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: §2. The entropy of the space of smooth functions

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

$$H_\varepsilon(F^c_{\rho_1,\rho_2,\dots,\rho_n}) - \frac{1}{(n+1)!}\prod_{k=1}^{n}\frac{1}{\log\ \rho_k}\left(\log\frac{c}{\varepsilon}\right)^{n+1} + O\left[\left(\log\frac{c}{\varepsilon}\right)^{n}\log\log\frac{c}{\varepsilon}\right].$$

3. Let $F^n_{s,c}$ be the class of real valued functions on the cube $\{-1 \leqslant x_k \leqslant 1\}$ $(k=1, \dots, n)$, bounded in modulus on that cube by the constant s_k and such that their analytic extensions are entire functions of order s_k with respect to $z_k = x_k + iy_k$ $(k=1, \dots, n)$. Then

$$H_\varepsilon(F^n_{s,c}) = \frac{1}{(n+1)!}\prod_{k=1}^{n} s_k\left(\log\frac{c}{\varepsilon}\right)^{n+1}\left(\log\log\frac{c}{\varepsilon}\right)^{-n} +$$
$$= O\left[\left(\log\frac{c}{\varepsilon}\right)^{n+1}\left(\log\log\frac{c}{\varepsilon}\right)^{-n-1}\right].$$

These estimates and other results connected with estimates of entropy and applications are to be found for example in [49]-[53].

§ 2. *The entropy of the space of smooth functions*

Here we give an estimate of the entropy of the class of S times differentiable functions of n variables. The lower estimate was obtained in [4], the upper one—in [23].

We fix integers $n \geqslant 1$ and $p \geqslant 0$ and numbers $0 \leqslant \alpha \leqslant 1$, $L > 0$, $C > 0$, $\rho > 0$. We will denote by $\mathcal{I}$ the cube $0 \leqslant x_i \leqslant \rho$ $(i=1, \dots, n)$ and by $F = F^{\rho,n}_{S,L,c}$ $(S=p+\alpha)$ the set of all real valued functions defined on $\mathcal{I}$ such that their partial derivatives of order p satisfy the condition Lip α with the constant L and

$$\left|\frac{\partial^{k_1+\dots+k_n} f(0)}{\partial^{k_1}x_1 \dots \partial^{k_n}x_n}\right| \leqslant c \quad \left(\sum_{i=1}^{n} k_i \leqslant p\right)$$

We say that the function $g(x)$ satisfies the condition Lip α with the constant L if for any x' and x''

$$|g(x') - g(x'')| \leqslant L(r(x', x''))^\alpha,$$

where $r(x', x'')$ is the distance between x' and x''.

THEOREM 2.2.1. *If* $\varepsilon > 0$ *is sufficiently small then*

$$A\rho^n\left(\frac{L}{\varepsilon}\right)^{n/s} \leqslant H_\varepsilon(F) \leqslant B\rho^n\left(\frac{L}{\varepsilon}\right)^{n/s},$$

where A *and* B *are positive constants depending only on* s *and* n.

We choose $\delta > 0$ such that the number ρ/δ is an integer. We divide the cube $\mathcal{I}$ into $\left(\frac{\rho}{\delta}\right)^n$ cubes P_i $\left(i = 1, 2, \ldots, \left(\frac{\rho}{\delta}\right)^n\right)$ by hyperplanes, parallel to its $(n-1)$-dimensional edges. Each of the cubes P_i has side of length δ, and the edges of these cubes are parallel to those on $\mathcal{I}$. Let C_i denote the centre of the cube P_i and S_i the n-dimensional closed sphere (inscribed in P_i) of radius $\delta/2$ and centre at the point C_i. Put

$$\varphi_i(x) = \varphi_i(x_1, x_2, \ldots, x_n) = \begin{cases} 0, & \text{if } x \in \mathcal{I} - S_i \\ A\left(1 + \cos\left(\frac{2\pi}{\delta} r(C_i, x)\right)\right)^p & \text{if } x \in S_i, \end{cases}$$

where $r(C_i, x)$ is the distance from the point x to the centre C_i of the sphere S_i. Put, further,

$$\varphi_{\eta_1, \eta_2, \ldots, \eta_h}(x) = \sum_{i=1}^{h} \eta_i \varphi_i(x)$$

$$\left(\eta_i = \pm 1;\ i = 1, 2, \ldots, h;\ h = \left(\frac{\rho}{\delta}\right)^n\right).$$

LEMMA 2.2.1. *We can find a positive number* $A(s, L, n)$, *such that when* $A = A(s, L, n)\,\delta^s$ *and given any set of numbers* $\eta_i\ (i = 1, 2, \ldots, h)$*-the corresponding function* $\varphi_{\eta_1, \eta_2, \ldots, \eta_h}(x)$ *belongs to* F.

Proof. By differentiating $\varphi_i(x)$ it is not difficult to see that inside the sphere S_i its partial derivatives of all orders exist. And the modulus of any partial derivative of order k is bounded inside S_i by $AB(s, k, n)\,\delta^{-k}$, where $B(s, k, n)$ is some constant, depending only on s, k, n. In particular, any derivative of the function $\varphi_i(x)$ of order $p + 1$ is bounded in the sphere S_i by the constant

$$AB(s, p+1, n)\,\delta^{-p-1} = \frac{A(s, L, n)\,B(s, p+1, n)}{\delta^{1-\alpha}}.$$

Let $g(x)$ be any p-th order partial derivative of the functions $\varphi_i(x)$. We take two points a and b belonging to the sphere S_i. Then $g(b) - g(a) = r(a, b)\,\frac{\partial g(c)}{\partial r}$, where $\frac{\partial g(c)}{\partial r}$ is the derivative of $g(x)$ along the direction (a, b), taken at some point c of $[a, b]$. Since any p + 1-th order partial derivative of $\varphi_i(x)$ is bounded inside the sphere by the constant

$\frac{A(s, L, n)\,B(s, p+1, n)}{\delta^{1-\alpha}}$, we have $\left|\frac{\partial g(c)}{\partial r}\right| \leq n\,\frac{A(s, L, n)\,B(s, p+1, n)}{\delta^{1-\alpha}}$

And then

$$| g(b) - g(a) | \leqslant \left| \rho \frac{\partial g(c)}{\partial r} \right| \leqslant \rho n \frac{A(s, L, n) B(s, p+1, n)}{\delta^{1-\alpha}}$$

$$\leqslant \rho^{\alpha} n A(s, L, n) B(s, p+1, n) .$$

Put

$$A(s, L, n) = \frac{L}{2n\, B(s, p+1, n)} .$$

Then

$$| g(b) - g(a) | \leqslant \frac{1}{2} L \rho^{\alpha} .$$

Now let $\Psi(x)$ be any of the p-th partial derivatives of the function $\varphi_{\eta_1, \eta_2, \ldots, \eta_h}(x)$. We choose two points x' and x'' of $\mathscr{I}$ ($x' \in S_i$, $x'' \in S_j$) and let $g_1(x)$ and $g_2(x)$ be the partial derivatives of the same kind as $\Psi(x)$ of the functions $\varphi_i(x)$ and $\varphi_j(x)$ (respectively). It is easy to verify that $g_1(x)$ and $g_2(x)$ are continuous on $\mathscr{I}$ and identically equal to zero on the sets $\mathscr{I} - S_i$ and $\mathscr{I} - S_j$ (respectively). We select some point x_0 belonging to the boundary of the sphere S_i and lying on the segment $[x', x'']$. Then

$$| \psi(x'') - \psi(x') | \leqslant | g_1(x'') - g_1(x') | + | g_2(x'') - g_2(x') |$$

$$\leqslant | g_1(x') - g_1(x_0) | + | g_2(x'') - g_2(x_0) | \leqslant | g(b) - g(a) |$$

$$\leqslant \frac{1}{2} L (r(x', x_0))^{\alpha} + \frac{1}{2} L (r(x'', x_0))^{\alpha} \leqslant L (r(x', x''))^{\alpha} .$$

If one of the points x', x'' (or both) belongs to the set $\mathscr{I} - \sum_{i=1}^{h} S_i$, then we can prove similarly that

$$| \varphi(x'') - \varphi(x') | \leqslant L (r(x', x''))^{\alpha} .$$

Q.E.D.

LEMMA 2.2.2. *There exists a positive constant* A, *depending only on* s, L, n *such that for sufficiently small* ε

$$H_{\varepsilon}(F) \geqslant A \rho^{n} \left(\frac{1}{\varepsilon} \right)^{n/s} .$$

Proof. We choose some positive number $k > 1$ such that when $\delta = \left(\frac{k \varepsilon}{A(s, L, n)} \right)^{1/s}$ is an integer.

We choose two different functions of the type $\varphi_{\eta 1,\dots,\eta h}(x)$ and $\varphi_{\tau 1,\tau 2,\dots,\tau h}(x)$, $A = A(s, L, n)\,\delta^s$ and $A(s, L, n)$ is taken so small that both functions belong to the family F. Since the functions we have chosen are assumed to be different, for some i $\tau_i \neq \eta_i$. And therefore

$$\begin{gathered} |\varphi_{\eta 1,\eta 2,\dots,\eta h}(c_i) - \varphi_{\tau 1,\tau 2,\dots,\tau_n}(c_i)| \\ = 2A = 2A(s, L, n)\,\delta^s = 2k\varepsilon > 2\varepsilon\,. \end{gathered}$$

Hence

$$H_\varepsilon(F) \geqslant \log\, 2^h = \left(\frac{\rho}{\delta}\right)^n = \left(\frac{A(s, L, n)}{k}\right)^{\frac{n}{s}} \rho^n \left(\frac{1}{\varepsilon}\right)^{\frac{n}{s}}$$

Q.E.D.

LEMMA 2.2.3. *There exists a constant* $B > 0$ *such that for sufficiently small* $\varepsilon > 0$

$$H_\varepsilon(F) \leqslant B\rho^n \left(\frac{1}{\varepsilon}\right)^{\frac{n}{s}}$$

Proof. Let us choose some $\delta > 0$ such that the ratio ρ/δ is an integer. In the cube $\mathscr{I}$ consider the uniform lattice with step δ, consisting of the points $d_i\left(i = 1, 2, \dots, h; h = \left(\frac{\rho}{\delta} + 1\right)^n\right)$.

We shall assume the corners of the lattice to be numbered so that the point d_1 coincides with the origin of co-ordinates, and for any i

$$r(d_{i-1}, d) = \delta\,.$$

We now choose some function $f(x)$ of the family F and we shall show a method of constructing a table for this function the volume of which is less than $B\rho^n\left(\frac{1}{\varepsilon}\right)^{n/s}$.

Let h_p denote the number of different kinds of partial derivative (of all orders up to and including the p-th) of a function of n variables. It is not difficult to verify that $h_p \leqslant (p+1)^n$. Let $\{\tau_1^{j,k}\}$ $(\tau_1^{j,k} = 0, 1)$ be the coefficients of the binary representation of the numbers

$$\frac{\partial^{k_1+k_2+\dots+k_n} f(d_1)}{\partial x_1^{k_1}\,\partial x_2^{k_2} \dots \partial x_n^{k_n}} \qquad (k_1 + k_2 + \dots + k_n) \leqslant p$$

written in some order (k is the order of the derivative, $j = 1, 2, \dots, h_1^k$). Then the numbers

$$\left\{\frac{\partial^{k_1+k_2+\ldots+k_n} f(d_1)}{\partial x_1^{k_1}\,\partial x_2^{k_2}\ldots\partial x_n^{k_n}}\right\} \qquad (k_1+k_2+\ldots+k_n=k)$$

are represented in the table to an accuracy of δ^{s-k}, i.e.

$$h_1^k \leqslant \left(\left[\log\frac{c}{\delta^{s-k}}\right]+1\right)\ (k+1)^n$$

binary digits $\tau_1^{s,k}$ $(j=1, 2, \ldots, h_1^k)$ are sufficient to represent them in binary. Thus, to represent all partial derivatives of $f(x)$ at the point $x = d_1$ in binary we need

$$h_1 = \sum_{k=0}^{p} h_1^k \leqslant (p+1)^{n+1}\left(1+\log\frac{c}{\delta^s}\right)$$

binary digits

$$\tau_1^{j,k}\ (j=1, 2, \ldots, h_1^k\ , \quad k=0, 1, 2, \ldots, p)\,.$$

Let us assume now that we have found a method for selecting the digits $\{\tau_1^{j,k}\}$ $(i=1, 2, \ldots, q-1)$ together with a rule for calculating from these digits the values of the numbers

$$\left\{\frac{\partial^{k_1+k_2+\ldots+k_n} f(d_i)}{\partial x_1^{k_1}\,\partial x_2^{k_2}\ldots\partial x_n^{k_n}}\right\} \qquad (k_1+k_2+\ldots+k_n=k)$$

$(i=1, 2, \ldots, q-1)$ to an accuracy of δ^{s-k} $(k=0, 1, \ldots, p)$. We examine the subsequent procedure for constructing the table for $f(x)$. Let $g_k(x)$ be one of the k-th order partial derivatives of $f(x)$. According to the induction hypothesis, the values of all partial derivatives of order $m \leqslant p-k$ of $g_k(x)$ at the point $x = d_{q-1}$ can be calculated to an accuracy of δ^{s-k-m} $(m=0, 1, \ldots, p-k)$ from that part of the table already constructed. From Lagrange's formula, the value of $g_k(d_q)$ is found sufficiently accurately from the approximate values of the derivatives of $g(x)$ at d_{q-1}. Therefore, to represent the numbers $g_k(d_q)$ to an accuracy of δ^{s-k} we need only a small number of binary digits. Since $r(d_{q-1}, d_q) = \delta$ all the corresponding coordinates (except one) of the points d_{q-1}, d_q are equal. For definiteness, we shall suppose that

$$x_1(d_q) = x_1(d_{q-1}) + \delta \quad \text{and} \quad x_i(d_q) = x_i(d_{q-1})$$

for $i = 2, 3, \ldots, n$. Then

$$g_k(d_q) = \sum_{m=0}^{p-m-1} \frac{\partial^m g_k(d_{q-1})}{\partial x_1^m}\cdot\frac{\delta^m}{m\,!}$$

$$+\frac{1}{(p-1)!}\frac{\partial^{p-k} g_k(d_{q-1}+\theta\delta)}{\partial x_1^{p-k}}\delta^{p-k}$$

$$=\sum_{m=0}^{p-k}\frac{\partial^m g_k(d_{q-1})}{\partial x_1^m}\cdot\frac{\delta^m}{m!}+\frac{L}{(p-1)!}\theta\delta^{s-k},$$

where $0 \leqslant \theta \leqslant 1$. But since $\dfrac{\partial^m g_k(d_{q-1})}{\partial x_1^m}$ is given by the table only to an accuracy of δ^{s-k-m} $(m=0, 1, \ldots, p-k)$ $g_k(d_q)$ is determined by the constructed part of the table only to an accuracy of

$$\sum_{m=0}^{p-1}\delta^{s-k-m}\frac{\delta^m}{m!}+\frac{L\delta^{s-k}}{(p-k)!}=\delta^{s-k}\left(\sum_{m=0}^{p-k}\frac{1}{m!}+\frac{L}{(p-k)!}\right)\leqslant e(L+1)^{s-k}$$

Therefore, in order to represent the value of $g_k(d_q)$ in the table to an accuracy of δ^{s-k}, it is sufficient to put another $h_q^{j,k} = [\log((L+1)\,e)] + 1$ binary digits in the table. Hence, to determine the values of all k th order partial derivatives of $f(x)$ it is sufficient to add $h_q^k \leqslant (k+1)^n h_q^{j,k}$ binary digits to the table $(k=0, 1, \ldots, p)$. Thus, the approximate representation of the values of all partial derivatives of the functions $f(x)$ at the point will use only

$$h_q = \sum_{k=0}^{p} h_q^k \leqslant (p+1)^{n+1}\,(1+\log\,[e(L+1)])$$

binary digits.

The volume of the table T which we have constructed is equal to

$$P(T) = \sum_{q=1}^{k} h_q \leqslant (p+1)^{n+1}\left(1+\log\frac{c}{\delta^s}\right)$$
$$+(h-1)(p+1)^{n+1}\,(1+\log\,[e(L+1)])\,.$$

We shall now describe the rule we use to enable us to compute the value of $f(x)$ at any point of the cube $\mathscr{I}$ from the parameters of the table. To do this, we divide the cube $\mathscr{I}$ in some way into sets ω_q $(\omega_q \ni d_q)$ the diameter of each set not exceeding $\delta\sqrt{n}$, and such that $\sum_{q=1}^{h}\omega_q = \mathscr{I}$. The approximate value of the function $f(x)$ is calculated using the parameters $\tau_q^{j,k}$ of T in the following way.

Let $x \in \omega_q$. Then, for the approximate value of $f(x)$ we take

$$f^*(x) = \sum_{k_1+k_2+\ldots+k_n \leqq p} a_{k_1,k_2,\ldots,k_n}\prod_{i=1}^{n}\frac{(x_i-x_i(d_q))^{k_i}}{k_i!}$$

where $a_{k_1,k_2,\dots,k_n}$ is the approximate value (to an accuracy of δ^{s-k}, $k = \sum_{i=1}^{n} k_i$)

of partial derivative $\dfrac{\partial^{k_1+k_2+\dots+k_n} f(d_q)}{\partial x_1^{k_1}\, \partial x_2^{k_2} \dots \partial x_n^{k_n}}$. Since $f(x) \in F$

$$\|f(x) - f^*(x)\| \leqslant \delta^s((p+1)^m + L + 1) = B(s, L, n)\, \delta^s = \varepsilon' .$$

Therefore,

$$H_{\varepsilon'}(F) \leqslant (p+1)^{n+1}\left(1 + \log\frac{c}{\delta^s}\right) + (h-1)(p+1)^{n+1}\left(1 + \log\left(e(L+1)\right)\right).$$

We now define δ in the form

$$\delta = \left(\frac{k\varepsilon}{B(s, L, n)}\right)^{1/s}$$

We choose $k < 1$ so that the ratio ρ/δ is an integer. Then

$$H_\varepsilon(F) \leqslant H_{\varepsilon'}(F) \leqslant (p+1)^{n+1}\left(1 + \log\frac{c}{\delta^s}\right)$$
$$+ (h-1)(p+1)^{n+1}\left(1 + \log\left(e(L+1)\right)\right),$$

i.e. for sufficiently small ε $H_\varepsilon(F) \geqslant B\rho^n \left(\frac{1}{\varepsilon}\right)^{n/s}$, where $B > 0$ is a constant which can be taken to depend on s, L, n only.

Q.E.D.

Proof of the Theorem 2.2.1. First let $L = 1$. Then from lemmas 2.2.2. and 2.2.3 we have

$$A\rho^n\left(\frac{1}{\varepsilon}\right)^{n/s} \leqslant H_\varepsilon(F) \leqslant B\rho^n\left(\frac{1}{\varepsilon}\right)^{n/s}$$

where A and B are positive constant, depending only on s and n, since in this case $L = 1$. But since

$$H_{\frac{\varepsilon}{L}}(F_{s,1,\frac{c}{L}}) = H_\varepsilon(F)$$

for sufficiently small ε

$$A(s, n)\, \rho^n \left(\frac{L}{\varepsilon}\right)^{n/s} \leqslant H_\varepsilon(F) \leqslant B(s, n)\, \rho^n \left(\frac{L}{\varepsilon}\right)^{n/s}$$

Q.E.D.