$ in $C(X)$ is in the closure of the set of measures generated by paths.

It should be remarked that many authors considering the problems of representation and approximation by linear superpositions indicated the difficulties and at the same time usefulness of going from measure-theoretic to path-descriptive results (see, e.g., [9,12,19]).

The purpose of this note is to obtain a path-descriptive necessary condition for representability of each continuous function by linear superpositions. We hope that our condition will complement the above sufficient condition $\tau^n(X) = \emptyset$ in some sense.

2. The main result

We begin this section with a definition. Let us first assume that we are given a compact Hausdorff space X and continuous functions $p_i : X \rightarrow \mathbb{R}$, $i = 1, \dots, k$.

Definition 2.1 (See [1,7,10]). A set of points $l = (x_1, \dots, x_n) \subset X$ is called a closed path with respect to the functions $p_1, \dots, p_k$ if there exists a vector $\lambda = (\lambda_1, \dots, \lambda_n) \in \mathbb{Z}^n \setminus \{0\}$ such that

$$\sum_{j=1}^n \lambda_j \delta_{p_i(x_j)} = 0, \quad \text{for all } i = 1, \dots, k.$$

Here δ_a is the characteristic function of the set $\{a\}$.

For example, the set $l = \{(0, 0, 0), (0, 0, 1), (0, 1, 0), (1, 0, 0), (1, 1, 1)\}$ is a closed path in $\mathbb{R}^3$ with respect to the functions $p_i(z_1, z_2, z_3) = z_i$, $i = 1, 2, 3$. The vector λ in Definition 2.1 can be taken as $(-2, 1, 1, 1, -1)$.

The idea of closed paths with respect to k directions in $\mathbb{R}^d$ was first considered in the paper by Braess and Pinkus [1]. Kłopotowski et al. [10] defined these objects with respect to canonical projections. In our recent paper [7], which deals with linear superpositions, closed paths have been generalized to those having association with k arbitrary functions. In these three works, it was shown that nonexistence of closed paths of the respective form is both necessary and sufficient for

- (1) interpolation with ridge functions [1];
- (2) representation of multivariate functions by sums of univariate functions [10];
- (3) representation by linear superpositions [7].

It should be remarked that consideration of only closed paths is not enough for investigating the problems of representation by linear superpositions in cases where some topology (that of boundedness, or continuity) is involved. As in the case $k = 2$, more general objects must be implemented.

Definition 2.2. A set of points $l = (x_1, \dots, x_n) \subset X$ is called a path with respect to the functions $p_1, \dots, p_k$ if there exists a vector $\lambda = (\lambda_1, \dots, \lambda_n) \in \mathbb{Z}^n \setminus \{0\}$ such that for any $i = 1, \dots, k$,

$$\sum_{j=1}^n \lambda_j \delta_{p_i(x_j)} = \sum_{s=1}^{r_i} \lambda_{i_s} \delta_{p_i(x_{i_s})}, \quad \text{where } r_i \leq k.$$

Note that for $i = 1, \dots, k$, the set $\{\lambda_{i_s}, s = 1, \dots, r_i\}$ is a subset of the set $\{\lambda_j, j = 1, \dots, n\}$. Thus, for each i , we actually have at most k terms in the sum $\sum_{j=1}^n \lambda_j \delta_{p_i(x_j)}$.

Let, for example, $k = 2$, $p_1(x_1) = p_1(x_2)$, $p_2(x_2) = p_2(x_3)$, $p_1(x_3) = p_1(x_4)$, $\dots$, $p_2(x_{n-1}) = p_2(x_n)$. Then it is not difficult to see that for a vector $\lambda = (\lambda_1, \dots, \lambda_n)$ with the components $\lambda_i = (-1)^i$,

$$\sum_{j=1}^n \lambda_j \delta_{p_1(x_j)} = \lambda_n \delta_{p_1(x_n)},$$

$$\sum_{j=1}^n \lambda_j \delta_{p_2(x_j)} = \lambda_1 \delta_{p_2(x_1)}.$$

Thus, by Definition 2.2, the set $l = \{x_1, \dots, x_n\}$ forms a path with respect to the functions p_1 and p_2 .

One can construct many paths by adding not more than k arbitrary points to a closed path with respect to some functions $p_1, \dots, p_k$.

Remark 1. Closed paths and paths with respect to two functions mentioned above in the Introduction satisfy Definitions 2.1 and 2.2, respectively. But for $k = 2$, Definitions 2.1 and 2.2 may allow also some unions of the previously known objects.

Each path $l = (x_1, \dots, x_n)$ and the associated vector $\lambda = (\lambda_1, \dots, \lambda_n)$ generate the functional

$$G_{l,\lambda}(f) = \sum_{j=1}^n \lambda_j f(x_j), \quad f \in C(X). \quad (1)$$

Clearly, $G_{l,\lambda}$ is linear and continuous with norm $\sum_{j=1}^n |\lambda_j|$.

Note that the space $S_c(p_1, \dots, p_k; X)$ is the sum of algebras

$$S_i = \{u_i \in C(X) : u_i = g(p_i(x)), g \in C(\mathbb{R})\}, \quad i = 1, \dots, k.$$

From Definition 2.2 it follows that for each function $u_i \in S_i$, $i = 1, \dots, k$,

$$G_{l,\lambda}(u_i) = \sum_{j=1}^n \lambda_j u_i(x_j) = \sum_{s=1}^{r_i} \lambda_{i_s} u_i(x_{i_s}), \quad (2)$$

where $r_i \leq k$. That is, for each algebra S_i , $G_{l,\lambda}$ can be reduced to a functional defined with the help of not more than k points of the path l . Note that if l is closed, then $G_{l,\lambda}(u_i) = 0$ for all $u_i \in S_i$, $i = 1, \dots, k$, whence $G_{l,\lambda}(u) = 0$, for any $u \in S_c(p_1, \dots, p_k; X)$.

Remark 2. Let $f \in C(X)$. If $G_{l,\lambda}(f) = 0$, for any closed path $l \subset X$, then $f \in S(p_1, \dots, p_k; X)$. That is, f can be represented by linear superpositions. But generally, f may not be in $S_c(p_1, \dots, p_k; X)$ (see [7]).

Theorem 2.3. Let $S_c(p_1, \dots, p_k; X) = C(X)$. Then

- (a) there are no closed paths in X ,
- (b) lengths (number of points) of all paths in X are uniformly bounded.

Proof. The part (a) is obvious. Indeed, let $l = (x_1, \dots, x_n)$ be a closed path in X and $\lambda = (\lambda_1, \dots, \lambda_n)$ be a vector associated with it. As is indicated above, $G_{l,\lambda}(u) = 0$, for any function $u \in S_c(X)$. Let f_0 be a continuous function such that $f_0(x_j) = 1$ if $\lambda_j > 0$ and $f_0(x_j) = -1$ if $\lambda_j < 0$, $j = 1, \dots, n$. Since $G_{l,\lambda}(f_0) \neq 0$, f_0 cannot be in $S_c(X)$. Therefore, $S_c(X) \neq C(X)$. But this contradicts the hypothesis of the theorem.

To prove part (b) of the assertion, consider the linear space

$$U = \prod_{i=1}^k S_i = \{(u_1, \dots, u_k) : u_i \in S_i, i = 1, \dots, k\}$$

endowed with the norm

$$\|(u_1, \dots, u_k)\| = \|u_1\| + \dots + \|u_k\|.$$

We will also deal with the dual of the space U . Each functional $F \in U^*$ can be written as the sum

$$F = F_1 + \dots + F_k,$$

where the functionals $F_i \in S_i^*$ and

$$F_i(u_i) = F[(0, \dots, u_i, \dots, 0)], \quad i = 1, \dots, k.$$

Thus, we see that the functional F determines the collection $(F_1, \dots, F_k)$. Conversely, every collection $(F_1, \dots, F_k)$ of continuous linear functionals $F_i \in S_i^*$, $i = 1, \dots, k$, determines the functional $F_1 + \dots + F_k$ on U . Considering this, in what follows, elements of U^* will be denoted by $(F_1, \dots, F_k)$.

It is not difficult to verify that

$$\|(F_1, \dots, F_k)\| = \max\{\|F_1\|, \dots, \|F_k\|\}. \quad (3)$$

Consider the operator

$$A : U \rightarrow C(X), \quad A[(u_1, \dots, u_k)] = u_1 + \dots + u_k.$$

Clearly, A is a linear continuous operator with the norm $\|A\| = 1$. Besides, since $S_c(X) = C(X)$, A is a surjection. Consider also the conjugate operator

$$A^* : C(X)^* \rightarrow U^*, \quad A^*[H] = (F_1, \dots, F_k),$$

where $F_i(u_i) = H(u_i)$, for any $u_i \in S_i$, $i = 1, \dots, k$. Let H be an arbitrary functional $G_{l,\lambda}$ of the form (1). Set $A^*[G_{l,\lambda}] = (G_1, \dots, G_k)$. From (2) it follows that

$$|G_i(u_i)| = |G_{l,\lambda}(u_i)| \leq \|u_i\| \times \sum_{s=1}^{r_i} |\lambda_{is}| \leq b_\lambda(k) \times \|u_i\|, \quad i = 1, \dots, k,$$

where $b_\lambda(k)$ is the maximum of all numbers $\sum_{s=1}^k |\lambda_{js}|$ formed by k components of the vector λ . Therefore,

$$\|G_i\| \leq b_\lambda(k), \quad i = 1, \dots, k.$$

From (3) we obtain that

$$\|A^*[G_{l,\lambda}]\| = \|(G_1, \dots, G_k)\| \leq b_\lambda(k). \quad (4)$$

Since A is a surjection, there exists a positive real number δ such that

$$\|A^*[H]\| > \delta \|H\|$$

for any functional $H \in C(X)^*$ (see [15]). Taking into account that $\|G_{l,\lambda}\| = \sum_{j=1}^n |\lambda_j|$, for the functional $G_{l,\lambda}$ we have

$$\|A^*[G_{l,\lambda}]\| > \delta \sum_{j=1}^n |\lambda_j|. \quad (5)$$

It follows from (4) and (5) that

$$\delta < \frac{b_\lambda(k)}{\sum_{j=1}^n |\lambda_j|}.$$

The last inequality shows that n (the length of the arbitrarily chosen path l) cannot be as great as possible; otherwise $\delta = 0$. This simply means that there must be some positive integer bounding the lengths of all paths in X . $\square$

Remark 3. The condition (a) of Theorem 2.3 is also necessary for the density of $S_c(p_1, \dots, p_k; X)$ in $C(X)$, whereas the condition (b) is not. For the case $k = 2$, one can use the nontrivial example of Khavinson [8].

Remark 4. The conditions (a) and (b) of [Theorem 2.3](#) are sufficient for the equality $S_c(p_1, p_2; X) = C(X)$ (see [9]). Thus in the case $k = 2$, these conditions are equivalent to the above mentioned condition $\tau^n(X) = \emptyset$ of Sternfeld. For $k > 2$, they are no longer equivalent, since the condition of Sternfeld is not necessary for the representation (see [16]). The question of whether for $k > 2$, our conditions (a) and (b) are sufficient for the representation unfortunately has a negative answer. Our argument is as follows. It can be proven by the same way that it is proved that the conditions (a) and (b) are necessary for the equality $S_b(p_1, \dots, p_k; X) = B(X)$. If they had been sufficient for $S_c(p_1, \dots, p_k; X) = C(X)$, they would have also been sufficient for $S_b(p_1, \dots, p_k; X) = B(X)$, since the representability of continuous functions implies the representability of bounded functions (see [17]). Hence, we would obtain that the conditions (a) and (b) are necessary and sufficient for both the equalities $S_c(p_1, \dots, p_k; X) = C(X)$ and $S_b(p_1, \dots, p_k; X) = B(X)$. But for $k > 2$, these equalities are not equivalent (see [19]).

Acknowledgment

I would like to thank the anonymous referee for his/her comments on improving the original manuscript.

References

- [1] D. Braess, A. Pinkus, Interpolation by ridge functions, *J. Approx. Theory* 73 (1993) 218–236.
- [2] R.C. Cowsik, A. Kłopotowski, M.G. Nadkarni, When is $f(x, y) = u(x) + v(y)$? *Proc. Indian Acad. Sci. Math. Sci.* 109 (1999) 57–64.
- [3] S. Demko, A superposition theorem for bounded continuous functions, *Proc. Amer. Math. Soc.* 66 (1977) 75–78.
- [4] R. Doss, A superposition theorem for unbounded continuous functions, *Trans. Amer. Math. Soc.* 233 (1977) 197–203.
- [5] B.L. Fridman, An improvement in the smoothness of the functions in A. N. Kolmogorov's theorem on superpositions, *Dokl. Akad. Nauk SSSR* 177 (1967) 1019–1022. (in Russian).
- [6] V.E. Ismailov, On the approximation by compositions of fixed multivariate functions with univariate functions, *Studia Math.* 183 (2007) 117–126.
- [7] V.E. Ismailov, On the representation by linear superpositions, *J. Approx. Theory* 151 (2008) 113–125.
- [8] 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. Acad. Nauk. SSSR Ser. Mat.* 33 (1969) 650–666. English transl. *Math. USSR Izv.* 3 (1969), 617–632.
- [9] S.Ya. Khavinson, Best Approximation by Linear Superpositions (Approximate Nomography), in: *Translations of Mathematical Monographs*, vol. 159, American Mathematical Society, Providence, RI, 1997, Translated from the Russian manuscript by D. Khavinson. 175 pp.
- [10] A. Kłopotowski, M.G. Nadkarni, K.P.S. Bhaskara Rao, When is $f(x_1, x_2, \dots, x_n) = u_1(x_1) + u_2(x_2) + \dots + u_n(x_n)$? *Proc. Indian Acad. Sci. Math. Sci.* 113 (2003) 77–86.
- [11] A.N. Kolmogorov, On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition, *Dokl. Akad. Nauk SSSR* 114 (1957) 953–956. (in Russian).
- [12] D.E. Marshall, A.G. O'Farrell, Uniform approximation by real functions, *Fund. Math.* 104 (1979) 203–211.
- [13] D.E. Marshall, A.G. O'Farrell, Approximation by a sum of two algebras. The lightning bolt principle, *J. Funct. Anal.* 52 (1983) 353–368.
- [14] P.A. Ostrand, Dimension of metric spaces and Hilbert's problem 13, *Bull. Amer. Math. Soc.* 71 (1965) 619–622.
- [15] W. Rudin, *Functional Analysis*, second ed., in: *International Series in Pure and Applied Mathematics*, McGraw-Hill, Inc., New York, 1991, 424 pp.
- [16] J.P. Sproston, D. Strauss, Sums of subalgebras of $C(X)$, *J. Lond. Math. Soc.* 45 (1992) 265–278.
- [17] Y. Sternfeld, Uniformly separating families of functions, *Israel J. Math.* 29 (1978) 61–91.
- [18] Y. Sternfeld, Dimension, superposition of functions and separation of points, in compact metric spaces, *Israel J. Math.* 50 (1985) 13–53.
- [19] Y. Sternfeld, Uniform separation of points and measures and representation by sums of algebras, *Israel J. Math.* 55 (1986) 350–362.
- [20] A.G. Vitushkin, G.M. Henkin, Linear superpositions of functions, *Uspekhi Mat. Nauk* 22 (1967) 77–124. (in Russian).