The Interval Method of Bisection for Solving the Nonlinear Equations with Interval-Valued Parameters

Konstantin Semeno[v](https://orcid.org/0000-0003-4292-8053) [a](https://orcid.org/0000-0002-5440-257X)nd Anastasia Tselishcheva

Abstract The article dwells on the interval extension of the bisection approach for solving nonlinear equations with interval-valued parameters, i.e. the ones that might have values from the specified bounds. It is shown that such a procedure allows to obtain an interval of possible values for equation root that is entirely determined by the equation parameters inaccuracy and does not depend on any other factor. The proposed interval bisection method can be easily implemented. All the differences from the traditional bisection approach for solving equations have a clear meaning. The simple stopping rule is proposed. It is shown that considering the interval nature of equation parameters makes it possible to finish the iterative process of equation solving earlier in full accordance with known information on the equation parameters. The proposed approach keeps the important bisection method property—all the intermediate estimates of the bounds of the root's possible values interval include the exact boundaries. The article provides an illustrative example of how to use the interval bisection.

Keywords Inaccurate data · Nonlinear equations · Equation solving · Interval bisection

373

K. Semenov (\boxtimes) · A. Tselishcheva

Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia e-mail: semenov.k.k@iit.icc.spbstu.ru

[©] The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021 N. Voinov et al. (eds.), *Proceedings of International Scientific Conference on Telecommunications, Computing and Control*, Smart Innovation, Systems and Technologies 220, https://doi.org/10.1007/978-981-33-6632-9_33

1 Introduction

Many problems in data processing require solving nonlinear equations to obtain the desired result: parametrical identification, fitting data using a specific model, some cases of indirect measurements, etc. In most cases, the data to be processed aren't accurate and, consequently, the equations to be solved have uncertain parameters. That's the reason why the roots of such equations cannot be found precise: we always will face the residual uncertainty inherited from the inaccurateness of equation parameters. So, the trials to interpret the roots estimates as quantities, which values are only corrupted with round-off errors and errors caused by using an iterative numerical algorithm of equation solving, aren't correct—if we solve the required equation analytically, calculate the roots directly using corresponding expressions and don't allow round-off errors to occur, then it still does not make the root estimate absolutely accurate. If we don't consider this circumstance, then we overestimate the preciseness of our knowledge on obtained results which in turn leads to an increase in the risks of making erroneous decisions in the future. As en example, we can consider the root of the function $f(x, a, b) = a \cdot x + b$, where values of *a* and *b* aren't known accurately. If we know that the value of *a* is inside the interval *Ia* $=[1; 2]$ and the value of *b* is inside the interval $I_b = [-2; -1]$ then the root of the function *f* is inside $I_x = [0.5; 2]$. Applying any numerical method to solve the equation $f(x, a, b) = 0$ considering the uncertainty of *a* and *b* will bring us to the interval wider than the mentioned interval I_x . So, these bounds for the equation root's possible value are the accurate limits that can be reached for some combination of values *a* from I_a and *b* from I_b .

This paper discusses the interval bisection method for solving nonlinear equations in which parameters are known approximately—in practice, usually, all we know is the interval of possible values of these parameters.

2 Solving Nonlinear Equations for Indirect Measurements

All the data that we encounter in real-world conditions are uncertain—from world constants to measurement results. In some cases, we can neglect this uncertainty in our calculations and reasoning, but, in other cases, we cannot because of the big price we will pay: decisions made without taking into account the uncertainty of source data may be ill-founded. The absolute accurateness is the distortion of reality, and, in practice, we always need to know the quality of our results. The natural scientific area, in which there are the regulatory requirements to accompany each result with individual characteristics of its uncertainty, is metrology and science on measurements. So, the most relevant example of applications where we deal with equations with inaccurate parameters is the case of indirect measurements when we calculate the value of interest from the measurement results of the related quantities connected with it with the known dependence. Without loss of generality, we can

consider the problem of solving indirect measurements equations to illustrate the approach proposed and describe details.

Many quantities cannot be measured directly for various reasons. In these cases, indirect measurements can be used. Firstly, the mathematical model is constructed and tested that describes the interconnection between quantities measured directly and values to be found out. After performing all corresponded measurements, the necessary values are computed as the root of the equation or the solution of the equations system relating the participating quantities. The calculation of the result is an essential part of the indirect measurements and should be taken into account during metrological characteristics estimation. It should be metrologically verified like any other measurement procedure or conversion. Indeed, as it was mentioned earlier, all results of calculations with uncertain data are always inaccurate too—the uncertainty inherited from the input data cannot be overcome. The only thing that can be performed is to analyze how the uncertainty transforms during the computations and to estimate the value of the final calculations results error.

To date, there are many approaches and methods to support metrologically computations—including solving equations with inaccurate parameters (usually being direct measurement results). These methods can be grouped into two big sets: methods that use randomization (Monte Carlo approach [\[1\]](#page-10-0), the Cauchy deviate method [\[2,](#page-10-1) [3\]](#page-10-2) and related techniques) and methods that perform automatic analysis of the computational algorithm by overloading the operations performed during calculations—assuming a linear approximation (automatic differentiation of first order [\[4,](#page-10-3) [5\]](#page-10-4) as the most valuable approach, finite differences and complex step derivatives estimates [\[6,](#page-10-5) [7\]](#page-10-6) and similar techniques for sensitivity analysis) and in the common case (interval arithmetic [\[8,](#page-10-7) [9\]](#page-10-8) and its extensions and modifications like affine arithmetic [\[10,](#page-11-0) [11\]](#page-11-1) or others [\[12,](#page-11-2) [13\]](#page-11-3) joined with random variables processing approaches like probability boxes framework $[14, 15]$ $[14, 15]$ $[14, 15]$). All of these techniques were developed for the wide class of computational problems with inaccurate input data and can be used in computational metrology.

Numerous different methods are developed and used to search the roots of nonlinear equations. The most popular of them are Newton and Newton–Raphson methods, secant method, bisection method, etc [\[16\]](#page-11-6). The first three methods need the initial guess as a first estimation of the root, and the last-mentioned one needs the interval of root localization. If the initial guess is unsuccessful, then the iterative process can diverge, and no root can be obtained at all. In many cases, it is impossible to determine if the guess is acceptable or not before the iterative process starts. In constrast, the bisection method guarantees that the final result of root estimation will be got and supposes a very simple procedure to test if the initial localization interval is acceptable or not.

We should consider the additional sides of the issue to determine what method for solving equations from the listed above is better for metrological practice. So for this, we analyze the details of the metrological supporting the corresponding computational procedures. Approaches for processing initial data uncertainty using randomization might cause the situation when the Newton, Newton–Raphson, and secant methods will diverge—so, we should recognize corresponded iterations during

the Monte Carlo method execution or similar approaches applying and should stop timely and throw out the wrong results from consideration. This complicates the procedure of root finding. The same situation may occur if we use operator overloading—the procedure that converges being executed with individual numbers as input variables may begin to diverge if it is executed with intervals or interval-valued quantities like probability boxes. Thus, this can bring us to an unacceptable situation if the equation to be solved is the equation of indirect measurements—we will not obtain any measurement result at all because of the computational procedure. To address these shortcomings, we suggest using the bisection method that always ensures the final result obtaining. Besides, supporting the bisection with one of the discussed approaches for estimating uncertainty inherited from the initial data doesn't bring us to the iterative process divergence. So, combining the bisection method with any kind of procedure of metrological supporting is the preferable way to solve nonlinear equations of indirect measurements.

This paper presents the interval version of the bisection method for solving nonlinear equations with interval-valued parameters that are commonplace in metrological practice. The proposed method is fully in line with metrological requirements that is an advantage in comparison with traditionally used approaches [\[17,](#page-11-7) [18\]](#page-11-8). The way is proposed for taking into account the uncertainty of initial data during equation solving and reasonably set the moment to stop the iteration process.

3 The Interval Bisection

Let $\vec{x}^T = (x_1, x_2, x_3, ..., x_n)$ be the direct measurement results, and $f(y, \vec{x}) = 0$ be the equation that connects these measurands with quantity *y* that is supposed to be measured indirectly. Let Δx_1 , Δx_2 , Δx_3 , ..., Δx_n be the absolute errors of $x_1, x_2, x_3, \ldots, x_n$ correspondingly, and let it be known that their maximum possible values satisfy the restrictions: $|\Delta x_i| < \Delta_i$, $i = 1, 2, ...n$.

The traditional bisection method $[19]$ ignores that quantities x_1 , x_2 , x_3 , ..., x_n are inaccurate and treats them as the only possible values of parameters of the equation to be solved. Let the interval $I_1 = [a, b]$ be the localization bounds for *y*. In metrology during indirect measurements, we usually face equations representing zeros of the monotonic functions *f*. So we have the only root because, for one set of direct measurement results, we must have the only one corresponding value of the indirect measurement. If the values $f(a, \vec{x})$ and $f(b, \vec{x})$ have different signs, then interval *I* will contain the only root. For problems from other fields, we should start with such an interval of values of argument *y* that will provide different signs of function *f* values obtained at the interval's left and right bounds. Then, we can be sure that not less than one root is inside this interval.

For each step of bisection, the current interval of root localization is divided into two equal parts, and the one that contains the root is chosen. To determine what half should be preferred, the sign of value $f(0.5 \cdot (a + b), \vec{x})$ in the middle of interval *I* should be calculated. The obtained narrowed interval is new localization bounds

for equation root, and then the new iteration starts, and all described operations are repeated. So, let the localization interval for *i*-th iteration step be $I_i = [a_i, b_i]$. If $f(a_i, \vec{x}) \cdot f(0.5 \cdot (a_i + b_i), \vec{x}) > 0$, then $a_{i+1} = 0.5 \cdot (a_i + b_i), b_{i+1} = b_i$. If $f(0.5 \cdot (a_i + b_i), \vec{x}) \cdot f(b_i, \vec{x}) > 0$, then $a_{i+1} = a_i, b_{i+1} = 0.5 \cdot (a_i + b_i)$. The interval $I_{i+1} = [a_{i+1}, b_{i+1}]$ is the localization interval for the next iteration.

The situation stops to be unambiguous if taking into account the information on the uncertainty of initial data. Since some iteration, we will not be able to determine exactly the sign of the value $f(0.5 \cdot (a_i + b_i), \vec{x})$ because of the influence of uncertainty of direct measurement results \vec{x} acting as equality parameters. So, we will not be able to decide what half of the current localization interval contains the root of the equation to be solved: for some possible values of \vec{x} , it will be in the left half, and for other possible values – in the right half.

The solution allowing to overcome this obstacle is to use one of the methods discussed in the previous section of the paper that provides each calculation of the function $f(y, \vec{x})$ with its individual uncertainty estimate $\Delta f(y, \vec{x})$. Then we will be able to determine the moment when the traditional bisection method faces at current iteration *i* such center $c_i = 0.5 \cdot (a_i + b_i)$ of the current root localization interval $I_i =$ $[a_i, b_i]$ that satisfies the condition $\Delta f(c_i, \bar{x}) > |f(c_i, \bar{x})|$. This inequality indicates the situation described above when we cannot choose half of the localization interval for the next bisection iteration. Really, if it holds, then there are no reasons to consider the value $f(c_i, \vec{x})$ differing from zero. The equivalent form of the inequality is $0 \in (f(c_i, \vec{x}) \pm \Delta f(c_i, \vec{x}))$, so we see that any positive or negative values $f(c_i, \vec{x})$ lying inside the interval determining by mentioned inequality could be formed by distorting the true value equal to zero by measurement errors. In this paper, we suggest using the moment when the analyzed inequality holds as the transition to the second stage of the modified bisection method.

So, the following simple algorithm can describe the first stage of the proposed approach.

> *Input:* $f(y, \vec{x}) = 0$ // the equation to be solved $I_0 = [a_0, b_0]$ // the initial root localization interval $\frac{1}{2}$ assign the value to the iteration index $Do:$ $i := 1:$ $c_i = 0.5 \cdot (a_i + b_i)$; // calculate the center of localization interval While $\Delta f(c_i, \vec{x}) < |f(c_i, \vec{x})|$ // the condition to end the method first stage // choose the half of localization interval for next iteration If $f(a_i, \vec{x}) \cdot f(c_i, \vec{x}) < 0$ Then $a_{i+1} := a_i$; $b_{i+1} := c_i$; Else $a_{i+1} := c_i$; $b_{i+1} := b_i$; End $I_{i+1} = [a_{i+1}, b_{i+1}]$; // the new root localization interval $c_{i+1} := 0.5 \cdot (a_{i+1} + b_{i+1});$ // its center $i := i + 1$; $\frac{1}{2}$ increment the iteration index End

> I_i // the interval of root localization obtained at the last iteration Output:

On the second stage of the interval bisection, we need to narrower the last obtained interval $I_i = [a_i, b_i]$ of root localization that provides the unambiguous sign of function *f* at its bounds:

$$
\Delta f(a_i, \vec{x}) < |f(a_i, \vec{x})| \text{ and } \Delta f(b_i, \vec{x}) < |f(b_i, \vec{x})|
$$

The goal of each iteration of the second stage of the proposed method is to narrower these bounds to such an interval $I_{i+1} = [a_{i+1}, b_{i+1}]$ that ensures holding the condition $I_{i+1} \subseteq I_i$ and guarantees at the same time that the sign of function $f(y, \bar{x})$ at $y = a_{i+1}$ and $y = b_{i+1}$ isn't ambiguous:

$$
\Delta f(a_{i+1}, \vec{x}) < |f(a_{i+1}, \vec{x})| \text{ and } \Delta f(b_{i+1}, \vec{x}) < |f(b_{i+1}, \vec{x})|.
$$

Surprisingly, the traditional bisection approach can be easily applied for this purpose. We can reformulate the problem to be solved in the following manner:

- to find the root's minimum possible value, we need to solve equation $\Delta f(y_{\text{min}}, \vec{x}) = |f(y_{\text{min}}, \vec{x})|$ for y_{min} within the localization interval [a_i, c_i];
- to find the root's maximum possible value, we need to solve equation $\Delta f(y_{\text{max}}, \vec{x}) = |f(y_{\text{max}}, \vec{x})|$ for y_{max} within the localization interval [c_i, b_i].

Here, as before, $c_i := 0.5 \cdot (a_i + b_i)$ is the center of interval I_i that is obtained on the last iteration of the first stage of interval bisection.

Thus, at every new iteration, we need to examine the left and right bound of the interval that localizes the equation root separately. To finish the iterative process, we propose the following rule. It is rational to stop improving the interval estimating when the interval length refining on the next iteration is less than the given constant $ε > 0$:

$$
||I_i||-||I_{i+1}||<\varepsilon.
$$

Solving the metrological problems, the uncertainty bounds for the obtained root's value should be rounded – this circumstance is the natural opportunity to determine the best moment to stop the interval bisection method. If the rounded bounds of the interval of root localization obtained on the previous iteration are the same as the rounded bounds of the interval of root localization obtained on the current iteration, then we should finish. The rounding is suggested to be performed in a metrological sense.

The algorithm of the second stage of the interval bisection is the following.

 $f(y, \vec{x}) = 0$ // the equation to be solved Input: $I_i = [a_i, b_i]$ // the interval of root localization that was obtained \mathcal{N} on the last iteration i of the method's first stage // user-defined positive constant ϵ $Do:$ $c_i := 0.5 \cdot (a_i + b_i);$ // calculate the center of localization interval I_i $a'_i := a_i$; $b'_i := c_i$; // the interval that localizes y_{\min} $a''_i := c_i$; $b''_i := b_i$; // the interval that localizes y_{max} $\mathbb{D}o$ // refining of the left border of the root localization area $c'_i := 0.5 \cdot (a'_i + b'_i)$. // the center of the interval that localizes y_{\min} $\Delta f(c_i', \vec{x}) < |f(c_i', \vec{x})|$ H° $a'_{i+1} := c'_{i}; b'_{i+1} := b'_{i};$ **Then** $a'_{i+1} := a'_{i}; b'_{i+1} := c'_{i};$ Else **Fnd** ℓ refining of the right border of the root localization area $c_i'' = 0.5 \cdot (a_i'' + b_i'')$. // the center of the interval that localizes y_{max} $\Delta f(c_i'',\vec{x}) < |f(c_i'',\vec{x})|$ H *Then* $a_{i+1}'' := a_i''$, $b_{i+1}'' := c_i''$, $a''_{i+1} := c''_i$; $b''_{i+1} := b''_i$; Else End $I_{i+1} = [a'_{i+1}, b''_{i+1}]$; // the new root localization interval $i := i + 1$; $\frac{1}{2}$ increment the iteration index While $(b_{i-1}''-a_{i-1}')-(b_i''-a_i')>\varepsilon$ // condition: $||I_{i-1}||-||I_i||<\varepsilon$ *Output*: I_i // the interval of root localization obtained at the last iteration

4 Illustrating Example

To make the proposed ideas of the interval bisection method clearer, let us examine some function $f(y, \vec{x}) = \exp(x_1 \cdot y) + x_2 \cdot y$ depending on the variable of our interest *y* and a set of parameters \vec{x} that are known with uncertainty. Let us find the root of the equation $f(y, \vec{x}) = 0$ using the discussed approach.

From the physical sense, this equation models the environmental pollution caused by the point source. The parameters \vec{x} describe the characteristics of the environment and the pollution. From the mathematical viewpoint, this problem is equivalent to calculating the standard Lambert W-function [\[20\]](#page-11-10).

The values $\vec{x}^{\text{T}} = (x_1, x_2)$ aren't known exactly. All we know about values \vec{x} is that *x*₁ ∈ [−0.11, −0.09] and *x*₂ ∈ [−5.05, −4.95]. So, $(x_1, x_2) = (-0.10, -5.00)$ and $(\Delta x_1, \Delta x_2) = (0.01, 0.05)$. Let us try as the start root's localization interval the interval $I_0 = [a_0, b_0] = [0, 1]$. These bounds satisfy the condition of the bisection method applicability condition: $f(a_0, \vec{x}) \cdot f(b_0, \vec{x}) < 0$.

The entire iterative process that corresponds to using the interval bisection for the mentioned equation is presented in Fig. [1.](#page-7-0) We can see that, as a result, we obtain further unimprovable interval estimation of the root that cannot be narrower because

Fig. 1 Intervals of root localization for different stages of the proposed algorithm

Fig. 2 Illustration of the first stage of the interval bisection algorithm

of the uncertainty of the solved equation parameters. Figure [1](#page-7-0) also illustrates that, during the first stage of the interval bisection method, this approach reproduces the traditional scheme of the bisection and that the second stage essence is in narrowing the last localization interval obtained at the first stage.

Fig. 3 Results of the first stage of the interval bisection algorithm and the transition to the second stage

In Fig. [2,](#page-7-1) the results of the first stage execution of the proposed method are illustrated. The stop condition is satisfied on the 4th iteration when we cannot, for the first time, determine the sign of the function *f* in the center of the root localization interval. So, we go to the second part of the method.

The results obtained on the several iterations of the second stage of interval bisection are illustrated in Figs. [3,](#page-8-0) [5,](#page-9-0) and [6.](#page-10-9) We can see how the left and right bounds of the localization interval are refined. For convenience, in Figs. [3,](#page-8-0) [5,](#page-9-0) and [6,](#page-10-9) the independent indexing of iterations is used: index $j = 0$ corresponds to the beginning of the second stage of interval bisection when dealing with the localization interval obtained on the last iteration of the method's first stage.

In Fig. [6,](#page-10-9) we see the final iteration of the proposed approach. It corresponds to the stopping rule taken from the metrological nature of the solving problem: if we round the uncertainty bounds of the root's estimate at the current iteration, then the new iteration will not bring the refining, and we should finish. The obtained bounds are [0.172, 0.219].

5 Conclusions

In this paper, the interval extension of the bisection method is proposed for solving nonlinear equations with inaccurate parameters. A simple and effective algorithm is presented that brings with the guarantee to the root estimation. The clear stopping rules are proposed that naturally follow from the problem and allow to finish the

Fig. 4 Results of the first iteration of the second stage of the interval bisection algorithm

Fig. 5 Results of the second iteration of the second stage of the interval bisection algorithm

iterative process of equation solving earlier in full correspondence with known initial data on the equation to be solved. The proposed approach remains the important property of the bisection method—all the intermediate interval estimates of the root possible values contain the exact bounds.

Fig. 6 Final results of the interval bisection algorithm

Acknowledgements The reported study was funded by the Russian Foundation for Basic Research (RFBR), project number 19-31-90165.

References

- 1. Kreinovich, V., Beck, J., Ferregut, C., Sanchez, A., Keller, G.R., Averill, M., Starks, S.A.: Monte-Carlo-type techniques for processing interval uncertainty, and their potential engineering applications. Reliable Comput. **13**(1), 25–69 (2007)
- 2. Kreinovich, V., Nguyen, H.T.: Towards intuitive understanding of the cauchy deviate method for processing interval and fuzzy uncertainty. In: IFSA World Congress/EUSFLAT Conference (IFSA-EUSFLAT 2009), pp. 2336–2341. Lisbon, Portugal (2009)
- 3. Kreinovich, V., Ferson, S.A.: A new Cauchy-based black-box technique for uncertainty in risk analysis. Reliab. Eng. Syst. Saf. **85**(1–3), 267–279 (2004)
- 4. Griewank, A.: On automatic differentiation. In: Iri, M., Tanabe, K. (eds.) Mathematical Programming: recent developments and applications, pp. 83–107. Kluwer Academic, Dordrecht, the Netherlands (1989)
- 5. Bücker, H.M., Corliss, G., Hovland, P., Naumann, U., Norris, B. (eds.): Automatic differentiation: applications, theory, and implementations. 50. Springer Science & Business Media (2006)
- 6. Martins, J.R.R.A., Sturdza, P., Alonso, J.J.: The complex-step derivative approximation. ACM Trans. Math. Software (TOMS) **29**(3), 245–262 (2003)
- 7. Lai, K.L., Crassidis, J.: Generalizations of the complex-step derivative approximation. In: AIAA Guidance, Navigation, and Control Conference and Exhibit, paper 6348 (2006)
- 8. Moore, R.E., Kearfott, R.B., Cloud, M.J.: Introduction to Interval Analysis. Siam, Philadelphia (2009)
- 9. Hickey, T., Ju, Q., van Emden, M.H.: Interval arithmetic: from principles to implementation. J. ACM **48**(5), 1038–1068 (2001)
- 10. de Figueiredo, L.H., Stolfi, J.: Affine arithmetic: concepts and applications. Numer. Algorithms **37**(1–4), 147–158 (2004)
- 11. Shou, H., Lin, H., Martin, R., Wang, G.: Modified affine arithmetic is more accurate than centered interval arithmetic or affine arithmetic. In: Mathematics of Surfaces, pp. 355–365. Springer, Berlin, Heidelberg (2003)
- 12. Nedialkov, N.S., Kreinovich, V., Starks, S.A.: Interval arithmetic, affine arithmetic, Taylor series methods: why, what next? Numer. Algorithms **37**(1–4), 325–336 (2004)
- 13. Barrio, R., Rodríguez, M., Abad, A., Blesa, F.: Breaking the limits: the Taylor series method. Appl. Math. Comput. **217**(20), 7940–7954 (2011)
- 14. Ferson, S., Kreinovich, V., Grinzburg, L., Myers, D., Sentz, K.: Constructing probability boxes and dempster-shafer structures. Sandia National Lab (SNL-NM), Albuquerque, NM (United States). Report No. SAND-2015-4166J (2015)
- 15. Troffaes, M.C.M., Miranda, E., Destercke, S.: On the connection between probability boxes and possibility measures. Inf. Sci. **224**, 88–108 (2013)
- 16. Kelley, C.T.: Solving nonlinear equations with Newton's method. SIAM, Philadelphia (2003)
- 17. Semenov, K.K., Tselishcheva, A.A.: Interval method of bisection for a metrologically based search for the roots of equations with inaccurately specified initial data. Meas. Tech. **61**(3), 203–209 (2018)
- 18. Semenov, K.K., Tselishcheva, A.A.: Generalized interval method of bisection for metrologically based search for solutions of systems of equations with inaccurately specified initial data. Meas. Tech. **62**(3), 193–201 (2019)
- 19. Wu, X.: Improved Muller method and bisection method with global and asymptotic superlinear convergence of both point and interval for solving nonlinear equations. Appl. Math. Comput. **166**(2), 299–311 (2005)
- 20. Corless, R.M., Gonnet, G.H., Hare, D.E., Jeffrey, D.J., Knuth, D.E.: On the LambertW function. Adv. Comput. Math. **5**(1), 329–359 (1996)