LOW-COMPLEXITY ZONOTOPES CAN ENHANCE UNCERTAINTY QUANTIFICATION (UQ)

Olga Kosheleva[1] and Vladik Kreinovich[2]

[1]Department of Teacher Education
University of Texas at El Paso
500 W. University, El Paso, TX 79968, USA
e-mail: olgak@utep.edu

[2]Department of Computer Science
University of Texas at El Paso
500 W. University, El Paso, TX 79968, USA
e-mail: vladik@utep.edu

Abstract. *In many practical situations, the only information that we know about the measurement error is the upper bound Δ on its absolute value. In this case, once we know the measurement result $\widetilde{x}$, the only information that we have about the actual value x of the corresponding quantity is that this value belongs to the interval $[\widetilde{x} - \Delta, \widetilde{x} + \Delta]$. How can we estimate the accuracy of the result of data processing under this interval uncertainty? In general, computing this accuracy is NP-hard, but in the usual case when measurement errors are relatively small, we can linearize the problem and thus, make computations feasible. This problem is well studied when data processing results in a single value y, but usually, we use the same measurement results to compute the values of several quantities $y_1, \ldots, y_n$. What is the resulting set of tuples $(y_1, \ldots, y_n)$? In this paper, we show that this set is a particular case of what is called a zonotope, and that we can use known results about zonotopes to make the corresponding computational problems easier to solve.*

Keywords: Uncertainty Quantification, Interval Uncertainty, Zonotopes

1 FORMULATION OF THE PROBLEM

1.1 Main objective of science and engineering

What do we want? We want to predict what will happen in the future – this is what science does, and what we want to select the actions that will leads to the best possible future – this is, crudely speaking, what engineering is for.

Both to predict the future state of the world and to select the best action, we must have information about the current state of the world, i.e., about the values of all the quantities that characterize this state. This information mostly comes from measurements. To predict the future value y of a quantity or to describe each control parameter y, we use the known relation $y = f(x_1, \ldots, x_N)$ between this future value (or control parameter) and the current values of several related quantities $x_1, \ldots, x_N$:

- we measure the values of the quantities $x_1, \ldots, x_N$, and

- we apply the algorithm $f(x_1, \ldots, x_N)$ to the results $\widetilde{x}_1, \ldots, \widetilde{x}_N$ of measuring the quantities $x_1, \ldots, x_N$, and return the value $\widetilde{y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_N)$.

1.2 Need for uncertainty quantification

Measurements are never absolutely accurate; see, e.g., [9]. The result $\widetilde{x}$ of each measurement is, in general, different from the actual (unknown) value x of the corresponding quantity. In other words, the *measurement error* $\Delta x \stackrel{\text{def}}{=} \widetilde{x} - x$ is, in general, different from 0.

Since, in general, the measurement result $\widetilde{x}_i$ is, in general, different from x_i, our estimate $\widetilde{y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_N)$ based on the measurement results is, in general, different from the desired value $y = f(x_1, \ldots, x_N)$.

How different can they be? What can we say about the estimation error $\Delta y \stackrel{\text{def}}{=} \widetilde{y} - y$? This is very important to know in many practical situations. For example, suppose that we are prospecting for oil, and we estimated that in some location, there is $\widetilde{y} = 150$ million tons. What shall we do? It depends on the accuracy of this estimate:

- if $y = 150 \pm 20$, this is very good news; we should dig a well and start producing oil;

- on the other hand, if $y = 150 \pm 200$, then maybe at this location, there is no oil at all; in this case, it is better to perform some additional measurements first, to decrease the risk of wasting money on the expensive well.

Estimating the approximation error Δy based on the known information about the measurement errors Δx_i is one of the main problems of uncertainty quantification.

1.3 Traditional probability-based approach to uncertainty quantification and its limitations

Traditional engineering approach to uncertainty quantification assumes that we know the probability distributions of each measurement error Δx_i [9]. And indeed, in many real-life situations we have this knowledge. However, there are many important practical situations when we do not know these probabilities. To explain why, let us recall where the information about the probabilities comes from.

In the ideal world, for each measuring instrument, we should compare, several times, the measurement result $\widetilde{x}$ with the actual value x of the corresponding quantity – and for each

such comparison, compute the measurement error $\Delta x = \widetilde{x} - x$. After a sufficient number of measurements, we would get a large sample of values Δx. Based on this sample, we will then be able to find the corresponding probability distribution.

Of course, in reality, we never know the exact actual values of the physical quantities. However, in many cases, there exists another – much more accurate – measuring instrument, whose measurement error Δx_s is much smaller than the measuring error of the tested instrument; such much-more-accurate measuring instruments are known as *standard* ones. In this case, with high accuracy, the value $\widetilde{x}_s$ measured by the standard measuring instrument is approximately equal to the actual value, and the difference $\delta x \overset{\text{def}}{=} \widetilde{x} - \widetilde{x}_s$ between the results of the two measurements is approximately equal to the desired measurement error Δx. Thus, we can measure, several times, the same quantities by both measuring instruments, and use the resulting sample to find the probability distribution of the corresponding measurement error.

In many cases, such a calibration is indeed performed, and we get the corresponding probability distributions. However, in many other cases, such a calibration is not done – and thus, we do not know the corresponding probabilities. There are two main reasons why calibration is not done.

The first reason is that sometimes, we use state-of-the-art measuring instruments, for which no other instrument is more accurate. This happens a lot in advanced science: e.g., it would be nice if near the Hubble telescope, we would have a 5 times more accurate instrument – but the Hubble telescope is the best we have. This often happens in applications as well. For example, geophysical companies often use state-of-the-art measuring equipment: this equaipment costs money, but it is still cheaper to use such expensive measuring instruments than to risk wasting even more money on, e.g., drilling oil well where there is no oil at all.

The second reason is more mundane: yes, potentially, in a manufacturing plant, we can, in principle, calibrate all the sensors, and get the corresponding probability distributions, but there is a problem. Many sensors are very cheap nowadays: kids play with robotic toys that measure distances to the walls etc. as they go, and these sensors can be bought for a few bucks. However, calibrating each sensor requires access to an expensive accurate measuring instrument – and it would cost several orders of magnitude more than the sensor itself. This is too expensive for a manufacturing plant — which already usually operates at a very low profit margin.

1.4   Enter interval uncertainty

If we do not know probabilities of different values of measurement error Δx, what do we know? For a device to be called a measuring instrument, we need to know at least some upper bound Δ on the absolute value of the measurement error: $|\Delta x| \leq \Delta$. If we do not even know such an upper bound, this means that after a measurement by this instrument, we cannot say anything about the actual value of the measured quantity: it can be as far away from the measurement result as we can imagine. In other words, what such a device would produce is a wild guess, not a measurement result. Thus, such a bound is always produced by the manufacturer of the measuring instrument.

And if we cannot find the probabilities of different values Δx, this upper bound is all we know. In this case, once we know the measurement result $\widetilde{x}$, the only information that we have about the actual value x of the measured quantity is that this value is somewhere in the interval $[\widetilde{x} - \Delta, \widetilde{x} + \Delta]$. Such uncertainty is naturally called *interval uncertainty*.

1.5 Need for interval computations

Let us go back to the situation when, instead of the actual (ideal) value $y = f(x_1, \ldots, x_N)$, we have an estimate $\widetilde{y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_N)$ based on the measurement results $\widetilde{x}_1, \ldots, \widetilde{x}_N$. If for each of N measurements, we only know the upper bound Δ_i on the absolute value of the corresponding measurement error Δx_i, then all we know about the actual value y is that it is equal to $f(x_1, \ldots, x_N)$ for some $x_i \in [\widetilde{x}_i - \Delta_i, \widetilde{x}_i + \Delta_i]$.

Thus, all we can say about the value y is that it belongs to the set

$$Y \stackrel{\text{def}}{=} \{f(x_1, \ldots, x_N) : x_i \in [\widetilde{x}_i - \Delta_i, \widetilde{x}_i + \Delta_i] \text{ for all } i\}. \tag{1}$$

For continuous functions $f(x_1, \ldots, x_N)$ this set is also an interval. The problem of computing the endpoint of this interval is known as the problem of *interval computations*; see, e.g., [7, 8].

In general, the interval computation problem is NP-hard; see, e.g., [6]. This means that unless P = NP (which most computer scientists do not believe to be true), it is not possible to have a feasible algorithm that solves all particular cases of this problems. However, in many practical situations, there exist efficient algorithms that either compute the desired range Y – or at least compute a good approximation to Y.

1.6 Possibility of linearization

One of the cases when a feasible algorithm for uncertainty quantification is possible is when the weasurement errors Δx_i are reasonably small – and usually, they are reasonable small. In this case, we can use one of the main ideas of computations in physics (see, e.g., [2, 12]): expand the corresponding expression in Taylor series in terms of the corresponding small quantities, and keep only linear terms in this expansion. In our case, by definition of the measurement error $\Delta x_i = \widetilde{x}_i - x_i$, we have $x_i = \widetilde{x}_i - \Delta x_i$, thus:

$$\Delta y = f(\widetilde{x}_1, \ldots, \widetilde{x}_N) - f(x_1, \ldots, x_N) = f(\widetilde{x}_1, \ldots, \widetilde{x}_N) - f(\widetilde{x}_1 - \Delta x_1, \ldots, \widetilde{x}_N - \Delta x_N). \tag{2}$$

Expanding the expression in the right-hand side of (2) in Taylor series in terms of Δx_i, we get

$$f(\widetilde{x}_1 - \Delta x_1, \ldots, \widetilde{x}_N - \Delta x_N) = f(\widetilde{x}_1, \ldots, \widetilde{x}_N) - \sum_{i=1}^{N} c_i \cdot \Delta x_i, \tag{3}$$

where we denoted $c_i \stackrel{\text{def}}{=} \dfrac{\partial f}{\partial x_i}$. Thus, the formula (2) takes the following form:

$$\Delta y = \sum_{i=1}^{N} c_i \cdot \Delta x_i. \tag{4}$$

In this linearized case, we can feasibly compute the bounds on Δy. Indeed, since each measurement error Δx_i takes values from the interval $[-\Delta_i, \Delta_i]$, and different measurement errors do not depend on each other, the largest possible value of the sum (1) is attained when each term $c_i \cdot \Delta x_i$ attains the largest possible value. The corresponding linear function $c_i \cdot \Delta x_i$ is increasing when $c_i > 0$ and decreasing when $c_i < 0$. Thus:

- when $c_i > 0$, the largest possible value of the quantity $c_i \cdot \Delta x_i$ is attained when Δx_i is the largest possible, i.e., when $\Delta x_i = \Delta_i$; the resulting largest value of the quantity $c_i \cdot \Delta x_i$ is equal to $c_i \cdot \Delta_i$;

- when $c_i < 0$, the largest possible value of the quantity $c_i \cdot \Delta x_i$ is attained when Δx_i is the smallest possible, i.e., when $\Delta x_i = -\Delta_i$; the resulting largest value of the quantity $c_i \cdot \Delta x_i$ is equal to $-c_i \cdot \Delta_i$.

In both cases, the largest possible value of the quantity $c_i \cdot \Delta x_i$ is equal to $|c_i| \cdot \Delta_i$. Thus, the largest possible value Δ of the sum (4) is equal to

$$\Delta = \sum_{i=1}^{N} |c_i| \cdot \Delta_i. \tag{5}$$

By using this formula, we can explicitly compute Δ in N steps – i.e., in feasible time.

Comment. While, strictly speaking, this algorithm is feasible, still, in situations when we have a large number N of inputs, it requires a large amount of computation time. It should be mentioned that there exist more efficient algorithms for computing Δ; see, e.g., [5].

1.7 Need to estimate the joint uncertainty of several data processing results – the main problem that we analyze in this paper

All the above discussions are about estimating a *single* quantity y. In reality, we usually estimate *several* different characteristics $y_1, \ldots, y_n$ based on the same data $\widetilde{x}_1, \ldots, \widetilde{x}_N$:

$$\widetilde{y}_1 = f_1(\widetilde{x}_1, \ldots, \widetilde{x}_N);$$

$$\ldots \tag{6}$$

$$\widetilde{y}_n = f_n(\widetilde{x}_1, \ldots, \widetilde{x}_N).$$

For example, when we predict weather, we do not just predict temperature at one locations, we predict weather, wind speed and direction, and humidity at several locations.

What is the accuracy of the resulting estimations? In other words, what can we say about the corresponding approximation errors

$$\Delta y_j \stackrel{\text{def}}{=} \widetilde{y}_j - f_j(x_1, \ldots, x_N). \tag{7}$$

As we have mentioned earlier, in many practical situations, we only know the upper bounds on the measurement errors – so that we have interval uncertainty, for which the only information that we have about each measurement error Δx_i is the upper bound Δ_i on its absolute value: $|\Delta x_i| \leq \Delta$. Also, in many practical situations, measurement errors are relatively small – so that we can ignore quadratic (and higher order) terms in the Taylor expansions. Then, we get the linearized formulas

$$\Delta y_1 = c_{1,1} \cdot \Delta x_1 + \ldots + c_{1,N} \cdot \Delta x_N;$$

$$\ldots \tag{8}$$

$$\Delta y_n = c_{n,1} \cdot \Delta x_1 + \ldots + c_{n,N} \cdot \Delta x_N,$$

where we denoted $c_{j,i} \stackrel{\text{def}}{=} \dfrac{\partial f_j}{\partial x_i}$.

For each value y_j, we can use the above techniques and find the interval of possible values of the approximation error Δy_j. However, this is not enough: we also need to also know what combinations of the values $(y_1, \ldots, y_n)$ – i.e., equivalently, of the approximation errors $(\Delta y_1, \ldots, \Delta y_n)$ – are possible. For example, when we predict weather, in some cases, the

future temperature in two nearby locations can range from 15 to 25 degrees. However, unless these two locations are separated by a mountain – as we have in our city of El Paso – the temperatures at these two locations cannot differ two much: we can have $(15, 16)$ and even, probably, $(15, 17)$, but we cannot have $(15, 25)$. How can we take this into account? How can we describe the corresponding set of tuples $(y_1, \ldots, y_n)$?

This is the problem that we analyze in this paper.

2 ANALYSIS OF THE PROBLEM: ENTER ZONOTOPES

2.1 General approach to solving problems: reduce and/or reformulate

A usual approach to solving a new problem is to try to find similar problems that have been already solved – or at least for which there are some partial solutions. If we cannot immediately come up with such a similar somewhat-solved problem, a natural idea is to try to reformulate our problems in equivalent terms so that it will be easier to find a similar problem.

2.2 Enter zonotopes

For our problem, this reformulation becomes possible if we reformulate the formulas (8) in vector terms, as
$$\Delta y = c_1 \cdot \Delta x_1 + \ldots + c_N \cdot \Delta x_N, \tag{9}$$
where $\Delta x_i \in [-\Delta_i, \Delta_i]$, and we denoted
$$\Delta y \stackrel{\text{def}}{=} (\Delta y_1, \ldots, \Delta y_n) \tag{10}$$
and
$$c_i \stackrel{\text{def}}{=} (c_{1,i}, \ldots, c_{n,i}). \tag{11}$$
For each j, the set S_i of all the vectors $\Delta x_i \cdot c_i$ for $\Delta x_i \in [-\Delta_i, \Delta_i]$ forms a straight line segment connecting the points $\Delta_i \cdot c_i$ and $-\Delta_i \cdot c_i$. The desired set S of all possible values of the sum (9) is thus equal to the set of all possible sums of vectors from the corresponding sets S_i:
$$S = \{s_1 + \ldots + s_N : s_1 \in S_1, \ldots, s_N \in S_N\}. \tag{12}$$

In geometry, the construction (12) is known as a *Minkowski sum* of the sets $S_1, \ldots, S_N$; this sum is denoted by
$$S = S_1 + \ldots + S_N. \tag{13}$$

The Minkowski sum of several straight line segments is known as a *zonotope*. Thus, our conclusion is that the desired set of posisble values of the tuple $\Delta y = (\Delta y_1, \ldots, \Delta y_n)$ is a zonotope.

2.3 Main conclusion of this section

So, to solve our main problem – of estimating the joint uncertainty of several data processing results – we need to be able to deal with zonotopes.

2.4 An interesting observation: every zonotope can be thus represented

We have shown that every set of possible values of the tuple $(\Delta y_1, \ldots, \Delta y_n)$ – and thus, of the tuple $(y_1, \ldots, y_n)$ – is a zonotope. Let us show that, vice versa, every zonotope can be thus represented. Indeed, in the above representation, we use straight-line segments centered at

0. Every straight-line segment T_i can be represented as the sum $T_i = m_i + S_i$ of its midpoint m_i and a segment $S_i \overset{\text{def}}{=} T_i - m_i$ centered at 0, i.e., a segment connecting one of its endpoints $c_i = (c_{i,1}, \ldots, c_{i,N})$ with the opposite endpoint $-c_i = (-c_{i,1}, \ldots, -c_{i,N})$. Thus, each zonotope

$$T = T_1 + \ldots + T_N \tag{14}$$

can be represented as

$$T = (m_1 + \ldots + m_N) + (S_1 + \ldots + S_N). \tag{15}$$

The set $S_1 + \ldots + S_n$ can be interpreted as the set of possible approximation errors for a data processing algorithm

$$f_i(x_1, \ldots, x_N) = c_{i,1} \cdot x_1 + \ldots + c_{i,N} \cdot x_n, \tag{16}$$

when we take $\Delta_1 = \ldots = \Delta_N = 1$. Thus, every zonotope can indeed be represented as the set of possible tuples $(y_1, \ldots, y_n)$ for some data processing algorithm.

2.5 Historical comment

The idea of using zonotopes was described, e.g., in [10], where it is shown that for a *specific* data processing algorithm – namely, for the least square estimation under interval uncertainty – the resulting set of possible tuples is a zonotope. In this paper, we show that this is true for *all* data processing algorithms – and we also show that, vice versa, every zonotope can be thus represented.

3 HOW TO DEAL WITH ZONOTOPES: WHAT IS KNOWN, WHAT WE PROPOSE, AND WHAT ARE THE REMAINING OPEN PROBLEMS

3.1 What is known

In computational geometry, there are several efficient algorithms for dealing with zonotopes; see, e.g., [3, 4]. Some of these algorithms have been efficiently used in [10].

3.2 What is the difficulty with the known algorithms

The main problem with these algorithms is that the exact description of the uncertainty-related zonotope in an n-dimensional space requires as many n-dimensional parameters c_i as there are measured quantities $x_1, \ldots, x_N$. In many practical problems, e.g., in seismology, N is in thousands, so this description becomes difficult to process.

3.3 What we propose: general idea

The possibility to make computations easier comes from the fact that the number n of desired properties $y_1, \ldots, y_n$ is much smaller than N. So, to speed up computations, we propose to use the known results [1, 11] about the possibility of approximating zonotopes with "low-complexity" sets, i.e., sets determined by a much smaller number of n-dimensional parameters.

By approximating a set S, we mean, as usual, producing a set A for which, for some small number δ:

- each element $s \in S$ is δ-close to some element $a \in A$, and

- each element $a \in A$ is δ-close to some element $s \in S$.

In mathematics, this closeness is usually described by saying that the Hausdorff distance $d_H(A, S)$ between the sets S and A is smaller than or equal to δ, where the Hausdorff distance $d_H(A, S)$ is defined as the small distance for which the above two conditions are true.

We can also say that the set A approximates the set S with *relative accuracy* ε if $d_H(A, S) \leq \delta \cdot \text{diam}(S)$, where the diameter $\text{diam}(S)$ is defined as the largest distance between two points from the set S – this is a natural generalization of the width of an interval and the diameter of a disk or of a sphere to general sets.

3.4 First simplifying result

The first simplifying result from [1, 11] is that each n-dimensional zonotope S can be approximated, with any given relative accuracy $\varepsilon > 0$, by a "low-complexity" zonotope, which is the sum of $N' = c(\varepsilon) \cdot n \cdot (\log(n))^3$ segments for some constant c depending on ε.

Since $n \ll N$, the new number of n-dimensional vector parameters is much smaller than N – thus, the problem becomes easier to handle.

3.5 Second simplifying result

The second simplifying result from [1, 11] is that each symmetric convex polyhedron – in particular, each zonotope – can be approximated, with any given relative accuracy $\varepsilon > 0$, by a convex polyhedron with "low" number of vertices $v_1, \ldots, v_{N'}$ – namely, by a polyhedron for which the number N' is also bounded, from above, by the value $c(\varepsilon) \cdot n \cdot (\log(n))^d \ll N$, for some small constant d.

In this case, the approximating set A is the convex combination of these vertices v_j. In other words, all elements $a = (a_1, \ldots, a_n)$ from the approximating set A have the form

$$a = c_1 \cdot v_1 + \ldots + c_{N'} \cdot v_{N'} \tag{17}$$

where $c_i \geq 0$ and

$$c_1 + \ldots + c_{N'} = 1. \tag{18}$$

3.6 What are the remaining open problems

The results from [1, 11] (that we propose to use) are effective – they drastically reduce the complexity of the corresponding problem – but at present, they are not supported by efficient computational algorithms for producing the corresponding approximations. These results are still useful – we spend time on computing the approximation only once, and then, we can enjoy the benefits of this reduction for every single way we want to process this set.

However, it would be nice to have efficient algorithms for producing the corresponding approximations. Designing such efficient algorithms is the main open problem that we want to emphasize. Hopefully, the fact that – as we have shown – such algorithms will be very helpful:

- not just in somewhat obscure computational geometry problems,

- but also in generic problems of uncertainty quantification

will hopefully encourage researchers to design such algorithms.

It was also supported by the program of the development of the Scientific-Educational Mathematical Center of Volga Federal District No. 075-02-2020-1478.

REFERENCES

[1] J. Bourgain, J. Lindenstrauss, and V. Milman, Approximation of zonoids by zonotopes, *Acta Mathematica*, **162**(1–2), 73–141, 1989.

[2] R. Feynman, R. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Addison Wesley, Boston, Massachusetts, 2005.

[3] J.E. Goodman and J. O'Rourke (eds.), *Handbook of Discrete and Computational Geometry*, CRC Press, Boca Raton, Florida, 1997.

[4] P.M. Gruber and J.M. Willis (eds.), *Handbook of Convex Geometry*, Elsevier, Amsterdam, 1993.

[5] V. Kreinovich and S. Ferson, A new Cauchy-based black-box technique for uncertainty in risk analysis, *Reliability Engineering and Systems Safety*, **85**(1–3), 267–279, 2004.

[6] V. Kreinovich, A. Lakeyev, J. Rohn, and P. Kahl, *Computational Complexity and Feasibility of Data Processing and Interval Computations*, Kluwer, Dordrecht, 1998.

[7] G. Mayer, *Interval Analysis and Automatic Result Verification*, de Gruyter, Berlin, 2017.

[8] R.E. Moore, R.B. Kearfott, and M.J. Cloud, *Introduction to Interval Analysis*, SIAM, Philadelphia, 2009.

[9] S.G. Rabinovich, *Measurement Errors and Uncertainties: Theory and Practice*, Springer, New York, 2005.

[10] S. Schön and H. Kurtterer, Using zonotopes for overestimation-free interval least squares – some geodetic applications, *Reliable Computing*, **11**, 137–155, 2004.

[11] T. Tao, Exploring the toolkit of Jean Bourgain, *Bulletin on the American Mathematical Society*, **58**, 155–171, 2021.

[12] K.S. Thorne and R.D. Blandford, *Modern Classical Physics: Optics, Fluids, Plasmas, Elasticity, Relativity, and Statistical Physics*, Princeton University Press, Princeton, New Jersey, 2017.