Zeitschrift: L'Enseignement Mathématique

Herausgeber: Commission Internationale de l'Enseignement Mathématique

Band: 23 (1977)

Heft: 1-2: L'ENSEIGNEMENT MATHÉMATIQUE

Artikel: ON REPRESENTATION OF FUNCTIONS BY MEANS OF SUPERPOSITIONS AND RELATED TOPICS

Autor: Vitushkin, A. G.

Kapitel: §1. (,)-entropy and the "dimension" of function spaces

DOI: https://doi.org/10.5169/seals-48931

Chapter 5. — Dimension of the space of linear superpositions

In this chapter we present a calculation of the functional dimension of the space of functions representable by means of linear superpositions and prove that a representation of analytic functions by means superpositions of smooth functions can not be stable.

§ 1. (ε, δ)-*entropy and the "dimension" of function spaces*

Let G_n be a closed region of n-dimensional euclidean space, and $C(G_n)$ the space of all functions continuous in G_n. Two functions $f_1(x), f_2(x) \in C(G_n)$ are called (ε, δ)-distinguishable if there exists an n-dimensional closed sphere $S_\delta \subset G_n$ of radius δ such that

$$\min_{x \in S_\delta} |f_1(x) - f_2(x)| \geqslant \varepsilon .$$

Let $F \subset C(G_n)$ be a set of continuous functions. A subset $K \subset F$ is called (ε, δ)-distinguishable if any two of its elements are (ε, δ)-distinguishable. We denote by $N_{\varepsilon,\delta}(F)$ the maximum number of elements in an (ε, δ)-distinguishable subset of F.

Definition 5.1.1. The number $H_{\varepsilon,\delta}(F) = \log_2 N_{\varepsilon,\delta}(F)$, by analogy with the definition of ε-entropy, is called the (ε, δ)-entropy of F.

Let $f_0 \in F$. We denote by $F_{\lambda\varepsilon}(f_0)$ the set of functions $f \in F$ such that $|f(x) - f_0(x)| \leqslant \lambda\varepsilon$. It follows immediately from the definition that the expression $\overline{\lim\limits_{\delta \to 0}}\ \overline{\lim\limits_{\varepsilon \to 0}} - \dfrac{\log_2 H_{\varepsilon,\delta}(F_{\lambda\varepsilon}(f_0))}{\log_2 \delta}$ as a function of λ does not decrease as $\lambda \to \infty$.

Definition 5.1.2. The number

$$r(F, f_0) = \lim_{\lambda \to \infty} \overline{\lim_{\delta \to 0}}\ \overline{\lim_{\varepsilon \to 0}} - \frac{\log_2 H_{\varepsilon,\delta}(F_{\lambda\varepsilon}(f_0))}{\log_2 \delta}$$

is called the functional "dimension" of F at f_0. The number $r(F) = \sup (F, f_0)$ is called the functional "dimension" of F.

The functional "dimension" $r(F)$ of a set of functions $F \subset C(G_n)$ has the following properties.

5.1.1. Let $\Phi \subset F$ be a set of functions. Then $r(\Phi) \leqslant r(F)$. Moreover, if Φ is everywhere dense in F in the uniform metric, then $r(\Phi) = r(F)$.

Proof. The first part of the assertion follows immediately from the definition. For a proof of the second part it is sufficient to show that $r(\Phi, \varphi_0) \geqslant r(F, \varphi_0)$ for any element $\varphi_0 \in \Phi$. Suppose that the functions $f_1, \ldots, f_N$ from a $(2\varepsilon, \delta)$-distinguishable subset of $F_{\lambda\varepsilon}(\varphi_0)$. Since Φ is everywhere dense in F, there exist functions $\varphi_1, \ldots, \varphi_N \in \Phi$ such that $\max_{x \in G_n} |f_i(x) - \varphi_i(x)| \leqslant \min\left(\frac{\varepsilon}{2}, \lambda\varepsilon\right) (i = 1, 2, \ldots, N)$. These functions form an (ε,δ)-distinguishable subset of $F_{2\lambda\varepsilon}(\varphi_0)$. Consequently $N_{\varepsilon,\delta}(\Phi_{2\lambda\varepsilon}(\varphi_0)) \geqslant N_{2\varepsilon,\delta}(F_{\lambda\varepsilon}(\varphi_0))$. Hence $r(\Phi, \varphi_0) \geqslant r(F, \varphi_0)$.

5.1.2. For any set $F \subset C(G_n)$ we have $r(F) \leqslant n$.

Proof. Suppose that $f_0 \in F$ and $f_1, f_2, \ldots, f_p$ is a maximal set (with respect to p) of pairwise (ε, δ)-distinguishable functions of $F_{\lambda\varepsilon}(f_0)$. Let $\sigma_1, \sigma_2, \ldots, \sigma_q$ be a maximal set (with respect to q) of spheres of radius $\delta/3$ in G_n, such that no two of them have common interior points. Then any pair of functions $f_i(x)$ and $f_j(x)$ of the given set satisfies on at least one of the spheres σ_l the inequality $\min_{x \in \sigma_l} |f_i(x) - f_j(x)| \geqslant \varepsilon$. For the functions $f_i(x)$ and $f_j(x)$ satisfy on some sphere $S_\delta \subset G_n$ the inequality $\min_{x \in S_\delta} |f_i(x) - f_j(x)| \geqslant \varepsilon$. Since q is maximal, it follows that one of the spheres $\sigma_l \subset S_\delta$. Consequently on this sphere the inequality we need is satisfied. We denote by a_l the centre of the sphere σ_l $(l = 1, 2, \ldots, q)$. Every set of functions $f_{i_1}, f_{i_2}, \ldots, f_{i_r}$ each pair of which has values differing by not less than ε at one and the same point consists of a number $r \leqslant 2\lambda + 1$ of functions. (All functions are taken from the set indicated above.) Since every pair of functions $f_i(x)$ and $f_j(x)$ has values differing by not less than ε at one of the points a_l at least, we have $p \leqslant 2\lambda + 1$. But since the spheres $\{\sigma_i\}$ do not intersect, $q \leqslant C/\delta^n$, where C is a constant depending only on n. Consequently,

$$r(F, f_0) \leqslant \lim_{\lambda \to \infty} \lim_{\delta \to 0} \lim_{\varepsilon \to 0} -\frac{\log_2 \log_2 (2\lambda + 1)^{\frac{C}{\delta^n}}}{\log_2 \delta} = n.$$

5.1.3. If F is everywhere dense (in the uniform metric) in the space $C(G_n)$, then $r(F) = n$. In particular $r(C(G_n)) = n$.

Proof. By 5.1.1 and 5.1.2 it is sufficient to show that $r\left(C\left(G_n\right)\right) \geqslant n$. We denote by $C_\varepsilon\left(G_n\right)$ the set of all $f(x) \in C\left(G_n\right)$ for which $\max\limits_{x \in G_n} |f(x)| \leqslant \varepsilon$. Let $\theta > 0$ be a constant such that for any $\delta > 0$ we can find $H = [\theta/\delta^n]$ closed and pairwise non-intersecting spheres $\sigma_1, \sigma_2, \ldots, \sigma_H$ of radius δ in G_n. For any system of numbers $\{\alpha_i\}$ ($\alpha_i = \pm 1$, $i = 1, 2, \ldots, H$) we construct a function $f_{\{\alpha_i\}}(x) \in C_\varepsilon\left(G_n\right)$ such that $f_{\{\alpha_i\}}(x) = a_i \varepsilon$ for $x \in \sigma_i$ ($i = 1, 2, \ldots, H$). These functions are obviously pairwise (ε,δ)-distinguishable. The number of functions $f_{\{\alpha_i\}}(x)$ for all possible sets $\{\alpha_i\}$ is equal to 2^H. Consequently $H_{\varepsilon,\delta}\left(C_\varepsilon\left(G_n\right)\right) \geqslant H = [\theta/\delta^n]$. Hence $r\left(C\left(G\right)\right) \geqslant n$.

COROLLARY 5.1.1. *The space of all polynomials in* n *variables has functional "dimension"* n.

In the same way, the following properties are easily proved.

5.1.4. Let G_n^1 and G_n^2 be two non-intersecting closed regions in n-dimensional space, and $F\left(G_n^1 \cup G_n^2\right)$ a space of functions, defined and continuous on $G_n^1 \cup G_n^2$. Denote by $F\left(G_n^1\right)$ the space of all functions $\varphi(x)$, defined on the set G_n^1, for which there exists a function $\Phi(x) \in F\left(G_n^1 \cup G_n^2\right)$ such that $\varphi(x) \equiv \Phi(x)$ for $x \in G_n^1$. The space $F\left(G_n^2\right)$ is defined similarly. Then

$$r\left(F\left(G_n^1 \cup G_n^2\right)\right) = \max\left\{r\left(F\left(G_n^1\right)\right);\ r\left(F\left(G_n^2\right)\right)\right\}.$$

5.1.5. If F is a linear space, then $r(F) = r(F, f_0)$ for any function $f_0 \in F$. If F is a finite-dimensional linear space, then $r(F) = 0$.

5.1.6. Let F be a linear metric space with metric $\rho(\varphi, \psi)$ between a pair of functions $\varphi, \psi \in F$. We denote by $F(\rho_0)$ the set of all those functions $\varphi \in F$ for which $\rho(\varphi, 0) \leqslant \rho_0$. Then $r(F) = r\left(F(\rho_0)\right)$.

COROLLARY 5.1.2. *The set of all polynomials in* n *variables whose partial derivatives of order* p, *for any* $p = 1, 2, \ldots,$ *are bounded by a constant* $0 < K_p < \infty$ *has functional "dimension"* n.

5.1.7. Let F be a complete linear metric space and $F = \bigcup\limits_{i=1}^{\infty} F_i$, where $\{F_i\}$ are sets of continuous functions. Then $r(F) = \max\limits_i r(F_i)$.

We now write down the main result on the functional "dimension" of a set of linear superpositions.

5.1.8. Let $q_i = q_i(x_1, x_2, \ldots, x_n)$ be continuously differentiable functions of n variables, and $p_i = p_i(x_1, x_2, \ldots, x_n)$ continuous functions of n variables ($i = 1, 2, \ldots, N$). We denote by $F(G_n, \{p_i\}, \{q_i\})$ the set of superpositions of the form $\sum_{i=1}^{N} p_i(x_1, x_2, ..., x_n) f_i(q_i(x_1, x_2, ..., x_n))$, where $(x_1, x_2, ..., x_n) \in G_n$, and $\{f_i(t)\}$ are arbitrary continuous functions of one variable. Then in any region D_n there exists a closed subregion $G_n \subset D_n$ such that

$$r(F(G_n, \{p_i\}, \{q_i\})) \leqslant 1 .$$

For ease of presentation we limit the proof to the case $n = 2$ (§ 3). It is interesting to compare the result 5.1.8 with the following proposition.

5.1.9. Let $\alpha_i(x_1, x_2, ..., x_n) = \sum_{j=1}^{n} \alpha_{ij}(x_j) \quad (i = 1, 2, ..., 2n+1)$

be the continuous functions involved in Kolmogorov's formula (I). We denote by $\psi(G_n, \alpha_i)$ the space of all functions of the form $\psi(\alpha_i(x_1, x_2, ..., x_n))$, where $\psi(t)$ is an arbitrary continuous function of one variable and $(x_1, x_2, ..., x_n) \in G_n$. Then for any i and every region G_n, $r(\psi(G_n, \alpha_i)) = n$ (see 5.1.7).

Let $p_i(x_1, x_2, ..., x_n)$ be fixed continuous functions of n variables, $q_{1,i}(x_1, x_2, ..., x_n)$, $q_{2,i}(x_1, x_2, ..., x_n)$, ..., $q_{k,i}(x_1, x_2, ..., x_n)$ fixed continuously differentiable functions of n variables, and $f_i(t_1, t_2, ..., t_k)$ arbitrary continuous functions of k variables, $k < n$ $(i = 1, 2, ..., N)$. One would expect that the set of superpositions of the form (V) (see Chapter I) has functional "dimension" not greater than k. However, in this direction, only the following partial result has so far been proved.

5.1.10. Denote by $F(\lambda, G_n, \{p_i\}, \{q_{1,i}\}, ..., \{q_{k,i}\})$ the set of all those continuous functions $\varphi(x_1, x_2, ..., x_n)$ for which there exist continuous functions $\{f_i(t_1, t_2, ..., t_k)\}$ such that in G_n.

$$\varphi(x_1, x_2, ..., x_n) = \sum_{i=1}^{N} p_i(x_1, x_2, ..., x_n) f_i(q_{1,i}(x_1, x_2, ..., x_n), ..., q_{k,i}(x_1, x_2, ..., x_n))$$

and

$$\max_i \sup_{(t_1, t_2, ..., t_k)} |f_i(t_1, t_2, ..., t_k)| \leqslant \lambda \sup_{(x_1, x_2, ..., x_n) \in G_n} |\varphi(x_1, x_2, ..., x_n)|$$

Then, for any $\lambda < \infty$, in any region D_n there exists a closed subregion $G_n \subset D_n$ such that

$$r(F(\lambda, G_n, \{p_i\}, \{q_{1,i}\}, ..., \{q_{k,i}\}), 0) \leqslant k .$$

From the last result and Banach's open mapping theorem there follows

COROLLARY 5.1.3. *For any continuous functions* p_i *and continuously differentiable functions* $q_{1,i}, q_{2,i}, \ldots, q_{k,i}$, $k < n$ $(i = 1, 2, \ldots, N)$ *and every region* G_n *there exists a continuous function that is not equal in* G_n *to any superposition of the form* (V).

§ 2. (ε, δ)-*entropy of the set of linear superpositions*

We denote by $S(\delta, z)$ the disc of radius δ with centre at z. Let $p(z) = p(x, y)$ and $q(z) = q(x, y)$ be functions defined in a closed region G of the x, y-plane and having the properties:

a) $p(x, y), \dfrac{\partial q(x, y)}{\partial x}, \dfrac{\partial q(x, y)}{\partial y}$ are continuous in G and have modulus of continuity $\omega(\delta)$,

b) the inequalities $0 < \gamma \leqslant |\operatorname{grad}[q(r)]| \leqslant \dfrac{1}{\gamma}$ and $|p(z)| \leqslant \dfrac{1}{\gamma}$, where γ is some constant, are satisfied everywhere in G.

LEMMA 5.2.1. *Let* $S(\delta, z) \subset G$ *and let* $\mu_q(t)$ *be the function equal to* $2\sqrt{\delta^2 - (t - q(z))^2 |\operatorname{grad}[q(z)]|^{-2}}$ *on*

$$q(z) - \delta |\operatorname{grad}[q(z)]| \leqslant t \leqslant q(z) + \delta |\operatorname{grad}[q(z)]|$$

and equal to zero elsewhere. Then

$$\int_{-\infty}^{\infty} |\mu_q(t) - h_1(e(q, t) \cap S(\delta, z))| \, dt \leqslant c_1(\gamma)\, \omega(\delta)\, \delta^2,$$

where $c_1(\gamma)$ *is a constant depending only on* γ.

Proof. Let $[a, b] \subset e(q, t) \cap S(\delta, z)$ be the segment of the level curve $e(q, t)$, endpoints a and b, lying on the boundary of $S(\delta, z)$; $[z, a]$ and $[z, b]$ the vectors with origin at z and endpoints at a and b, respectively;

$$\alpha_1 = \gamma(\overrightarrow{[z, a]}, \operatorname{grad}[q(z)]), \quad \alpha_2 = \gamma(\overrightarrow{[z, b]}, \operatorname{grad}[q(z)]).$$

We have

$$|t - q(z)| = |q(a) - q(z)| = \left| \int_{s \in [z,a]} \frac{\partial q}{\partial s} ds \right|$$
$$= \delta \cos \alpha_1 |\operatorname{grad}[q(z)]| (1 + 0(1)\, \omega(\delta))$$