6-1-2024

What to Do If an Inflexible Tolerance Problem Has No Solutions: Probabilistic Justification of Piegat's Semi-Heuristic Idea

Olga Kosheleva
The University of Texas at El Paso, olgak@utep.edu

Vladik Kreinovich
The University of Texas at El Paso, vladik@utep.edu

Comments:

Technical Report: UTEP-CS-24-29

What to Do If an Inflexible Tolerance Problem Has No Solutions: Probabilistic Justification of Piegat's Semi-Heuristic Idea

Olga Kosheleva and Vladik Kreinovich

Abstract In many practical situations, it is desirable to select the control parameters $x_1, \ldots, x_n$ in such a way that the resulting quantities $y_1, \ldots, y_m$ of the system lie within desired ranges. In such situations, we usually know the general formulas describing the dependence of y_i on x_j, but the coefficients of these formulas are usually only known with interval uncertainty. In such a situation, we want to find the tuples for which all y_i's are in the desired intervals for all possible tuples of coefficients. But what if no such parameters are possible? Since we cannot guarantee the inclusions with probability 1, a natural idea is to select parameters for which the probability that all inclusions are satisfied is the largest. To implement this idea, we need to select a probability distribution on the set of all tuples. Since we have no reason to believe that some tuples are more probable than others, it is reasonable to assume that all tuples are equally probable, i.e., that we have a uniform distribution on the set of all tuples. Interestingly, this idea leads to the same recommendation as was proposed – based on heuristic fuzzy-logic-based arguments – in a recent paper by Piegat. An important remaining open problem is how to efficiently compute the recommended solution.

1 Formulation of the Problem

What is tolerance solution and why do we need it. In many practical situation, we need to make sure that the values of certain quantities $y_1, \ldots, y_m$ lies within the desired ranges $\mathbf{y}_1 = [\underline{y}_1, \overline{y}_1], \ldots, \mathbf{y}_m = [\underline{y}_m, \overline{y}_m]$. For example, for a chemical reactor:

Olga Kosheleva
Department of Teacher Education, University of Texas at El Paso, 500 W. University
El Paso, Texas 79968, USA, e-mail: olgak@utep.edu

Vladik Kreinovich
Department of Computer Science, University of Texas at El Paso, 500 W. University
El Paso, Texas 79968, USA, e-mail: vladik@utep.edu

- the temperature y_1 cannot be too small – then there will be no reaction, and it cannot be too high – then the reactor itself may break;
- the amount y_2 of chemical leaking into the environment should be below a desired threshold, etc.

Similarly, a medical doctor would like to make sure that all the patient's quantities – blood pressure, body temperature, blood sugar level, etc. – are within the norm.

We usually have several parameters $x_1,\ldots,x_n$ that we can set up to control the system, and we know the form of the dependence of the desired quantities y_i on these parameters: namely, we know that this dependence has the form $y_i = f_i(x_1,\ldots,x_n,c_1,\ldots,c_\ell)$ for some coefficients c_i. For example, we may know that the dependence is linear, so that $y_i = c_{i0}+c_{i1}\cdot x+\ldots+c_{in}\cdot x_n$. Which values of the parameters $x_1,\ldots,x_n$ should we choose?

In the idealized situation, when we know the exact values of the coefficients $c_1,\ldots,c_\ell$, we can simply select the values x_j that satisfy m inequalities

$$\underline{y}_1 \le f_i(x_1,\ldots,x_n,c_1,\ldots,c_\ell) \le \overline{y}_i, \quad i=1,\ldots,m. \tag{1}$$

In particular, if instead of the bounds $\underline{y}_i$ and $\overline{y}_i$, we know the exact values y_i that we want to maintain, then we need to solve the system of equations $f_i(x_1,\ldots,x_n,c_1,\ldots,c_\ell) = y_i$.

In practice, however, we usually know the coefficients only with uncertainty. Usually, we only know bounds lower and upper bound $\underline{c}_k$ and $\overline{c}_k$ on each of these coefficients, i.e., we only know the intervals $\mathbf{c}_k = [\underline{c}_k,\overline{c}_k]$ that contain each coefficient c_k [2, 6, 7, 9, 13]. In this case, we need to find the values $x_1,\ldots,x_k$ for which all the inequalities (1) are satisfied for all possible combination of values $c_k \in [\underline{c}_k,\overline{c}_k]$. The set of all such tuples $x = (x_1,\ldots,x_n)$ is known as the *tolerance solution* to the problem.

But what if there is no tolerance solution? In many cases, there are values $x_1,\ldots,x_n$ satisfying the desired conditions. However, sometimes, there are none. Let us give a simple example: $m=n=k=1$, $f(x_1,c_1) = c_1\cdot x_1$, $\mathbf{c}_1 = [2,4]$, and $\mathbf{y_1} = [4,6]$. In this case, no matter what value x_1 we choose, the set of possible values of $c_1\cdot x_1$ will be equal to the interval $[2x_1,4x_1]$. We need both endpoints of the interval $[2x_1,4x_1]$ to be inside the desired interval $[4,6]$. In particular, we need to have $2x_1 \ge 4$ and $4x_1 \le 6$. But the first of these two inequalities implies $x_1 \ge 2$ while the second implies $x_1 \le 1.5$. Clearly, these two inequalities cannot be both true.

What shall we do in this case?

Constraints can be flexible or inflexible. In some cases, the constraints are flexible. For example, we may impose limitations on the cost of the bridge design, on its longevity, etc. If it turns out that we cannot satisfy all these constraints, we need to adjust them: either increase the budget or decrease the longevity requirement.

In this case, one of the reasonable suggestions it to find the values x_j that corresponds to the smallest possible extensions of the ranges that allow a solution. This problem is described and solved in [14, 15].

However, sometimes, the constraints are not flexible. For example, suppose that we set up a reactor in an emergency situation. In this case, even if we cannot guarantee that it will work under all circumstances, it is worth trying – and it is definitely better to try something and hopefully succeed than to do nothing. Similar in the medical case: yes, there may be a possibility that the medicine will not work, it happens, but it does not mean we should not try, the only question is which of the possible medicines we should try. If a medical doctor cannot find an ideal combination of medicines that would guarantee to make the patient perfectly healthy, what is the next best choice?

What we do in this paper. In this paper, we provide a reasonably natural probability-based suggestion on what to do if a given inflexible tolerance problem has no solutions. Interestingly, this probability-based suggestion is in perfect agreement with a heuristic fuzzy-motivated solution to the problem proposed in a recent paper [12].

2 Our solution: motivation, the resulting formula, its relation to Piegat's solution, and remaining open problems

Natural idea. If we cannot guarantee that all the constraints are satisfied with probability 1, the natural next best idea is to select a tuple for which the probability of satisfying all the constraint is the largest possible.

But where do we get the probabilities? The problem with this idea is that we do not know the probabilities of different values of c_k, all we know is that each of these values belongs to the corresponding interval $\mathbf{c}_k$. We do not have any information about the probability of different tuples $c = (c_1, \ldots, c_\ell)$. So what shall we do?

Let us use Laplace Indeterminacy Principle. Let us recall our problem. We do not know the probabilities of different tuples $c = (c_1, \ldots, c_\ell)$ from the given box

$$\mathbf{c}_1 \times \ldots \times \mathbf{c}_\ell.$$

There are many possible probability distributions on this box. To apply the above idea, we need to select one of these distributions. Which one should we select?

We have no reason to believe that some tuples are more probable than others. It is therefore reasonable to select the distribution in which all the tuples are equally probable – i.e., to select the uniform distribution on this box. This makes perfect sense: e.g., if we have several suspects, and we have no reason to believe that some are more probably than other, then it is natural to consider them equally probable. This argument goes back to Pierre-Simon Laplace, one of the founders of probability theory. It is thus known as *Laplace Indeterminacy Principle*; see, e.g., [3].

So what should we recommend. We want to select the tuple x maximizing the probability that for a given tuple c, all the desired inequalities (1) are satisfied, i.e., the probability that the tuple c belongs to the set S_x of all the tuples for which all inequalities are satisfied:

$$S_x \stackrel{\text{def}}{=} \{c : \underline{y}_1 \le f_i(x,c) \le \overline{y}_i \text{ for all } i\}. \tag{2}$$

For the uniform distribution, the probability that a tuple belongs to some set S is proportional to the volume $V(S)$ of this set. Thus, our recommendation is: *to select the tuple x for which the volume $V(S_x)$ is the largest.*

How is this related to Piegat's solution. In the specific case when $n = m = k = 1$, $\mathbf{c}_1 = [\underline{c}_1, \overline{c}_1]$ with $\underline{c}_1 > 0$, $\mathbf{y}_1 = [\underline{y}_1, \overline{y}_1]$ with $\underline{y}_1 \ge 0$, and $f(x_1, c_1) = c_1 \cdot x_1$, the above recommendation means selecting the value x_1 for which the width (1-D volume) of the interval $S_{x_1} = \{c_1 \in \mathbf{c}_1 : \underline{y}_1 \le c_1 \cdot x_1 \le \overline{y}_1\}$ is the largest possible. This idea was proposed in a recent paper [12] based on some semi-heuristic ideas related to fuzzy logic (see e.g., [1, 4, 8, 10, 11, 16]). So, we get a probability-based justification of this semi-heuristic suggestion.

Remaining open problem. What is the computational complexity of computing the proposed optimal solution?

For generic computable functions $f_i(x,c)$, even the general problem of solving a system of non-linear equations is not algorithmically decidable; see, e.g., [5]. If we restrict ourselves to polynomial functions $f_i(x,c)$, the problem becomes algorithmically decidable, but already for quadratic functions it is, in general, NP-hard; this means that unless it turns out that P = NP (which most computer scientists believe not to be true), no feasible algorithm is possible for solving all instances of this problem [5].

For linear functions $f_i(x,c)$, solving the corresponding system of linear equations is, of course, feasible. Interestingly, for linear functions $f_i(x,c)$, the general tolerance problem is also feasible: we can feasibly check whether the problem has a solution and, if it has, feasibly produce one of these solutions; see, e.g., [5].

A natural question is whether for this linear case, the problem of computing the optimal tuple x – in our new sense – is still feasible. This is an important open problem.

Acknowledgments

This work was supported in part by the National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science), HRD-1834620 and HRD-2034030 (CAHSI Includes), EAR-2225395 (Center for Collective Impact in Earthquake Science C-CIES), and by the AT&T Fellowship in Information Technology.

It was also supported by a grant from the Hungarian National Research, Development and Innovation Office (NRDI).

References

1. R. Belohlavek, J. W. Dauben, and G. J. Klir, *Fuzzy Logic and Mathematics: A Historical Perspective*, Oxford University Press, New York, 2017.
2. L. Jaulin, M. Kiefer, O. Didrit, and E. Walter, *Applied Interval Analysis, with Examples in Parameter and State Estimation, Robust Control, and Robotics*, Springer, London, 2012.
3. E. T. Jaynes and G. L. Bretthorst, *Probability Theory: The Logic of Science*, Cambridge University Press, Cambridge, UK, 2003.
4. G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.
5. V. Kreinovich, A. Lakeyev, J. Rohn, and P. Kahl, *Computational complexity and feasibility of data processing and interval computations*, Kluwer, Dordrecht, 1998.
6. B. J. Kubica, *Interval Methods for Solving Nonlinear Constraint Satisfaction, Optimization, and Similar Problems: from Inequalities Systems to Game Solutions*, Springer, Cham, Switzerland, 2019.
7. G. Mayer, *Interval Analysis and Automatic Result Verification*, de Gruyter, Berlin, 2017.
8. J. M. Mendel, *Explainable Uncertain Rule-Based Fuzzy Systems*, Springer, Cham, Switzerland, 2024.
9. R. E. Moore, R. B. Kearfott, and M. J. Cloud, *Introduction to Interval Analysis*, SIAM, Philadelphia, 2009.
10. H. T. Nguyen, C. L. Walker, and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2019.
11. V. Novák, I. Perfilieva, and J. Močkoř, *Mathematical Principles of Fuzzy Logic*, Kluwer, Boston, Dordrecht, 1999.
12. A. Piegat and M. Pluciński, "Solution of the basic tolerant interval and fuzzy equation using Shifted Membership Function method with examples of applications", *Information Sciences*, 2024, Vol. 660, Paper 120137.
13. S. G. Rabinovich, *Measurement Errors and Uncertainty: Theory and Practice*, Springer Verlag, New York, 2005.
14. S. P. Shary, "Solving the tolerance problem for interval linear systems", *Interval Computations*, 1994, No. 1, pp. 6–26.
15. S. P. Shary, "Solving the linear interval tolerance problem", *Mathematics and Computers in Simulation*, 1995, Vol. 39, No. 1–2, pp. 53–85.
16. L. A. Zadeh, "Fuzzy sets", *Information and Control*, 1965, Vol. 8, pp. 338–353.