How to represent and process uncertainty in practical situations: towards a feasible combination of probabilistic, interval, and fuzzy uncertainty

Olga Kosheleva and Vladik Kreinovich

Abstract Computers deal with numbers that represent the values of physical quantities. What will be the numbers of the future? They need to take into account all kinds of uncertainty, they need to take into account dependence between different quantities. This paper provides a vision of how such numbers of the future will look like.

1 Need for data processing

In many application areas, we need to process data. We need to transform the available information $x_1, \ldots, x_n$ into an estimate y for some quantity:

- describing the current or the future state of the world, or
- describing an action or design that is recommended based on this information.

In this paper, we will denote the data processing algorithm by

$$y = f(x_1, \ldots, x_n).$$

Data processing is what computers were invented for. Data processing is what computers are mostly used for now.

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

2 AI-based data processing has become ubiquitous

In the last decades, more and more data processing is done by AI-based algorithms, mostly by deep neural networks. To come up with these algorithms, we first train a multi-layer neural network on thousands and millions of examples. As a result, we come with the weights for which the neural network best fits these examples. This training usually takes a lot of time. Once the training is done, the weights are fixed ("frozen"), and the neural network is ready to be used for data processing.

3 Need for uncertainty propagation

The data $x_1,\ldots,x_n$ that we process comes either directly from measurements, or from some previous processing of measurement results. Measurements are never absolutely accurate; see, e.g., [20]. The result $\widetilde{x}$ of measuring a quantity is, in general, somewhat different from the actual (unknown) value x of this quantity. Because of this, the value $\widetilde{y} = f(\widetilde{x}_1,\ldots,\widetilde{x}_n)$ that we get by processing measurement results is, in general, somewhat different from the value $y = f(x_1,\ldots,x_n)$ that we would get if we knew the exact values of the corresponding quantities.

To make an appropriate decision, it is important to know how accurate is our estimate. For example, if we estimate the amount of oil in an oilfield and our estimate is 150 million tons, there is a big difference between:

- a situation in which it is 150 ± 50 – so we should start exploiting this field, and
- a situation in which it is 150 ± 200, so that maybe there is no oil at all, and we should perform additional tests.

4 Uncertainty propagation: what we need to estimate

We need to estimate the accuracy of the results of data processing: how the result $\widetilde{y} = f(\widetilde{x}_1,\ldots,\widetilde{x}_n)$ of processing measurement results $\widetilde{x}_i$ is different from the ideal value $y = f(x_1,\ldots,x_n)$ that we would have gotten if we knew the actual values x_i. In other words, we need to estimate the difference

$$\Delta y = \widetilde{y} - y = f(\widetilde{x}_1,\ldots,\widetilde{x}_n) - f(x_1,\ldots,x_n).$$

By definition of Δx_i as the difference $\Delta x_i = \widetilde{x}_i - x_i$, we have $x_i = \widetilde{x}_i - \Delta x_i$. Substituting this expression for x_i into the formula for Δy, we conclude that

$$\Delta y = \widetilde{y} - y = f(\widetilde{x}_1,\ldots,\widetilde{x}_n) - f(\widetilde{x}_1 - \Delta x_1,\ldots,\widetilde{x}_n - \Delta x_n).$$

5 Different types of uncertainty

In the very beginning of the data processing process, as raw data, we have measurement results and expert estimates. Measurements are never 100% accurate. There is usually a difference between the measurement result $\widetilde{x}$ and the (unknown) actual value x of the measured quantity. This difference is known as the *measurement error*.

For a measurement result, sometimes, we know the *probability distribution* of the measurement error. However, often, all we know is the upper bound Δ on the absolute value $|\Delta x|$ of the measurement error. Indeed, probability distributions come from *calibration* – comparing current measurement instrument with a more accurate one.

We *cannot* do it for state-of-the-art instruments, for which there are no more-accurate ones. We *do not* do it in many industrial applications, since calibration is expensive and is only done if needed. In these cases, we have *interval uncertainty*; see, e.g., [6, 12, 15, 17].

For expert estimates, we can have different upper bounds $\Delta(\alpha)$ with different degrees of confidence α. This forms what is known as a *fuzzy number*; see, e.g., [1, 7, 16, 18, 19, 24].

6 Possibility of linearization

Measurement errors are usually relatively small. As a result, terms which are quadratic – or higher order – in terms of measurement errors are much smaller than linear terms and can, therefore, be safely ignored. For example:

- even if we have a not very accurate measurement – with accuracy 10%,
- the square of 10% is 1% which is an order of magnitude smaller than 10%.

Thus, we can do what physicists usually do in such situations:

- expand the expression for Δy in Taylor series in terms of Δx_i and
- keep only linear terms in this expansion.

Here,

$$f(\widetilde{x}_1 - \Delta x_1, \ldots, \widetilde{x}_n - \Delta x_n) = \widetilde{y} - \sum_{i=1}^{n} c_i \cdot \Delta x_i,$$

$$\text{where we denoted } c_i \stackrel{\text{def}}{=} \frac{\partial f}{\partial x_i}_{|x_1=\widetilde{x}_1,\ldots,x_n=\widetilde{x}_n}.$$

Substituting this expression into the linearized formula for Δy, we conclude that

$$\Delta y = \sum_{i=1}^{n} c_i \cdot \Delta x_i.$$

Depending on what we know about the measurement uncertainty Δx_i, we can get similar information about Δy.

7 Case of probabilistic uncertainty

In many cases, we know the probability distributions of all measurement errors, and we also know that measurement errors corresponding to different measurements are statistically independent. This means, in particular, that we know the mean m_i (also known as *bias*) and the standard deviation σ_i of each measurement error.

Since we know the bias, we can simply subtract this bias from all measurement results and thus get this bias equal to 0. In this case, the mean value of Δy is also 0, and the standard deviation σ of Δy is described by the following formula:

$$\sigma^2 = \sum_{i=1}^{n} c_i^2 \cdot \sigma_i^2.$$

8 Interval case

In many other cases, we do not know the probabilities, all we know are bounds Δ_i on the absolute values of the measurement errors Δx_i:

$$|\Delta x_i| \le \Delta_i.$$

In this case, after we know the measurement result $\widetilde{x}_i$, the only information that we gain about the actual value x_i is that this value is somewhere in the interval $[\widetilde{x}_i - \Delta_i, \widetilde{x}_i + \Delta_i]$. Because of this fact, such cases are known as cases of *interval uncertainty*.

In this case, all we can do is find the set of possible values of Δy. One can check that this set is an interval $[-\Delta, \Delta]$, where

$$\Delta = \sum_{i=1}^{n} |c_i| \cdot \Delta_i.$$

9 Need to deal with all types of uncertainty

Sometimes, we only deal with raw data. However, usually, when we process data, we deal with the results of some previous data processing. Data processing that typically uses results with all three types of uncertainty.

This is especially important for modern AI technique. They use thousands and millions of data points. Many of these data points are the results of previous data processing. With so many results, it is inevitable that they include all three types of uncertainty; see, e.g., [8, 13]. So, to describe both inputs and outputs of data processing, we need to take all these three types of uncertainty into account.

10 Linearization helps

In the original data processing, all Δx_i directly come from measurements or expert estimates. In this case, we have $\Delta y = \Delta y^r + \Delta y^i + \Delta y^f$. Here each component $\Delta y^\times$ comes from only taking into account inputs with the corresponding uncertainty. In the following data processing, we similarly have

$$\Delta y = \Delta y^r + \Delta y^i + \Delta y^f, \text{ where } \Delta y^\times = \sum_{i=1}^{n} c_i \cdot \Delta x_i^\times.$$

Thus, we can process all three types of uncertainty separately. This leads us to the following vision.

11 This is how we envision a number of the future: first approximation

In the first approximation, a number of the future is a tuple consisting of:

- an approximate numerical value $\widetilde{x}$,
- characteristic(s) of a probabilistic part of the approximation error – e.g., the standard deviation σ,
- upper bound Δ from the interval part of the uncertainty, and
- characteristic(s) for the fuzzy part – e.g., the upper bound δ corresponding to the triangular form of the fuzzy uncertainty part.

For data processing $y = f(x_1, \ldots, x_n)$, we have:

$$\widetilde{y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_n), \quad \sigma = \sqrt{\sum_{i=1}^{n} c_i^2 \cdot \sigma_i^2},$$

$$\Delta = \sum_{i=1}^{n} |c_i| \cdot \Delta_i, \quad \delta = \sum_{i=1}^{n} |c_i| \cdot \delta_i.$$

12 But this is not all

The same raw data values are often used to compute different inputs to our data processing algorithm. In this case, even when all original measurement errors are independent, there is a correlation between the inputs.

So, to properly describe uncertainty, it is not enough to describe the numbers themselves. It is important to also describe *relations* between different numbers.

13 How can we describe these relations: analysis of the problem and the resulting suggestions

In the probabilistic case, relations can be described by correlation. In the Gaussian case, a multi-D distribution with 0 means can be described by a covariance matrix C. And if we have known means, we can subtract them.

Often, we need to make a binary decision – what is possible, what is not. Naturally, if alternative a is not possible and b is less probable ($f(b) < f(a)$), then b is also not probable. So, the set of all possible alternative is $\{a : f(a) \geq p_0\}$ for some p_0.

For a normal distribution, this set is an ellipsoid $x^T Sx \leq c_0$, where $S = C^{-1}$.

In the interval case, it can be proven that the simplest way to describe such relation is also by an ellipsoid $x^T Sx \leq c_0$; see, e.g., [2, 5, 9, 11, 22].

How do we get the parameters of this ellipsoid? For every two numbers, we describe the parameters of the corresponding ellipse [9].

For the fuzzy part, it is also reasonable to use the corresponding ellipse [9]. For interval and fuzzy data processing, it makes sense to consider the inverse matrix $C = S^{-1}$. This enables us to use same algorithms for processing such uncertainty.

In the linearized case, for $y_k = f_k(x_1, \ldots, x_n)$, once we have a matrix $C^x_{ij} = E[\Delta x_i \cdot \Delta x_j]$, we have

$$C^y_{k\ell} = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ki} \cdot c_{\ell j} \cdot C^x_{ij}, \text{ where } c_{ki} \stackrel{\text{def}}{=} \frac{\partial f_k}{\partial x_i}.$$

Similar formulas can be used for interval and fuzzy uncertainty. This enables us, e.g., to avoid overestimation.

Overestimation was the main problem of the traditional interval computation techniques. Now we have a more adequate picture of numbers of the future: tuple-numbers plus relations between pairs of numbers.

14 Discussion

Our proposal is similar to the 1960s transition in physics; see, e.g., [4, 23].

Before that, we had a hierarchical model of matter:

- molecules consist of atoms,
- atoms consist of elementary particles, etc.

In this model, we can separate atoms, separate particles, and study the properties of each component.

In the modern physics, e.g., a proton consists of quarks, but quarks cannot be separated. We always need to take into account their interaction.

Similarly, we describe the overall uncertainty by simply describing uncertainty of each number. We need to take relation into account.

Another analogy is quantum computing. There, it is crucial to use – and take into account – entanglement between states of individual objects.

15 What next?

1) In the future, quantum computing becomes ubiquitous. Then, the we will need to add the fourth uncertainty component to the description of numbers and their relation. Namely, we need to add a component corresponding to quantum uncertainty.
2) There are also situations when measurement errors are larger, and quadratic terms can no longer be ignored. Good news is that with ellipsoids, we can get exact range of quadratic functions – while it is NP-hard for interval uncertainty (see, e.g., [10]). In such situations, we need to take into account relation between different types of uncertainty. This is already done with p-boxes (see, e.g., [3]) and related techniques. We need to extend this to all four types of uncertainty.
3) It may be desirable to consider relations between 3, 4, etc. quantities – to have hypergraph instead of a graph; see, e.g., [21].

Acknowledgments

This work was supported by the AT&T Fellowship in Information Technology, by the Institute for Risk and Reliability, Leibniz Universitaet Hannover, Germany, by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Focus Program SPP 100+ 2388, Grant Nr. 501624329, by the European Union under the project ROBOPROX (No. CZ.02.01.01/00/22 008/0004590), by the Center of Excellence in Econometrics, Faculty of Economics, Chiang Mai University, Thailand, by the Ho Chi Minh City University of Banking, Vietnam, and by Thang Long University, Hanoi, Vietnam.

The authors want to thank all the participants of the Numbers of the Future Conference (Liverpool, UK, December 3–5, 2025) for valuable discussions.

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. F. L. Chernousko, *State Estimation for Dynamical Systems*, CRC Press, Boca Raton, Florida, 1994.
3. S. Ferson, V. Kreinovich, L. Ginzburg, D. S. Myers, and K. Sentz, *Constructing Probability Boxes and Dempster-Shafer Structures*, Sandia National Laboratories, Report SAND2002-4015, January 2003.
4. R. Feynman, R. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Addison Wesley, Boston, Massachusetts, 2005.
5. A. Finkelstein, O. Kosheleva, and Vladik Kreinovich, "Astrogeometry, error estimation, and other applications of set-valued analysis", *ACM SIGNUM Newsletter*, 1996, Vol. 31, No. 4, pp. 3–25.
6. 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.
7. G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.
8. O. Kosheleva and V. Kreinovich, "How to Propagate Uncertainty via AI Algorithms", *Proceedings of the 10th International Workshop on Reliable Engineering Computing REC'2024*, Beijing, China, October 26–27, 2024, pp. 5–12.
9. V. Kreinovich, J. Beck, and H. T. Nguyen, "Ellipsoids and Ellipsoid-Shaped Fuzzy Sets as Natural Multi-Variate Generalization of Intervals and Fuzzy Numbers: How to Elicit Them from Users, and How to Use Them in Data Processing", *Information Sciences*, 2007, Vol. 177, No. 2, pp. 388–407.
10. V. Kreinovich, A. Lakeyev, J. Rohn, and P. Kahl, *Computational complexity and feasibility of data processing and interval computations*, Kluwer, Dordrecht, 1998.
11. V. Kreinovich, A. Neumaier, and G. Xiang, "Towards a Combination of Interval and Ellipsoid Uncertainty", Computational Technologies, 2008, Vol. 13, No. 6, pp. 5–16.
12. B. J. Kubica, *Interval Methods for Solving Nonlinear Constraint Satisfaction, Optimization, and Similar Problems: from Inequalities Systems to Game Solutions*, Springer, Cham, Switzerland, 2019.
13. C. Q. Lauter, M. Ceberio, V. Kreinovich, and O. M. Kosheleva, "Uncertainty Quantification for Results of AI-Based Data Processing: Towards More Feasible Algorithms", In: F. Pavese, A. Forbes, A. Bosnjakovic, S. Eichstaedt, and J. Sousa (eds.), *Advanced Mathematical and Computational Tools for Metrology and Testing XIII*, World Scientific, Singapore, 2025, pp. 95–108.
14. S. Li, Y. Ogura, and V. Kreinovich, *Limit Theorems and Applications of Set Valued and Fuzzy Valued Random Variables*, Kluwer Academic Publishers, Dordrecht, 2002.
15. G. Mayer, *Interval Analysis and Automatic Result Verification*, de Gruyter, Berlin, 2017.
16. J. M. Mendel, *Explainable Uncertain Rule-Based Fuzzy Systems*, Springer, Cham, Switzerland, 2024.
17. R. E. Moore, R. B. Kearfott, and M. J. Cloud, *Introduction to Interval Analysis*, SIAM, Philadelphia, 2009.
18. H. T. Nguyen, C. L. Walker, and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2019.
19. V. Novák, I. Perfilieva, and J. Močkoř, *Mathematical Principles of Fuzzy Logic*, Kluwer, Boston, Dordrecht, 1999.
20. S. G. Rabinovich, *Measurement Errors and Uncertainty: Theory and Practice*, Springer Verlag, New York, 2005.
21. H. A. Reyes, C. Joslyn, and V. Kreinovich, "Graph Approach to Uncertainty Quantification", In: N. H. Phuong and V. Kreinovich (eds.), *Deep Learning and Other Soft Computing Techniques: Biomedical and Related Applications*, Springer, Cham, Switzerland, 2023, pp. 253–284.

22. F. N. Schietzold, J. Urenda, V. Kreinovich, W. Graf, and M. Kaliske, “Why Ellipsoids in Mechanical Analysis of Wood Structures”, Proceedings of the 9th International Workshop on Reliable Engineering Computing REC’2021, Taormina, Italy, May 16–20, 2021, pp. 604–614.
23. K. S. Thorne and R. D. Blandford, *Modern Classical Physics: Optics, Fluids, Plasmas, Elasticity, Relativity, and Statistical Physics*, Princeton University Press, Princeton, New Jersey, 2021.
24. L. A. Zadeh, “Fuzzy sets”, *Information and Control*, 1965, Vol. 8, pp. 338–353.