

About Calculation Singularity of High-Order Derivative for Identification of the Graphic Objects Shape

I. M. Gostev

*National Research University Higher School of Economics
20, Myasnitskaya str., Moscow, Russia, 101000*

Methods of calculation of high order derivatives are considered on a basis: interpolation formulas; “without difference methods of calculation of derivatives”; applications of convolution with replacement of differentiation by integration operation; differentiation with use of quadratures on C. Lanczos; the method of Numerova. The comparative analysis of methods of calculation of high order derivatives on accuracy of calculations with use as the sample of the derivatives calculated in package Maple with 20 digit decimal accuracy is carried out. It is shown that all methods are almost equivalent on accuracy and are reduced to convolution calculation between differentiated function and some window which coefficient depend on an applied method. For carrying out of experiments the special program complex is developed for calculation of high order derivative (up to 7th) the tabulated functions with various step. Grids with steps from 0.005 to 0.1 have been investigated. Irrespective of a method of calculation of derivatives it has been defined that optimum value of step mesh for 64 digit arithmetic’s the step is from 0.01 till 0.05. Value of smooth functions differs less than their accuracy of representation at smaller value of a step, and at greater step — the differentiation error increases. Results of experiments confirm N.N. Kalitkin’s theoretical conclusions.

Key words and phrases: graphic pattern recognition, image processing, computer geometry, calculation derivative, calculation accuracy, object identification, line correlation metrics.

1. Introduction

The problem solution of graphic objects identification is almost always reduced to identification of curved lines which are fragments of some object contours. Identification of the closed contours has been repeatedly stated in the literature, for example, in [1] and is based on the signature analysis and methods of geometrical correlation [2]. Often in real conditions reception of the objects closed contours may be impossible; however it is easy to receive fragments of contours of object. In this case other methods of identification, for example, founded on *k-jet* [3] are necessary. For this methods the system of reference points constructed on zero of derivative various orders from the identified curve is used. As numerical differentiation researches on calculation of derivative higher orders by various numerical methods have been conducted is in that case applied and the right accuracy from which it is necessary to calculate derivatives for identification of curves is defined.

2. Line Correlation Metrics

Input next concepts.

Definition. Let $f(t)$ be a real function that is k times differentiable on the interval $[a, b]$. The $k - jet$ zeros of the j -th order ($1 \leq j \leq k$) of the function f are the points $t_1^{(j)}, t_2^{(j)}, \dots, t_{n_j}^{(j)}$, at which the j -th derivative of this function $k - jet$ equals zero: $f^{(j)}(t_r^{(j)}) = 0, r = 1, \dots, n_j, \forall t_r^{(j)} \in [a, b]$.

Consider a system of informative parameters that is based on the concept of $k - jet$ zeros. To this end, we should consider the mechanism of its application to identification in more detail.

Let functions $f(t)$ and $h(t)$ be defined on a set

$$G = \{0 < t_0 < t_1 < \dots < t_{m-1} < t_m = 1\}$$

by an irregular grid. Consider the grid functions $f_m = \{f(t_0), \dots, f(t_m)\}$ and $h_m = \{h(t_0), \dots, h(t_m)\}$. The distance between these functions can be defined by the relation

$$\rho_m(f_m, h_m) = \|f_m - h_m\| = \max_{t_j} |f(t_j) - h(t_j)|.$$

Taking into consideration that the role of a support set for the identification of grid functions is played by the set G_f formed by an ordered set of k – *jet* zeros, we can evaluate the closeness of two functions with the use of some metrics.

Definition 1. Introduce a discrete two dimensional difference function of values for the functions f and h has follows:

$$\eta_{i,j} = f(t_{i+j}) - h(\tau_i), \quad j = \overline{0, (l-m)}, \quad i = \overline{0, m}, \quad \tau \in G_h, \quad t \in G_f.$$

Define a discrete function of absolute identification error

$$\sigma_j = \frac{1}{m+1} \sum_{i=0}^m |\delta_j - \eta_{i,j}|, \quad j = \overline{0, (l-m)}.$$

and a discrete function of relative identification error

$$\delta_j = \frac{1}{(m+1)} \sum_{i=0}^m |\eta_{i,j}|, \quad j = \overline{0, (l-m)}.$$

We represent a recognition function for the identification of unclosed curves on the basis of linear correlation 1(LC1) as

$$\lambda_{LK1} = \begin{cases} 1, & (\rho_{LK1} < \varepsilon_{LK1}) \vee (\rho_{LK1}^z < \varepsilon_{LK1}), \\ 0, & (\rho_{LK1} \geq \varepsilon_{LK1}) \wedge (\rho_{LK1}^z \geq \varepsilon_{LK1}), \end{cases} \quad (1)$$

where $\rho_{LK1} = \min_j \delta_j$ is the metric of LC1, ρ_{LK1}^z is the metric of LC1 calculated on the mirror set of k – *jet* zeros of the object, and ε_{LK1} is a classification tolerance that defines a given identification accuracy. The equality $\lambda_{LK1} = 1$ implies successful curve identification, and j — is a shift of the position of the object with respect to the origin of the reference line.

3. Methods of Derivatives Calculation

Calculation of high-order derivatives tabular to the set function is a challenging task. The analysis of numerical methods of differentiation has shown the presence of a considerable quantity of various receptions. The majority of them has been developed in the late fifties – the beginning of sixtieth years of the last century. We will briefly state some of them.

1. *A method of derivatives calculation on the basis of interpolation formulas.* Here the technique for some function set decomposition and consecutive differentiation of members of a number before reception of the necessary order of a derivative is used. Calculation of derivatives is carried out through calculation of separate differences. Most full such procedure is stated in [4]. As a result of manipulations with the separate differences substituted set of Taylor from investigated function some polynomial with the parameter defining position on the investigated interval (usually $[0, 1]$) is received. Depending on the used formula of interpolation a set of

methods is available — CityNewton, placeCitySterling, Erlanger, Bessel, Gauss, etc.

As a result of substituting the concrete parameter value a certain formula is received on which it is possible to receive the value of a derivative, using some points of differentiated function. For example, using the formula of placeCitySterling on 7 points, value of the first derivative can be calculated according to the following formula

$$\frac{df}{dt} = \frac{1}{60}f_{-3} - \frac{3}{20}f_{-2} + \frac{1}{4}f_{-1} + f_0 - \frac{5}{4}f_1 + \frac{3}{20}f_2 - \frac{1}{60}f_3. \quad (2)$$

2. Other technique — “derivatives calculations without difference”, is stated in M. Abramowitz and I. Stegun’s handbook [5] in which predefined coefficients of the formula of numerical differentiation of k -th order on n to points (Table 25.2 on 708 p.) are brought. Actually, formulas of derivatives calculation on these factors are similar (2), except for their values.
3. The third technique of differentiation is stated for the first time in Gelfand I.M. and Shilov G.E. book [6]. It consists in application of integration instead of differentiation operation through use of the following convolution property:

$$D(f * g) = Df * g = f * Dg.$$

If a smooth function is chosen as differentiated function, for example Gaussian, as a result of function convolution calculation with derivative from Gaussian, we will receive the required result. Moreover this operation allows to smooth initial function simultaneously with differentiation operation.

4. The fourth technique of differentiation — called regularization — was published in placecountry-regionRussia in the Lanczos’ book [7] in 1961. The method is called “differentiation with the use of quadratures” as the author approximated values of function on the values of a square-parabola in points of derivatives calculation. As a result of values substitution of a square-parabola the following formula for the first derivative was deduced:

$$f'(x) = \frac{\sum_{a=-k}^{+k} a f(x + ah)}{2 \sum_{a=1}^k a^2 h}.$$

The first derivative is calculated on five points with factors-2,-1, 0, 1, 2; for calculation of the second derivative already it is required nine points - 4,4,1,-4,-10,-4,1,4,4, etc.

5. The last among the investigated methods is a the Numerova’s method in which function decomposition set is used again, but factors at values of function get out in such a manner that for the second derivative the error corresponds to the fourth order (for the first — to the third).

4. Modeling Toolkit

For the research of received derivatives quality the software was developed, that allows to realize all methods set forth above with the maximum accuracy, possible on the 64-digit computer. Applications enable to calculate derivatives up to the sixth order inclusive and allow to load a reference derivative, and then to compare it with received in different methods. As a reference derivative the functions calculated analytically by means of the Maple package undertook, and then their tabulated value

with 20 meaning (decimal) numbers remained in a file. The difference between functions was calculated on placeCityEuclid metrics. The sample program with the results of differentiation of Gaussian function and its 6th derivatives is brought on Fig. 1.

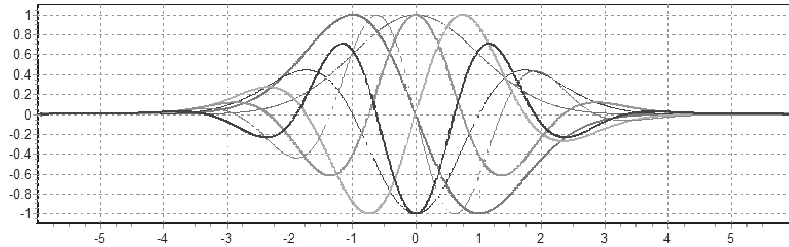


Figure 1. Gauss function and its six its derivatives

5. Result of Modeling

Through modeling two problems were set:

1. Using various methods of numerical differentiation to define which method has the greatest accuracy.
2. To define the optimum step of differentiation for modern computing systems with 64 digit arithmetic.

As a result of the moved researches the following conclusions were drawn:

1. All considered methods of numerical differentiation use **convolution operation**. Where one of the functions is a differentiated function and another is presented by some sequence of coefficients (a convolution window) which value defined on a concrete method. On the basis of this and as result of the measurements it was established that the speed of calculations does not depend on a method and varies for different differentiated functions a little.
2. According to N.N. Kalitkin's theoretical results [8] about the behavior of a differentiation error shown on Fig. 2. Grids with steps from 0.005 to 0.1 have been investigated.

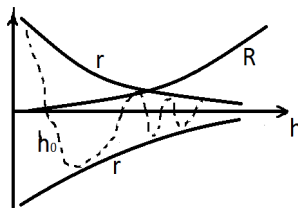


Figure 2. The behavior of a differentiation error depending on a step size. Here R is a differentiation error determined by the grid step, r is an error in calculating the difference of the function

Irrespective to the method of derivatives calculation it was defined that with the optimum value of grid step for 64 digit arithmetic the step is from 0.01 till 0.05. At a lower step size values of smooth functions differ less than their accuracy representation, and at greater step the differentiation error increases.

References

1. *Gonzales R. C., Woods R. E.* Digital Image Processing (2nd Edition). — New Jersey: Prentice-Hall, Inc., 2002.

2. *Gostev I. M.* On Recognition Methods for Graphical Patterns // Journal of Computer and Systems Sciences International. — 2004. — No 1. — Pp. 138–144.
3. *Gostev I. M.* On the Identification of Unclosed Curves // Pattern Recognition and Image Analysis. — 2013. — Vol. 23, No 2. — Pp. 217–225.
4. *Березин И. С., Жидков Н. П.* Методы вычислений. Том 1. — М.: Физматлит, 1962. [Beresin I. S., Zhidkov N. P. Methods of Calculation. Vol. 1 — Moscow: Fizmatlit, 1962. — (in russian).]
5. *Abramowitz M., Stegun I.* Handbook of Mathematical Functions. — Washington: U.S National Bureau of Standards, 1964.
6. *Гельфанд И. М., Шиллов Г. Е.* Обобщённые функции. — М.: Физматлит, 1959. [Gelfand I. M., Shilov G. E. Generalized Functions: Textbook. — Moscow: Fizmatlit, 1959. — (in russian).]
7. *Lanczos C.* Applied Analysis. — New Jersey: Prentice Hall, Inc., 1956.
8. *Калиткин Н. Н.* Численные методы. — М.: Наука, 1978. [Kalitkin N. N. Numerical Methods. — Moscow: Nauka, 1978. — (in russian).]

УДК 519.683.8:519.6

Об особенностях вычисления производных высших порядков для идентификации формы графических объектов

И. М. Гостев

*Национальный исследовательский университет «Высшая школа экономики»
ул. Мясницкая, д. 20, Москва, Россия, 101000*

Рассмотрены методы вычисления производных высоких порядков на основе: интерполяционных формул; «безразностных методов вычисления производных»; применения свертки с заменой дифференцирования на операцию интегрирования; дифференцирования с использованием квадратур по Ланцошу; метода Нумерова. Проведён сравнительный анализ методов вычисления производных высоких порядков по точности вычислений с использованием в качестве эталона производных, вычисленных в пакете Maple с 20 разрядной десятичной точностью. Показано, что все методы практически эквивалентны по точности и сводятся к вычислению свертки между дифференцируемой функцией и некоторым окном, коэффициенты которого зависят от применяемого метода. Для проведения экспериментов разработан специальный программный комплекс для вычисления производных высоких порядков (до 7-го) табулированных функций с различным шагом. Были исследованы сетки с шагами от 0,005 до 0,1. Независимо от метода вычисления производных было определено, что оптимальным значением шага сетки для 64 разрядной арифметики является шаг от 0,01 до 0,05. При меньшем значении шага величины гладких функций различаются меньше чем их точность представления, а при большем возрастает погрешность дифференцирования. Результаты экспериментов подтверждают теоретические выводы Н. Н. Калиткина.

Ключевые слова: распознавания графических образов, обработка изображений, компьютерная геометрия, вычисление производных, точность вычислений, идентификация объектов, метрики линейной корреляции.