Digitalization of Information Specified on the Grid

*Gurami Tsitsiashvili, **Yuriy Kharchenko

•

Russia, 690041, Vladivostok, Radio street 7
Institute for Applied Mathematics
Far Eastern Branch of Russian Academy Sciences
*guram@iam.dvo.ru, **har@iam.dvo.ru

Abstract

Digitalization is the process of implementing digital transmission systems at the level of primary networks, switching and control facilities that ensure the transmission and distribution of information flows in digital form at the level of secondary networks, which makes production more flexible, competitive, and profitable. First point considered here is an introduction of mathematical equivalent to the concept of a pixel, used when replacing the original information with its step-by-step approximation and estimate of its accuracy. Second point is the study of special knots: extreme knots or saddle knots on the grid and construction of level lines around them: ellipses or hyperbolas. This construction is connected with some meteorological problems and is based on the concept of positive definite quadratic form. Third point is an estimation of the average number of Poisson flow points in several cells of a square lattice in different problems of earth sciences. It is solved by introduction of relative error of the estimate.

Keywords: pixel, step–by–step approximation, positive definite quadratic form, relative error.

1. Introduction

Currently, the program of digitalization of information is being widely developed. Digitalization in a broad sense is the process of implementing digital transmission systems at the level of primary networks, switching and control facilities that ensure the transmission and distribution of information flows in digital form at the level of secondary networks, which makes production more flexible, competitive, and therefore more profitable. This trend allows us to take a fresh look at the already established methods in the processing of information and requires an assessment the accuracy of information conversion in various aspects and ensure the further development of information technology.

Such a statement of the question leads, among other things, to estimates of accuracy when replacing the original information with its digital expression. In turn, the accuracy estimates significantly affect the correctness of the actions of technical systems when receiving external information, and therefore the reliability of their operation. This requires the need to build mathematical equivalents to the concepts used in the process of digitalization. In particular, there is a need to take a fresh look at the concept of a pixel used when replacing the original information with its step-by-step approximation. As a result, it becomes necessary to study the transition from the original signals to the signals specified on a certain lattice. In turn, the reference to the signals given on the grid raises new questions for those tasks where such information is used.

Here, this problem is explored for some meteorology questions. It turns out that in these problems of meteorology, an important aspect of the analysis of such information is the study of singular points on the grid. A detailed study of special points (extreme points or saddle points) allows us to set and to solve the problems of forecasting meteorological systems in a new way.

These questions are the subject of research in this paper. Special attention is paid here to the accuracy of reproducing information from its discrete images and to the possibility of compressing information for its further use. For this purpose, it is possible to use such classical mathematical constructions as multidimensional Taylor series, positive definite quadratic forms, and etc.

Last problem is in estimating the mean number of Poisson flow points in some domain consisting of few cells in square grid. Its solution is based on a concept of relative error and on properties of Poisson distribution. It is connected with earth sciences problems, for example, with calculation of a number of rare animals.

2. The mathematical equivalent of the term "pixel"

In computer science, the term "pixel" (Engl. "pixel" is short for pictures element) [1] is the smallest logical two-dimensional element of a digital image in raster graphics, or a physical element of the matrix of displays that form the image. It is an indivisible object of rectangular (or round) shape, characterized by a certain color.

Signals that transmit sound or time-varying images are currently being digitized to make it convenient to transmit them from one point to another. At the same time, to determine and evaluate the quality of the transmitted information, it is desirable to construct a mathematical equivalent of the concept of "pixel".

For this purpose, it is natural to use a step-by-step approximation of the functions representing the transmitted signals. The quality of such an approximation increases with a decrease in the sampling step (in time and/or coordinate). Given the significance of this dependence of the approximation quality on the sampling step, it is natural to express this dependence mathematically.

Let's start solving the problem by analysing the stepwise approximation of the function $f(t)$, given on the half–interval [0,1). Let's assume that this function is continuously differentiable and for some positive F the equality holds

$$\sup_{0\le t\le 1}|f'(t)| = F. \tag{1}$$

We divide the semi-interval [0,1) into n parts by points $i/n,\ i = 0,1,\dots,n-1$. On the half–interval, $S_i = \left[\frac{i}{n},\frac{i+1}{n}\right)$ we approximate the function $f(t)$ by constant $f\left(\frac{i}{n}+\frac{1}{2n}\right)$, by constructing a stepwise approximation in this way $\widehat{f}\,(t)$. Using the decomposition of the function $f(t)$ into a Taylor series with a Lagrange residual term in the neighbourhood of the radius $\frac{1}{2n}$ of the point $\frac{i}{n}+\frac{1}{2n}$, $i = 0,1,\dots,n-1$, we obtain the following inequality

$$\sup_{0\le t<1}\left|f(t)-\widehat{f}\,(t)\right| \le \frac{F}{2n}. \tag{2}$$

Let us now consider the continuously differentiable function of the m–dimensional argument $f(t_1,t_2,\dots,t_m)$, $(t_1,\dots,t_m)\in[0,1)^m$. Suppose that there are such positive numbers $F_1,F_2,\dots,F_m$, that the relations are satisfied

$$\sup_{0\le t_1,t_2,\dots,t_m<1}\left|\frac{\partial f(t_1,t_2,\dots,t_m)}{\partial t_k}\right| = F_k,\ \ k = 1,2,\dots,m. \tag{3}$$

We divide the direct product $[0,1)^m$ of half–intervals into direct products of half–intervals of the form $S_{i_1} \times S_{i_2} \times \ldots \times S_{i_m}$, $i_1, i_2, \ldots, i_m = 0,1\ldots, n-1$. Let's construct a stepwise approximation of $\hat{f}\,(t_1, t_2, \ldots, t_m)$ functions $f(t_1, t_2, \ldots, t_m)$, assuming it to be equal

$$f\left(\frac{i_1}{n} + \frac{1}{2n}, \frac{i_2}{n} + \frac{1}{2n}, \ldots, \frac{i_m}{n} + \frac{1}{2n}\right)$$

in a direct product $S_{i_1} \times S_{i_2} \times \ldots \times S_{i_m}$, $i_1, i_2, \ldots, i_m = 0,1, \ldots, n-1$.

Using the decomposition of the function $f(t_1, t_2, \ldots, t_m)$ in Taylor m–dimensional series (see, for example, [2]) with a Lagrange-shaped residual term in the set $S_{i_1} \times S_{i_2} \times \ldots \times S_{i_m}$, $i_1, i_2, \ldots, i_m = 0,1\ldots, n-1$ we get the inequality

$$\sup_{0 \le t_1, t_2, \ldots, t_m < 1} \left| f(t_1, t_2, \ldots, t_m) - \hat{f}\,(t_1, t_2, \ldots, t_m) \right| \le \frac{1}{2n} \sum_{k=1}^{m} F_k. \tag{4}$$

Each step in the $\hat{f}$ approximation of the f function can be called a pixel. Moreover, with a decrease in the linear pixel size $\frac{1}{n}$, the accuracy of such an approximation increases in accordance with the formula (4). Note that the accuracy estimates of the step approximation depending on the linear pixel dimensions are expressed in a uniform metric.

3. Geometric interpretation of meteorological information in square grid nodes

The most important element of the structure of the pressure field at an altitude of 5 km above the Far East is a stable and extensive hollow. Its intensity and geographical localization largely determine the nature of atmospheric circulation and weather [3], [4]. When studying this hollow, the nodes of the square grid on the map are identified, the node in which the pressure field takes the minimum value and the pressure level isolines are built.

However, it is quite difficult to algorithmize and analytically investigate such a construction. Therefore, in this paper, an attempt is made to investigate the function given at the lattice nodes in a small neighbourhood of the minimum point, assuming that the step of the square lattice is sufficiently small. Assuming that the smooth function given at the lattice nodes reaches a minimum at the point of the lattice node, one can calculate the coefficients of the Taylor series.

The first-order coefficients are zero, and the second-order coefficients define a positive-definite quadratic shape, whose level lines are ellipses. Calculations show that the orientation of the major and minor axes of such an ellipse and their ratio largely determine the nature of atmospheric circulation. But since the second coefficients of the Taylor series are determined by the values of the function at the nodes of the square lattice, it is necessary to investigate the algorithm for determining them and its accuracy depending on the length of the lattice step.

Let the thrice continuously differentiable function $f(x, y)$ be defined on the rectangle $D = \{0 \le x \le Nh,\ 0 \le y \le Mh\}$ and measured in points (ih, jh), $i = 0, \ldots N$, $j = 0, \ldots, M$. Everywhere else, we assume that the value of the lattice step h is sufficiently small. It is known that at the point (kh, lh), the function f reaches the global minimum. Moreover, the point (kh, lh) is internal in the discrete set $\{(ih, jh),\ i = 0, \ldots N,\ j = 0, \ldots, M\}$, $0 < k < N$, $0 < l < M$. Imagine the decomposition of the function $f(x, y)$ into a Taylor series with a Lagrange residual term under the condition $|x - kh| \le h$, $|y - lh| \le h$:

$$f(x, y) = f(kh, lh) + \frac{1}{2}[A(x - kh)^2 + B(y - lh)^2 + 2C(x - kh)(y - lh)] + O(h^3),$$

$$A = f_{x,x}(kh, lh),\ B = f_{y,y}(kh, lh),\ C = f_{x,y}(kh, lh).$$

Since the point (kh, lh) is the point of the global minimum of the function $f(x, y)$, then the quadratic form $A(x-kh)^2 + B(y-lh)^2 + 2C(x-kh)(lh)$ is positive definite and hence the inequalities are satisfied $A+B>0,\ AB>C^2$.

We construct finite-difference estimates of partial derivatives A, B, C:

$$a = \frac{f((k+1)h, lh) - 2f(kh, lh) + f((k-1)h, lh))}{h^2} = A + O(h),$$

$$b = \frac{f(kh, (l+1)h) - 2f(kh, lh) + f(kh, (l-1)h)}{h^2} = B + O(h),$$

$$c = \frac{f\big((k+1)h, (l+1)h\big) - f\big((k+1)h, lh\big) - f(kh, (l+1)h) + f(kh, lh)}{h^2} = C + O(h).$$

Then the function f may be approximated by the function $\widehat{f}$ with an accuracy of $O(h^3)$ in variables $X = \frac{x-kh}{h},\ Y = \frac{y-lh}{h}$:

$$\widehat{f}\,(x, y) = f(kh, lh) + \frac{1}{2}(aX^2 + bY^2 + 2cXY),\ a+b>0,\ ab>c^2$$

and so the quadratic form $aX^2 + bY^2 + 2kxy$ is also positive definite.

We reduce this quadratic form to a diagonal form (see, for example, [5]), for which we construct its matrix $A = \begin{pmatrix} a & c \\ c & b \end{pmatrix}$ and write out the characteristic equation $(a-\lambda)(b-\lambda) - c^2 = 0$. The roots of this equation are

$$\lambda_{\pm} = \frac{a+b}{2} \pm \sqrt{\frac{(a+b)^2}{4} - ab + c^2} > 0,$$

the eigenvalues of the matrix A, and its orthonormal eigenvectors $\vec{n}_{\pm}$ satisfy the linear equations $A\vec{n}_{\pm} = \lambda_{\pm}\vec{n}_{\pm}$.

Let's move to the coordinate system (u_+, u_-) with an orthonormal basis $\vec{n}_+\vec{n}_-$. In this coordinate system, the quadratic form is $aX^2 + bY^2 + 2xy$ is represented by the sum of squares $\lambda_+u_+^2 + \lambda_-u_-^2$. The level lines of this square shape are ellipses of the form $\lambda_+u_+^2 + \lambda_-u_-^2 = const > 0$. Denote $k = \sqrt{\frac{\lambda_+}{\lambda_-}}$, then to construct the specified ellipses, the circles given by the equation $u_+^2 + u_-^2 = const$ should be compressed along the u_+ axis by k times. The coefficient k may be interpreted as the ratio of the major and minor axes of an ellipse whose level lines are defined by a quadratic shape $aX^2 + bY^2 + 2cXY$.

It is interesting to note that if $f_x = f_y = 0$ and the condition $ab < c^2$, is satisfied, then it is not difficult to establish that $\lambda_+ > 0,\ \lambda_- < 0$ and hence the quadratic form $aX^2 + bY^2 + 2cXY$ in the variables u_+, u_- has the form $\lambda_+u_+^2 + \lambda_-u_-^2$ and is an alternating sign, and its level lines are hyperbolas.

4. Error in estimating the mean number of Poisson flow points

Specialists in the field of earth sciences have the task of estimating the error of the mean number of Poisson flow points from observations in different cells of a square grid. Let the study area be divided into m cells, and the number of points in them in the area k is $n_k,\ k = 1, \dots, m$. It is natural to assume that the random variables $n_1, \dots, n_m$ are independent and have Poisson distributions with the parameters $\lambda_1, \dots, \lambda_m$. Using the properties of the Poisson distribution, it is easy to establish that the random sum $N = \sum_{k=1}^{m} n_k$ has a Poisson distribution with the parameter $\Lambda = \sum_{k=1}^{m} \lambda_k$ and consequently $E(N) = \Lambda,\ Var(N) = \Lambda$.

Using the known properties of the mathematical expectation and the variance of the Poisson distribution, we proceed to estimate the relative error. To do this, consider the random variable $\frac{N}{E(N)} = \frac{N}{\Lambda}$. Variance of this random variable $Var\left(\frac{N}{\Lambda}\right) = \frac{1}{\Lambda}$ and so the relative error of such an estimate satisfies the relation $\sqrt{Var\left(\frac{N}{\Lambda}\right)} = \frac{1}{\sqrt{\Lambda}}$. Therefore, the relative error decreases with the growth of Λ.

This result does not depend on the nonuniformity of the distribution density of the Poisson flow of points, and therefore does not depend on the parameters $\lambda_1, \dots, \lambda_m$. It can be considered by choosing the efficiency indicator of a complex system like a relative error. This result is based on the well-known models of Poisson point flows in the theory of random sets, which are used in the earth sciences [6].

References

[1] Pixel. Merriam-Webster.com Dictionary, Merriam-Webster, https://www.merriam-webster.com/dictionary/pixel.

[2] Ilyin, V. A., Sadovnichy, V. A., Sendov, B. Kh. Mathematical analysis. Part 1, ed. 3, ed. Tikhonov, A. N. Prospect, Moscow, 2004. (In Russian).

[3] Plotnikov, V. V., Mezentseva, L. I., Dubina, V. A. Atmospheric circulation over the Far East and its reflection in ice processes. Dalnauka, Vladivostok, 2015. (In Russian).

[4] Shatilina, T. A., Tsitsiashvili, G. Sh., Radchenkova, T. V. (2019). Variability of the tropospheric hollow regime near the East coast of Asia and climatic extremes in the period 1980 -2017. Meteorology and Hydrology, 8: 102–106. (In Russian).

[5] Beklemishev, D. V. Analytical geometry and linear algebra. High School, Moscow, 1998. (In Russian).

[6] Stoyan, D., Kendall, W., Mecke, J. Stochastic Geometry and Its Applications. J. Wiley and Sons: New York, 1987.