



Soviet-era science, translated into English

MATHEMATICS

A. G. VITUSHKIN

1957

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-195701.29038>

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.

Abstract

Full Text

MATHEMATICS

A. G. VITUSHKIN

SOME ESTIMATES FROM THE THEORY OF TABULATION

(Presented by Academician A. N. Kolmogorov on 25 XII 1956)

Let a function f of some family F be given, which must be entered into the memory of a machine or, for example, a table of p numbers must be compiled for it, from which f could be reconstructed with a prescribed accuracy ε , while spending no more than k elementary operations on the computations. From the definition of entropy

$$H_\varepsilon(F) = \log N_\varepsilon(F)*$$

it is easy to obtain that, for any method of representing a function about which we know only that it belongs to F , at least $p_0 \geq H_\varepsilon(F)$ binary digits must be expended to store it ⁽¹⁾.

$H_\varepsilon(F)$ binary digits are already sufficient to express a function $f \in F$ with accuracy up to ε , but in this case, generally speaking, it is unclear how, from these binary numbers, to arithmetically reconstruct the value of the function at a given point. This example shows that it is usually important to know not only the amount of information contained in the object under study, but also how complicatedly this information is given to us, i.e., for example, how many elementary operations must be expended in order, from the given p numbers determining a function from F , to obtain its value everywhere in the domain of definition.

For the simplest functional spaces it turns out that the number p of real parameters to be stored and the number k of elementary operations necessary for computing a function from F with accuracy up to ε must satisfy the inequality:

$$p \log k \geq AH_\varepsilon(F),$$

where $A > 0$ is an absolute constant.

Let us formulate more concretely the problem considered here.

Definition. Let F be some collection of real functions f , defined on a set G . An ε -representation of this family will mean an arbitrary set of piecewise-rational functions $r_{q,k}^\varepsilon(y)$ with common barrier of order q ⁽²⁾, defined on the

p -dimensional Euclidean space E_p , such that for every function $f(x)$ from F one can indicate at least one set of real numbers $\{y_1, y_2, \dots, y_p\} = y \in E_p$, for which, for every $x \in G$,

$$|f(x) - r_{q,k}^x(y)| \leq \varepsilon.$$

* Here and throughout below we shall mean the binary logarithm. $N_\varepsilon(F)$ is the cardinality of the minimal ε -net of the space F in the metric C ⁽¹⁾.

In this case, in each of the 2^l regions specified by the barrier,

$$p_{k_1} p_{k_2} \dots p_{k_l} = 0 \quad \left(\sum_{i=1}^l k_i = q \right).$$

The function $r_{q,k}^x(y)$ has the form

$$r_{q,k}^x(y) = r_k^{x, \beta_1, \beta_2, \dots, \beta_l} = r_k^{x, \beta}(y) = \frac{P_k^{x, \beta}(y)}{Q_k^{x, \beta}(y)},$$

where, for all $x \in G$ and $\beta = \beta_1, \beta_2, \dots, \beta_l$ ($\beta_i = \pm 1$, $i = 1, 2, \dots, l$), $P_k^{x, \beta}(y)$ and $Q_k^{x, \beta}(y) \neq 0$ are polynomials in y of degree not exceeding k , whose coefficients depend arbitrarily on x .

Apparently, all the most natural methods of tabulation are ε -representations in the indicated sense. Indeed, every such method is usually reduced to the following: by applying arithmetic operations, for certain values of prescribed parameters we compute a preassigned group of formulas. The obtained values are compared with a preassigned set of constants, or else these values are compared with one another, and, depending on the answers obtained from the comparisons, some of the computed parameters are substituted into one of the groups of formulas fixed in advance, and so on. Repeating this process several times, we obtain the desired result. Such methods, obviously, fit within the framework of the present definition.

Let us fix some ε -representation of the family F . Choose in G some finite set of points $\{x_i\}$ ($i = 1, 2, \dots, m$), and denote by ψ the projection of the space F into the m -dimensional Euclidean space E_m , which assigns to each function $f(x)$ the point $t = t_1, \dots, t_m$ of E_m with coordinates $\{t_i = f(x_i)\}$. Put $\bar{f} = \psi(F)$. The functions $\{t_i = r_{q,k}^{x_i}(y)\}$ define for us a mapping of E_p into E_m (see the definition). Denote this mapping by Φ and put $e = \Phi(E_p)$. Since the functions $\{r_{q,k}^x(y)\}$ realize an ε -representation of the family F , e approximates the set \bar{f} with accuracy up to ε , i.e. every regular ⁽²⁾ m -dimensional cube from E_m with center on \bar{f} and side length greater than 2ε must contain points from e . Then, assuming in what follows that $m > p$, from Lemma 6 ⁽²⁾ we obtain

$$[5m(q+1)(k+q+1)]^{2p} \geq \left(\frac{1}{12\varepsilon}\right)^{\mu_{12\varepsilon}^m(f) - \mu_{12\varepsilon}^p(f)},$$

where $\mu_d^k(f)$ is the maximum, over i , of the quantity

$$\frac{\log Nd(f_k^i)}{-\log d};$$

$Nd(f_k^i)$ is the cardinality of a minimal d -net in the sense of the metric C of the projection f_k^i of the set f onto the i -th coordinate plane of dimension k . From the last relation we obtain

$$2p \log[5m(q+1)(q+k+1)] \geq \log(12\varepsilon)^{\mu_{12\varepsilon}^p(f) - \mu_{12\varepsilon}^m(f)}. \quad (\text{A})$$

Definition. The family F shall be called a C -space if it satisfies the following two conditions:

- a) the metric order $\mu(F)$, equal to

$$\sup \left[\frac{\log \log N_\varepsilon(F)}{-\log \varepsilon} \right]$$

(in the sense of the metric C) is finite;

- b) there exist two positive constants, independent of ε , $d < 1$ and $\alpha \geq 1$, such that for every $\varepsilon > 0$, in the domain of definition of the functions of the family under consideration, one can indicate a set G_ε , consisting of no more than $m = \alpha H_\varepsilon(F)$ points $\{x_i\}$, with the property that every pair of functions from F that coincide on G_ε with accuracy up to $d\varepsilon$ turn out everywhere on G to coincide with accuracy up to ε .

Theorem 1. Let F be a uniformly bounded family of functions that is a C -space. Then the numbers ε, p, q, k of an arbitrary ε -representation* of this family ($\varepsilon < 1/2$) must satisfy the inequality

$$p \log \frac{k+q+1}{\varepsilon} \geq AH_\varepsilon(F),$$

where $H_\varepsilon(F) = \log N_\varepsilon(F)$; $A > 0$ is a constant determined by the metric properties of F and independent of the parameters of the representation.

Indeed, in the construction given above put $m = \alpha H_{12\varepsilon}(F)$, and choose the points $\{x_i\}$ in such a way that item b) holds. Then $N_{12\varepsilon}(f) \geq N_\gamma(F)$, where $\gamma = 12\varepsilon/d$ (see b)), i.e. $\mu_{12\varepsilon}^m(f)(-\log 12\varepsilon) \geq H_\gamma(F)$. Then from inequality (A) we obtain

$$\begin{aligned}
 4p \log(k + q + 1) &\geq -\log 12\varepsilon [\mu_{12\varepsilon}^m(f) - \mu_{12\varepsilon}^p(f)] - 2q \log(5m) \geq \\
 &\geq H_\gamma(F) - p \log \frac{3C}{12\varepsilon} - 2p \log[5\alpha H_{12\varepsilon}(F)],
 \end{aligned}$$

where C is a constant bounding the maximum of the modulus of functions from F , and $p < m$.

From the inequality obtained, taking into account that for $\beta > 1$

$$H_\varepsilon(F) \geq H_{\beta\varepsilon}(F) \geq \left(\frac{1}{\beta}\right)^{\mu(F)} H_\varepsilon(F) \quad \text{and} \quad H_{12\varepsilon}(F) \leq \left(\frac{1}{12\varepsilon}\right)^{\mu(F)},$$

we obtain that for $p < m$

$$4p \log(k + q + 1) \geq A_1 H_\varepsilon(F) - A_2 p \log \frac{1}{\varepsilon} - A_3 p \log \frac{1}{\varepsilon},$$

i.e.

$$p \log \left(\frac{k + q + 1}{\varepsilon} \right) \geq A_4 H_\varepsilon(F);$$

for $p \geq m$ we obtain

$$p \log \left(\frac{k + q + 1}{\varepsilon} \right) \geq m \geq A_5 H_\varepsilon(F),$$

since $\varepsilon < 1/2$.

Taking A to be the minimum of A_4 and A_5 , we obtain from the last inequalities the assertion of the theorem.

In some cases the estimate given in Theorem 1 can be improved by removing ε from the left-hand side of the inequality. This can be done, for example, when the number of values of functions from F is determined mainly not by the entropy of F , but by certain properties of functions from F , for example, their smoothness. We shall prove this in one particular case.

Denote by $F_{(l+\alpha),c_1,c_2}^n$ the space of all real functions defined on the n -dimensional closed unit cube I_n , bounded on I_n above and below by the constant c_1 , and having everywhere on this cube partial derivatives of all orders from 1 to l , satisfying a Hölder condition with exponent α and constant c_2 (the metric is taken to be C).

Theorem 2. The numbers ε, p, q, k of an arbitrary ε -representation of the family $F_{(l+\gamma),c_1,c_2}^n$ must satisfy the inequality

$$p \log(k + q + 1) \geq B \left(\frac{1}{\varepsilon} \right)^{\frac{n}{1+\alpha}} \geq CH_\varepsilon(F_{(l+\alpha), c_1, c_2}^n),$$

where $B > 0$ and $C > 0$ are constants depending only on n, l, c_1, c_2 .

* We note that the assertion of the theorem remains valid if the barrier of the ε -representation is allowed to depend on x .

Proof. In the construction presented, put

$$m = B_1 \left(\frac{1}{\varepsilon} \right)^{\frac{n}{1+\alpha}}$$

and distribute the points x_i uniformly in I_n . Denote by \mathfrak{w} the regular cube from E_m , with center at the origin and side length $2C\varepsilon$ ($C < 12$, see Theorem 1 in ⁽²⁾). We fix B_1 so small that

$$\mathfrak{w} \subset f = \psi \left(F_{(1+\alpha), c_1, c_2}^n \right).$$

Then from Theorem 1 ⁽²⁾, Lemma 5 ⁽²⁾, and the fact that $e = \Phi(E_p)$ must approximate f with accuracy up to ε , we obtain

$$\frac{2C\varepsilon}{\sqrt[m-p]{C^{m^2}(q+1)^p(2q+2k+1)^p}} \leq 2\varepsilon,$$

whence it is not difficult to derive the first half of the inequality of the theorem being proved. The second half of this inequality was proved by A. N. Kolmogorov in ⁽¹⁾.

Denote by $F_{\infty, r, d}^n$ the space (in the sense of the metric C) of all functions analytic on I_n , not exceeding on I_n , in modulus, the number $d > 0$, and having convergence coefficient $r > 1$; i.e., for every k , for every function f from $F_{\infty, r, d}^n$, one can indicate a polynomial of degree k approximating f everywhere on I_n with accuracy up to $1/r^k$.

Theorem 3. The numbers ε, p, k, q of an arbitrary ε -representation of the family $F_{\infty, r, d}^n$ must satisfy the inequality

$$p \log \frac{k + q + 1}{\varepsilon} \geq BH_\varepsilon(F_{\infty, r, d}^n) \geq C \left[\log \frac{1}{\varepsilon} \right]^{n+1},$$

where $B > 0$ and $C > 0$ are constants independent of ε .

The proof is analogous to the proof of Theorem 2.

Remark. From the results given here it follows quite clearly that it is impossible to devise a method of tabulation which, for the simplest functional spaces, would be essentially better than interpolation methods.* Indeed, for interpolation methods $q = 0$, since at every point $x \in I_n$ the function is computed as a polynomial according to a common formula (for all functions); the number k of arithmetic operations is finite and does not depend on ε , while the number of numbers to be stored has order

$$\left(\frac{1}{\varepsilon}\right)^{\frac{n}{1+\alpha}} = CH_\varepsilon(F_{1+\alpha, c_1, c_2}^n),$$

i.e., the parameters of the interpolation method turn the inequality of Theorem 2 into an equality (up to a factor on the right-hand side). This means that the “complexity” $p \log(k + q + 1)$ for any ε -representation of the family $F_{(1+\alpha), c_1, c_2}^n$ is no less than the corresponding expression for interpolation methods.

Received
25 XII 1956

REFERENCES

¹ A. N. Kolmogorov, DAN, **108**, No. 3 (1956). ² A. G. Vitushkin, DAN, **114**, No. 4 (1956).

* Here are meant methods based on approximating functions by polynomials in x , the coefficients of which are determined by the values of the function being computed at certain previously fixed points.

Note: Figure translations are in progress. See original paper for figures.

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.