ISSN 2658-5014 (Print)
2022. Т. 4, № 2 (9). С. 24–29
СИИТ
СИСТЕМНАЯ ИНЖЕНЕРИЯ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ
ISSN 2686-7044 (Online)
http://siit.ugatu.su

УДК 519.6
doi 10.54708/26585014_2022_42924

THE PRACTICAL APPLICATION OF NUMERICAL FILTERING METHOD BY EXAMPLE OF CALCULATING THE SIMPLE FUNCTIONS DERIVATIVE

N. M. Sherykhalina [1], A. A. Sokolova [2], E. R. Shaymardanova [3]

[1] n_sher@mail.ru, [2] alexandrakrasich@gmail.com, [3] shaymardanova.ekaterina.04@gmail.com

Ufa State Aviation Technical University (UGATU)

Поступила в редакцию 8 июня 2022 г.

Abstract. There are a lot of computing systems and packages now, which are indispensable in solving of mathematical problems. Despite on the intensive using of numerical methods for modeling and designing of various systems, as well as the presence of a large number of mathematical software packages, the problem of computational errors estimate is very acute. A method of numerical results filtering for the solutions of different problems is presented for estimating the errors and increasing the accuracy. It's shown that the proposed method avoids the uncertainty and limitations of the Runge rule for estimating errors in numerical data. Using the example of calculating the value of the derivative of a simple function at a point, it was shown that after processing of the calculated values applying the proposed method, they can actually be refined to the reference ones in several iterations. The technique under consideration was previously used for complex modeling tasks, which led to a lack of understanding of its practical value. Therefore, in this work, the investigation was carried out using a simple function as an example, demonstrating the effectiveness of the proposed technique.

Key words: error estimation; mathematical modeling; numerical filtering; improving the reliability of calculations; numerical experiment; data processing; computational error; numerical results; extrapolation; the Runge rule.

INTRODUCTION

The problem of computing results reliability is under consideration of different authors [1–5]. The idea of numerical filtering application as a postprocessing of computed data was proposed by V. P. Zhitnikov and N. M. Sherykhalina [6]. The main idea of the method is constructing of mathematical model of an error in the form of sums of terms of some form and consecutive suppression of these terms. The idea has no strict mathematical proof, the technique is purely heuristic. However, its verification on many complex modeling problems has shown effective results [7–11]. The conducted investigations have shown that the method developed in [6] makes it possible to obtain reliable estimates of numerical results, and on the base of them to make practical conclusions about the simulated phenomena. Application of the technology of numerical results filtering allows making these conclusions with high accuracy.

Unfortunately, due to the complexity of the considered problems, the technique itself has been sidelined by the scientific community. In this paper, we propose to move away from complex models and focus on the investigation of the filering process on the example of calculating the values of elementary functions.

THE INITIAL DATA FOR FILTERING

So, let us consider a simple example of calculating the right derivative of the function cos(x) at the point x = 0.5. We will move to the point by halving the interval. The software implementation of the calculation under consideration is simple (the variables of type double are chosen):

```
for (int i = 0; i < N; i++)
{
h[i] = 1.0/n;
fright[i] = (f(x + h[i]) - f(x)) / h[i];
fcentral[i] = (f(x + h[i]) - f(x - h[i])) / (2 * h[i]);
n = n * 2.0;
}
```

We obtain a set of values of the right derivative fright[i] and central derivative fcentral[i] for n = 1, 2, 4, 8,...524248 (Table 1).

Table 1

Calculated values of the derivatives at a point

n	*fright*	*fcentral*
1	–0.80684536	–0.40342268
2	–0.674560512	–0.459697694
4	–0.583574772	–0.474447106
8	–0.532955539	–0.47817801
16	–0.506529003	–0.479113474
32	–0.493058623	–0.479347511
64	–0.486262005	–0.479406031
128	–0.482848701	–0.479420662
256	–0.481138346	–0.479424319
512	–0.480282248	–0.479425234
1024	–0.479853969	–0.479425462
2048	–0.479639773	–0.47942552
4096	–0.479532661	–0.479425534
8192	–0.479479101	–0.479425537
16384	–0.47945232	–0.479425538
32768	–0.479438929	–0.479425539
65536	–0.479432234	–0.479425539
131072	–0.479428886	–0.479425539
262144	–0.479427212	–0.479425539
524288	–0.479426376	–0.479425539

The standard value of cos(0.5) declared by well-known online calculators and math packages is cos(0.5)= –0.479425539. This value is observed starting from n = 16384 when the central derivative is considered and it is not observed when we calculate the right derivative. Thus, the questions arises: how many correct signs are present in the obtained result? Which of the many possible numerical values of the same parameter is used to further solving of the problem? By how much has the total calculation error accumulated due to used inaccurate values?

NUMERICAL FILTERING

The apparatus of the multicomponent analysis presented in [6] is a process that has been called the "filtering." In this context, the filtering is a set of algorithms and analytical rules that can be applied to sequences of calculated values of a required parameter. The main idea of filtering algorithms is to use the model of the calculated value error as a sum of several summands with unknown coefficients

$$b_n - b = c_1 n^{-k_1} + c_2 n^{-k_2} + \ldots + c_L n^{-k_L} + \Delta(n), \quad (1)$$

in this representation b_n is the approximate result (the values b_n for our case are presented in the second and third columns of Table 1); b is the exact value. The difference between the exact and approximate solution is the error and it is expressed by the right part of equality (1). c_j are the unknown coefficients; k_1,... k_L are arbitrary real numbers (the known ones), such that k_1<k_2<...<k_L. In most cases, $\Delta(n)$ is taken as an infinitely small quantity. However, we assume that the value $\Delta(n)$ has no a priori estimate. Moreover, it is possible for this value to increase as n increases (rounding off error, non-summed terms, residual term of the series). Also $\Delta(n)$ may be affected by imperfections in the numerical method itself, as well as by its hardware-software implementation. The last factor directly depends on the particular developer, the equipment and development tools he uses. Therefore, it is impossible to estimate in advance the magnitude of a possible error or deficiency. Consequently, we cannot assume in advance that $\Delta(n)$ is infinitely small quantity. The main task of filtering is sterwise removing of the power components of the sum (1). This paper considers filtering with multiplication of the number of nodes (reduction of the interval of derivative determination) as $n_i = 2n_{i-1}$.

The theoretical basis of the process is reflected in [12–15]. According to [9], in the current case the filterinf process coincides with the Richardson formula:

$$b_{n_i}^{(j)} = b_{n_i}^{(j-1)} + \frac{b_{n_i}^{(j-1)} - b_{n_{i-1}}^{(j-1)}}{Q^{k_j} - 1}, \; Q=2. \quad (2)$$

Thus we run the procedure over all pairs of values $b_{n_{i-1}}$, b_{n_i}. We obtain a dependence in which the value at each following filtering step j is expressed by the values obtained at the previous step of procedure *j*-1. Thus, we get a dependence that no longer contains a term with n^{-k_j}:

$$b_n^{(j)} = b + c_{j+1}^{(j)} n^{-k_{j+1}} + \ldots + c_L^{(j)} n^{-k_L} \Delta^{(j)}(n), \quad (3)$$

in the example under consideration, in practice, the resulting filtered sequences of values of the right derivative are as follows:

Table 2

Filtering results

n	*1 filtration*	*2 filtration*
1		
2	–0.542275664	
4	–0.492589032	–0.476026821
8	–0.482336306	–0.478918731
16	–0.480102466	–0.479357853
32	–0.479588243	–0.479416835
64	–0.479465388	–0.479424436
128	–0.479435397	–0.4794254
256	–0.47942799	–0.479425521
512	–0.47942615	–0.479425536
1024	–0.479425691	–0.479425538
2048	–0.479425577	–0.479425539
4096	–0.479425548	–0.479425539
8192	–0.479425541	–0.479425539
16384	–0.479425539	–0.479425539
32768	–0.479425539	–0.479425539
65536	–0.479425539	–0.479425539
131072	–0.479425539	–0.479425539
262144	–0.479425539	–0.479425539
524288	–0.479425539	–0.479425539

The *j* times filtered sequence contains one less term than $b_{n_i}^{(j-1)}$. The filtering operations can be sequentially repeated for $n^{-k_1} \ldots n^{-k_L}$, if the initial sequence contains a sufficient number of terms. Since the considered function is simple, some conclusions can already be done from the two new value sequences (Table 2). First of all, it is obvious that after the first filtering procedure the sequence (3) is already obtained, which is close to the reference value (even the first value, which was initially almost twice wrong). Secondly, after the second filtering procedure (j = 2), we obtain values, almost each of which can already be trusted if four decimal places are used. In general case, the numerical values obtained in this way should be subjected to analysis in order to estimate the error and justify the validity of these estimates.

This method of calculated values filtering has a number of limitations for its application. Its main limitation is the presence of an error component $\Delta(n)$, the value of which is, of course, unknown.

Of course, the ways to determine the error estimate already exist. For example, there is the Runge rule for this. However, this rule gives a significant result only if the summand that is removed at this stage is dominated in the considered model. And we do not deny that the last summand $\Delta(n)$ may also dominate over the others. In that case, the estimates by the Runge rule may turn out to be much smaller than the real ones.

We would like to note once more that filtering provides only additional information for the researcher in the form of numerical data sequences (in contrast to the methods for accelerating convergence). On the basis of the new regularities, further analysis of the obtained set of solutions b_n is carried out.

SIMULATION RESULTS

One of the convenient ways to analyze the obtained results of filtering is a graphical method (Fig. 1). Here $\lg \delta$ is decimal logarithm of absolute or relative error, $\lg n$ is decimal logarithm of discretization parameter (the number of partition intervals, for example). Thus we can show the accuracy expressed in the number of exact decimal digits. In this case, each component of dependence (1) is represented on such graph by a straight line segment.

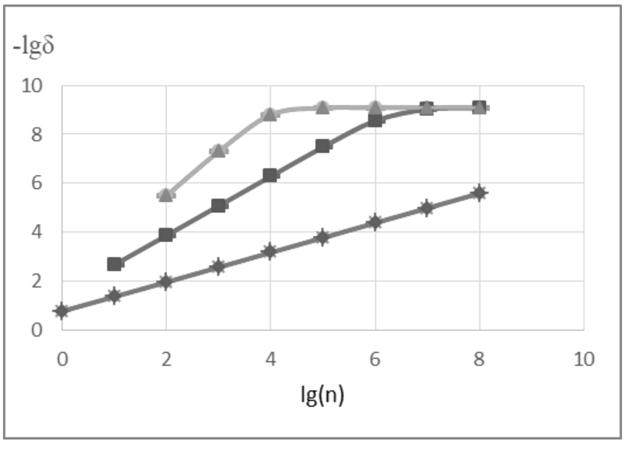

Fig. 1. Calculation error of the function cos'(x), x = 0.5

Below there are the developed rules that should be followed in the further analysis of the graph of the numerical data display.The upper line is used to estimate the fuzziness of the error estimate; the second line is used to estimate the error.

– We do not apply the upper line to estimate error. It visualizes the fuzziness of the error estimate. The second line from the upper one is used for the error estimate.

– The evidence of decreasing error with increasing n is that all lines of the graph are close to straight lines. If the lines bend upwards, it may be a signal that the error passes through the local minimum (or changes the sign). If this fact is detected, additional analysis should be performed.

– The density of points on the graph and the constancy of the sign of the error estimate also play a role. If the points are rarely located, there is a danger of confusing a straight line with one that is not a straight line. However, different methods can be used to increase the number of points without increasing the upper boundary of the number of nodes *n*.

Calculating the relative fuzziness of the estimate, it becomes obviously that performing 2 filtering stages (eliminating 2 components (1)) is sufficient to obtain an accuracy of 10^{-9} for 10 iterations.

APPLICATION OF FILTERING IN COMPLEX MODELS

In last considered problems, we used Schwarz integral instead of power series [2]. We have a general grid, and for each interval the integral is calculated using the two-point Gaussian formula which has the 4th order of accuracy with respect to the length of the integration segment. The Gaussian formula does not require calculating of the integrand on the boundaries of the interval, and this is important. For some integrals there is a singularity as 0/0 on the boundaries, which requires the ahhlication of the Lopital rules, and that leads to extra computation time. This is not needed when we use the Gaussian formula. But it is necessary to calculate repeatedly the values included in the integrand. In order to avoid repeated calculations, these values are calculated and memorized in advance. And then the recalculation with the memorized numbers takes place. Still, it is certainly more complicated and longer than calculating the partial sum of a series. But it gives an effect on the accuracy at the same dimensionality. One of the influencing factors is the ability to use irregular grids and the selection of the type of this irregularity. When we use a power series (on a circle it turns into Fourier series, i.e. periodic functions, which requires uniform grids). This is sometimes fatal. However, the singuliarities of developed numerical analytical methods require also calculation of integrals of functions with, for example, power fractional singuliarities. Substitution of the integration variable is inefficient because functions that depend on x in the first and other integer powers become power fractional functions. In this situation, filtering (in addition to irregular partitioning of the integration interval) helps a lot. So, there is a general grid, there is an additional partitioning of the interval closest to the singularity, and there is a calculation of the integral on each partial interval on its grid with decreasing step. The midpoint rectangles method is used because it does not require calculating the integrand on the boundaries of the interval of integration where there may be a singularity.

CONCLUSION

The considered process indirectly demonstrates the viability of the filtering idea. Obviously, the idea of refining values and estimating the error by filtering gives a qualitative improvement and refinement of the calculated values. Of course, for such simple examples, the practical usefulness of the approach is not obvious, since the values of the cosine derivative at the point 0.5 are known with high accuracy. However, for complex computational algorithms, the ability of reliable error estimate and refining the results is indispensable. The estimation results for such problems have been verified by obtaining the calculated more accurate values (using the data type of increased accuracy) [7]. In practice, it is not rational and not always possible to conduct calculations of numerical modeling problems using the data type of maximum accuracy.

Thus, the proposed approach makes it possible to obtain more accurate data and reliable estimates without resorting to complex computational processes using an over-precision data type, conducting many experiments on different sets of input values, and so on.

REFERENCES

1. **Amosov A. A., Dubinsky Y. A., Kopchenova N. V.** Computational methods. Moscow: MPEI Publishing House, 2008. 672 p.
2. **Bjoerck A., Dahlquist G.** Numerical mathematics and scientific computation. Vol. 1. 1999. 485 p.
3. **Richargson L. W.** The deferred approach to the limit // Phil. Trans. Roy. Soc. London. 1927. Vol. 226. Pp. 299-361.
4. **Paluri N. S. V., Sondur S.** Experiments with range computations using extrapolation // Reliable Computing. 2007. Vol. 13, No. 1. Pp. 1-23.
5. **Smith D. A., Ford W. F.** Numerical comparisons of non-linear convergence accelerations // Mathematics of Computation. 1982. Vol. 38, No. 158. Pp. 481-499.
6. **Sheykhalina N. M.** Mathematical modeling of technical objects and processes based on methods of multicomponent analysis of computational experiment results: Dr. Tech. Sci. diss. Ufa, 2012. 376 p.
7. **Sokolova A. A.** Cavitational shell flow by asymmetric Ryabushinsky scheme // Interuniversity Scientific Congress. Higher School: Scientific Research. 2019. Pp. 220-230.
8. **Zhitnikov V. P., Sheykhalina N. M., Sokolova A. A.** Marginal-quasi-stationary copying of circular electrode-tool // Vestnik UGATU. 2017. Vol. 21, No. 1 (75). Pp. 173-179.
9. **Zhitnikov V. P., Sheykhalina N. M., Sokolova A. A.** Problem of reliability justification of computation error estimates // Mediterranean Journal of Social Sciences. 2015. Vol. 6, No. 2. Pp. 65-78.
10. **Zhitnikov V. P., Sheykhalina N. M., Porechnyy S. S.** About one approach to practical assessment of errors in numerical results // SPbSPU. 2009. No. 3 (80). Pp. 105-110.
11. **Zhitnikov V. P., Sheykhalina N. M., Sokolova A. A.** Calculation program for the Stokes soliton parameters // Certificate of Official Registration of Computer Software of the Russian Federation No. 2018619757. Moscow: Rospatent, 2018.
12. **Zhitnikov V. P., Sheykhalina N. M.** Accuracy increase of complex problems solutions by numerical data post-processor handling // Computational technologies. 2008. Vol. 13, No. 6. Pp. 61-65.
13. **Sherykhalina N. M., Zhitnikov V. P**. Application of extrapolation methods of numerical results for improvement of hydrodynamics problem solution // Computational Fluid Dynamics Journ. 2001. Vol. 10, No. 3. Pp. 309-314.
14. **Zhitnikov V. P., Sheykhalina N. M., Sokolova A. A.** Estimation of error and its justification by filtering of numerical results obtained at different numbers of grid nodes // Izvestiya Samara Scientific Center RAS. 2017. Vol. 19, No. 1 (2). Pp. 401-405.
15. **Zhitnikov V. P., Sheykhalina N. M**. Methods of verification of mathematical models under uncertainty // Vestnik UGATU. 2000. Vol. 1, No. 2 (2). Pp. 53-60.

ABOUT AUTHORS

SHERYKHALINA Nataliya Mikhailovna, Prof., Dept. of Computational Mathematics and Cybernetics, Faculty of Computer science and robotics (USATU). Dipl. System engineer (USATU, 1993). Dr. of Tech. Sci. (USATU, 2012).

SOKOLOVA Aleksandra Alekseevna, Postgrad. (PhD) Student, Dept. of Computational Mathematics and Cybernetics, Faculty of Computer science and robotics (USATU). Dipl. Computer scientist-mathematician (USATU, 2014).

SHAYMARDANOVA Ekaterina Rinatovna, Student, Faculty of Computer science and robotics (USATU).

МЕТАДАННЫЕ

Заголовок: Практическое применение метода численной фильтрации на примере вычисления производной элементарной функции.

Авторы: Н. М. Шерыхалина [1], А. А. Соколова [2], Е. Р. Шаймарданова [3]

Принадлежность: Уфимский государственный авиационный технический университет (УГАТУ), Россия.

Эл. адрес: [1] n_sher@mail.ru, [2] alexandrakrasich@gmail.com, [3] shaymardanova.ekaterina.04@gmail.com

Язык: Английский.

Источник: СИИТ (научный журнал Уфимского государственного авиационного технического университета), т. 4, № 2 (9), стр. 24–29, 2022. ISSN 2686-7044 (онлайн), ISSN 2658-5014 (печатный вариант).

Аннотация: В настоящее время практически вся математика представлена в электронном виде в вычислительных системах и пакетах, которые являются незаменимыми в решении математических задач. Несмотря на интенсивное применение численных методов для моделирования и проектирования различных систем, а также наличие большого количества математических программных пакетов, проблема оценки вычислительных погрешностей стоит очень остро. В рассматриваемой работе представлен метод фильтрации численных результатов для оценки ошибок и повышения точности. Показано, что предложенный метод позволяет избежать неопределенности и ограничений правил Рунге для оценки ошибок численных данных. На примере вычисления значения производной простой функции в точке было показано, что после обработки вычисленных значений с помощью предложенного метода, их реально уточнить до эталонных за несколько итераций. Рассматриваемая методика ранее применялась для уточнения численных результатов при решении задач со сложными моделями и численно-аналитическими решениями, что привело к непониманию ее практической ценности. Поэтому в данной работе проведено исследование на примере простой функции, демонстрирующее эффективность предлагаемой методики.

Ключевые слова: оценка погрешности; математическое моделирование; численная фильтрация; надежность вычислительных результатов; численный эксперимент; обработка данных; погрешность вычислений; численные результаты; экстраполяция; правило Рунге.

Об авторах:

ШЕРЫХАЛИНА Наталия Михайловна, проф. каф. вычислительной математики и кибернетики. Дипл. инженер-системотехник (УГАТУ, 1993). Д-р техн. наук по мат. мод., числ. мет. и компл. программ (УГАТУ, 2012). Иссл. в обл. мат. мод. течений жидкости и электрохимического формообразования, численных методов и оценок погрешности.

СОКОЛОВА Александра Алексеевна, асс. каф. ВМиК. Дипл. информатик-математик (УГАТУ, 2014). Готовит дис. о мат. мод. течений жидкости и электрохимического формообразования с использованием методов численной фильтрации.

ШАЙМАРДАНОВА Екатерина Ринатовна, студент бакалавриата по направлению Математическое обеспечение и администрирование информационных систем ФИРТ УГАТУ.