

Global Independence, Possible Local Dependence: Towards More Realistic Error Estimates for Indirect Measurements

Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
El Paso, Texas 79968, USA
vladik@utep.edu

Abstract—In many practical situations, it is not realistically possible to directly measure the desired physical quantity. In such situations, we have to measure this quantity indirectly, i.e., measure related quantities and use the known relation to estimate the value of the desired quantity. How accurate is the resulting estimate? The traditional approach assumes that the measurement errors of all direct measurements are independent. In many practical situations, this assumption works well, but in many other practical situations, it leads to a drastic underestimation of the resulting estimation error: e.g., when we base our estimate on measurements performed at nearby moments of time, since there is usually a strong correlation between the corresponding measurement errors. An alternative approach is when we make no assumptions about dependence. This alternative approach, vice versa, often leads to a drastic overestimation of the resulting estimation error. To get a more realistic estimate, it is desirable to take into account that while on the local level, we may have correlations, globally, measurement errors are usually indeed independent – e.g., measurements sufficiently separated in time and/or space. In this paper, we show how to analyze such situations by combining Monte-Carlo techniques corresponding to both known approaches. On the geophysical example, we show that this combination indeed leads to realistic estimates.

Keywords—*error estimation; indirect measurements; global independence; local dependence; Monte Carlo techniques*

I. FORMULATION OF THE PROBLEM

A. Need for Indirect Measurements

The main objective of science and engineering is to understand the world, to predict the future, and to describe how to make this future more favorable to us. For all these goals, we need to know the current state of the world, i.e., we need to know the current values of all the physical quantities.

Some quantities we can measure directly. For example, we can directly measure the current outdoors temperature, and we can directly measure the width of a street. However, for many other physical quantities, it is not easy (and often even impossible) to measure them directly. For example, we cannot

directly measure the temperature on the surface of the Sun, we cannot directly measure the distance to a nearby star, we cannot directly measure the density and temperature many kilometers below the Earth's surface, etc.

Since we cannot measure such quantities y directly, we have to measure them indirectly (see, e.g., [8]); namely:

- we find easier-to-measure quantities x_1, \dots, x_n which are related to the desired quantity y by a known dependence $y = f(x_1, \dots, x_n)$,
- we measure these quantities, and
- we use the results X_1, \dots, X_n of measuring these quantities to produce an estimate $Y = f(X_1, \dots, X_n)$ for the desired quantity y .

For example, to estimate the distance to a nearby star, we can use the following “parallax” method:

- we measure the direction to the star at two different times, when the Earth is at the opposite sides of the Sun, and then
- we use trigonometry and the known diameter of the Earth orbit to find the desired distance to the star.

Similarly, to find the density at different depth and different geographic locations:

- we create a small artificial explosion, and
- we use sensors placed at different locations to measure the time needed for the explosion-generated wave to propagate to these locations.

B. How Accurate Is the Result of Indirect Measurement?

To be useful, each measurement result – be it the result of a direct or an indirect measurement – has to come with some estimate of its accuracy. For example, if the estimate is that the given oilfield contains 100 million tons of oil, then:

- if this is 100 plus minus 10, it is probably advisable to start drilling, but

- if it is 100 plus minus 200, maybe there is no oil at all – and it makes sense to perform further measurements before investing a large amount of money in drilling a well.

Even when the dependence $y = f(x_1, \dots, x_n)$ is exact, the resulting estimate Y is not absolutely accurate, since, in general, the measurement results X_i are, in general, somewhat different from the actual (unknown) values x_i of the corresponding quantities. As a result, the estimate

$$Y = f(X_1, \dots, X_n)$$

is, in general, different from the actual value $y = f(x_1, \dots, x_n)$ of the desired quantity.

What can we say about the difference $d = Y - y$? In most practical situations, the measurements are reasonably accurate, i.e., that the measurement errors $d_i = X_i - x_i$ of direct measurements are small. Thus, in the dependence

$$\begin{aligned} d &= Y - y = f(X_1, \dots, X_n) - f(x_1, \dots, x_n) \\ &= f(X_1, \dots, X_n) - f(X_1 - d_1, \dots, X_n - d_n) \end{aligned}$$

we can safely ignore terms which are quadratic (or higher order) in terms of d_i and keep only linear terms. As a result, we get a known formula

$$d = c_1 * d_1 + \dots + c_n * d_n$$

where c_i denotes the value of the partial derivative of the function $f(x_1, \dots, x_n)$ with respect to x_i – partial derivative computed at the point (X_1, \dots, X_n) [8].

C. Usual Assumptions and the Resulting Estimate

In many practical situations, we can safely assume that the measuring instruments are well calibrated, so there is no bias, i.e., the mean value of the measurement error d_i is 0. We usually also know the mean square error (standard deviation) s_i of each direct measurement.

In this case, the mean value of the resulting measurement error d is also 0, and if we know the correlations c_{ij} between the measurement errors, then we can also find the mean square error s of the estimate Y as follows [8]:

$$s^2 = c_1^2 * s_1^2 + \dots + c_n^2 * s_n^2 + c_1 * c_2 * c_{12} * s_1 * s_2 + \dots$$

The problem is that in many practical situations, we do not know the correlations – and we still need to estimate how accurate is the result of an indirect measurement.

D. Traditional Approach to Error Estimation: Formula

In many cases, measurement errors of different measuring instruments are caused by different causes and thus, we can safely assume that these measurement errors are all independent. In such cases, the above formula can be simplified into

$$s^2 = c_1^2 * s_1^2 + \dots + c_n^2 * s_n^2$$

E. Traditional Approach to Error Estimation: Algorithms

When the number n of direct measurements is small, we can directly apply this formula. The data processing algorithm $f(x_1, \dots, x_n)$ is usually given as a black box (e.g., it may include solving a system of partial differential equations – a common situation in geophysics). In this case, to compute partial derivatives c_i , we can use numerical differentiation, i.e., compute each derivative as a ratio

$$c_i = (f(X_1, \dots, X_{i-1}, X_i + h, X_{i+1}, \dots, X_n) - Y) / h$$

So, we need one call to the algorithm $f(x_1, \dots, x_n)$ to compute Y and n calls to compute n values $f(X_1, \dots, X_{i-1}, X_i + h, X_{i+1}, \dots, X_n)$. Overall, we need $n+1$ calls.

When the number of variables is small, this is quite reasonable. However, in many practical situations, the number of inputs is huge, in hundreds and in thousands – this, e.g., the case in many geophysical situations, when we process a time series, a sequence of measurements corresponding to consequent moments of time. When each application of the data processing algorithm requires hours – and often hours on a high performance computer – there is no way that we can spend thousand time more time (i.e., years) just to estimate the accuracy of this estimate.

Good news is that there is a way out: namely, instead of directly following the above formula, we can use Monte-Carlo simulations. In each iteration of this method:

- we simulate random variables d_i which are normally distributed with mean 0 and standard deviation s_i , and then
- we apply the data processing algorithm to compute the value of the following difference:

$$f(X_1 + d_1, \dots, X_n + d_n) - Y$$

One can easily check that this difference is normally distributed with 0 mean and desired standard deviation s . Thus, to estimate s , we can repeat the above procedure several times and find the mean square value of the resulting differences.

The accuracy of the resulting statistical estimate is inverse proportional to the square root of the sample size: e.g., if we repeat the procedure 25 times, we get the accuracy 20%. This is quite reasonable if we take into account that we are estimating accuracy of accuracy, so we are talking about the difference between the accuracy of 10% and the accuracy of 12% – the difference which is usually safely ignored.

25 times means that we only need 26 calls to the time-consuming data processing algorithm. This is much much faster than straightforward computations that require thousands of such calls.

F. What Is We Take Into Account Possible Dependence: Formula

In many practical situations, measurement errors can indeed be safely assumed to be independent. However, there are many other situations, in which there is a clear correlation between measurement errors. For example, when we use the same sensor to measure the values of a quantity at two nearby moments of time, it is highly probable that the same source of noise that contributed to the measurement error of the first measurement will still be active during the second measurement as well – and thus, the corresponding measurement errors will be correlated.

One possible way to take this into account is not to make any assumptions about the possible correlations, and instead of consider all possible values of these correlations. For different values of the correlations, we get different value of the mean square error. Sometimes, we may be lucky, and the resulting mean square may be small, but the goal of measurement theory is to provide guaranteed estimates, not to rely on luck. If we want a guaranteed upper bound, we need to consider the largest possible value of the resulting mean square error s . One can check that this worst-case occurs if all the measurement errors are perfectly correlated, with coefficient of correlation 1 or -1; namely:

- errors corresponding to measurement with the same sign of c_i are perfectly positively correlated, while
- errors corresponding to different signs of the partial derivatives are perfectly negatively correlated.

The resulting worst-case mean square error has the form

$$s = |c_1| * s_1 + \dots + |c_n| * s_n$$

G. What Is We Take Into Account Possible Dependence: Algorithms

How can we compute this estimate? Similarly to the traditional case, we can do it directly, if we estimate the partial derivatives c_i by using the above-described numerical differentiation, and then explicitly estimate the above expression. However, as we have mentioned earlier, this is not realistic if we have a large number of direct measurements.

In this case, it is also possible to use Monte-Carlo simulations, but very different ones; see, e.g., [1-3,10]. Namely, instead of normal distributions with given standard deviation s_i , we should simulate Cauchy distributions with probability density proportional to

$$1/(1 + (x/s_i)^2).$$

It is known that if we take a linear combination

$$c_1 * d_1 + \dots + c_n * d_n$$

of independent Cauchy distributed random variables with parameters s_i , then this linear combination is also Cauchy distributed, with the parameter s described by the above worst-case formula. So, after repeating the procedure of simulations and computing the difference

$$f(X_1 + d_1, \dots, X_n + d_n) - Y,$$

we get a sample which is Cauchy distributed with the desired parameter s . We can then use, e.g., Maximum Likelihood Method to estimate the desired parameter.

Similarly to the more traditional Monte-Carlo technique, here the main advantage is that the number of iterations is determined only by how accurately we want to estimate s . So, even when we have thousands of inputs, and direct computation would take forever, Monte-Carlo simulations require a few dozen iterations and are, thus, quite doable.

H. A Caution

A caution is in order: the traditional Monte-Carlo simulation simulates a realistic case, since normal distributions are quite possible – and are, moreover, ubiquitous. This is due to the well-known Central Limit Theorem (see, e.g., [9]), according to which, under some reasonable conditions, the distribution of the sum of a large number of small independent random variables is close to normal – and tends to normal when the number of these variables increases. In real life, for well-calibrated measuring instruments, for which all the main sources of possible measurement error have been eliminated, what remains is a large number of difficult-to-eliminate small error components. As a result, for many measuring instruments, the error distribution is indeed close to Gaussian.

In contrast, Cauchy distribution is NOT a realistic representation of the measurement error – its standard deviation is actually infinite, so it cannot represent the measurement errors with given finite standard deviation (and this is how most students learn about this distribution – as an example of a distribution for which the standard deviation is infinite). In contrast to the normal distribution, Cauchy distribution is a mathematical trick, not a realistic simulation – but it is a very useful and computationally efficient trick.

I. Remaining Problem

From the theoretical viewpoint, everything seems to be perfect. However, when we applied both techniques to a real geophysical example, the results were very disappointing [7].

When we applied the traditional techniques – techniques that assume that all measurement errors are independent – we got unrealistically small mean square errors – much much smaller than, e.g., the difference between the two repetitions of the same measurement (and take into account that these geophysical measurements, the measured values such as density at the depth of 10 km do not change in a few hours). This can be easily explained on a simpler example: suppose that an electronic thermometer measures outside temperature with accuracy 1 degree. If we repeat the same measurement 10000 times, and assume that all the measurement errors are independent, then we will be able to conclude that by taking

the arithmetic average of these measurements, we would reach the accuracy of 1 divided by the square root of 10000, i.e., of 0.01 degrees – which the thermometer only has accuracy of 1 degree. (I wish it would be so easy to perform accurate measurements.) The reason is that measurement at nearby moments of time are affected by the same noises and are, thus, clearly not correlated.

OK, since there is correlation, let us use the second, worst-case approach. Another disappointment: the worst-case values are so huge that all the estimates are useless: for density of 5 units, the estimation error was much larger than 5.

One possible explanation would be that estimates are indeed useless, but this is not the case: similar procedures were used many times in geological practice. When appropriate, wells were drilled, and so we got actual values of density at different depths – values which were reasonably close to the original estimates.

Here comes a challenge: how to come up with more realistic estimates for the error of indirect measurement?

II. GLOBAL INDEPENDENCE, POSSIBLE LOCAL DEPENDENCE: DESCRIPTION OF THE PHENOMENON AND OUR IDEA OF HOW TO ESTIMATE IN THIS CASE

A. *Global Independence, Possible Local Dependence: A Phenomenon*

While:

- locally – e.g., for two nearby sensors or for the same sensor at two nearby moments of time – there is usually correlation between measurement errors,
- there is usually no correlation on the global scale, when we consider two sensors far away from each other or the same sensor at two moments of time separate by a large time interval.

How can we take this phenomenon into account?

- On the one hand, we have correlations, so we cannot use the usual technique which assumes that all measurement errors are independent.
- On the other hand, many of the measurement errors are independent. It would be nice to take this independence into account – so as to avoid drastic overestimations of the estimation error s .

B. *How to Deal With This Phenomenon: Brainstorming*

In both the traditional approach and the worst-case approach, we can use Monte-Carlo simulations.

- In the traditional approach, we use normal distributions – which thus correspond to the case of independence.
- In the worst-case approach, we use Cauchy distributions – which thus correspond to the case of possible dependence.

How can we use this information in our case? We can view the whole process of combining errors of direct measurements into a single value as a multi-stage process.

- At first, we combine the direct measurement errors corresponding to nearby measurements. These measurement errors are possibly correlated, so to combine them, we need to use the techniques corresponding to this case – e.g., simulations based on Cauchy distributions.
- Once we combined the direct measurement errors corresponding to nearby measurements into somewhat larger values, we then need to combine these larger values into the final estimate.

On this second stage, due to global independence, the corresponding combined errors are independent – and so, for combining them, we can use techniques corresponding to the independent case, such as Monte-Carlo simulations based on normal distributions.

To make computation time reasonable, we would like to implement this two-stage procedure in a Monte-Carlo way. The question is: what type of distribution should we use for this purpose?

- For small values corresponding to the original errors of direct measurement, the procedure should coincide with the possibly correlated case for which the Cauchy distribution is the most appropriate.
- For larger values corresponding to combined errors, the procedure should coincide with the independent case for which the normal distribution is the most appropriate choice.

It is therefore reasonable to consider a hybrid distribution:

- that coincides with the Cauchy distribution until we reach a certain threshold, and
- that coincides with the normal distribution once we reach this threshold.

C. *Resulting General Recommendation*

Based on what we described in the previous subsection, a reasonable idea is to perform the simulation as follows:

- We simulate both Cauchy and normal distributions.
- If both simulated values are smaller than the pre-selected threshold, we select the Cauchy value.

- If both simulated values are larger than the pre-selected threshold, we select the normal value.
- If the simulated values are on different sides of the threshold, we ignore both values, and repeat the simulations again.

D. What Threshold to Select?

What threshold should we select? To answer this question, let us recall what happens in the case of a normal distribution. In this case, theoretically, we can have values which are as large as possible. In practice, however, we select some confidence level, select a threshold corresponding to this confidence level, and consider value corresponding to this confidence level as impossible. This is how all metrological validations are made; see, e.g., [8]. Usually, a 95% confidence level is used; however, in some cases, lower and higher confidence levels are needed – e.g., in critical situations like manned space flights to nuclear power stations, we may require 99.9% (or even higher) confidence.

It is reasonable to apply the same idea to the case of Cauchy distribution: select a threshold corresponding, e.g., to 95% confidence interval. This corresponds to a threshold of $12.7 * s_i$ [7].

E. Simplifying the Above Simulation Algorithm

For this threshold value, the above simulation algorithm can be simplified. Indeed, for normal distribution, the probability to get a value larger than 12.7 times the standard deviation is practically 0. Thus, instead of the double simulation algorithm, we can simply simulate the cut-off version of Cauchy distribution, with a given cutoff. In other words:

- We simulate Cauchy distribution.
- If the simulated value is smaller than the pre-selected threshold, we select this value.
- If the simulated value is larger than the threshold, we ignore this value, and repeat the simulations again.

F. This Seems to Work

Interestingly, in our geophysical example [7], this idea worked: we got very realistic estimates for the resulting density values.

G. Why Now?

If these results were obtained several years ago, why are we promoting them now? The answer is very simple: when we first came up with these results, there was no serious justification for them: we were desperate to get reasonable error estimates, we could not get anything reasonable by using known techniques, so we came up with a purely heuristic method – for which we had some geophysical motivations. At that time, we did not promote it as a general recommended technique, since we realized that our arguments are valid only for this geophysical example.

Now, we have a general justification – and thus, we can proudly recommend this method for the general use.

H. Need for Future Work

Of course, the proposed method – while now (somewhat) justified -- is still heuristic. For example, it is still not clear what threshold is the best to use – or, e.g., whether it is a good idea to have an abrupt transition from Cauchy to normal distribution or it is better to have a smooth transition of some sort.

One can argue that we should not worry that much about the lack of justification for selecting a threshold, since this is a common problem for all statistical methods used in metrology – but we do worry, especially since for several metrological problems, we do have a good theoretical justification for the corresponding empirical heuristic formulas; see, e.g., [4].

In the corresponding analysis, it makes sense to start with the simplest case of global independence and possible local dependence – when we have three measurements. Globally, the 1st and the 3rd are independent, but there is a possible dependence on the local level, between the 1st and the 2nd measurements, and between the 2nd and the 3rd measurements. In this particular case, it is possible to come up with an explicit formula for the resulting worst-case standard deviation. Indeed, since the 1st and 3rd measurement errors are independent, the worst-possible case of the combined error

$$c_1 * d_1 + c_3 * d_3$$

is equal to the square root of $c_1^2 * s_1^2 + c_3^2 * s_3^2$. Since there may be possible correlation between this combined error and the second measurement error, to find the worst-case overall measurement error, we need to add $|c_2| * s_2$ to the above square root [10].

In particular, for the simplest case when all the coefficients c_i are equal to 1, the worst-case overall standard deviation is equal to $1 + \text{square root of } 2$, i.e., approximately to 2.41. As expected, this value is larger than the value of square root of 3 (approximately 1.73) corresponding to the independence case, and smaller than the value 3 corresponding to the case when all dependencies are possible.

ACKNOWLEDGMENT

The author is greatly thankful to the conference organizers for their invitation and support.

This work was also 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) and HRD-1242122 (Cyber-ShARE Center of Excellence).

REFERENCES

- [1] V. Kreinovich, “Interval computations and interval-related statistical techniques: tools for estimating uncertainty of the results of data processing and indirect measurements”, In: F. Pavese and A.B. Forbes

- (eds.), *Data Modeling for Metrology and Testing in Measurement Science*, Birkhauser-Springer, Boston, 2009, pp. 117-145.
- [2] V. Kreinovich, "Interval computations and interval-related statistical techniques: estimating uncertainty of the results of data processing and indirect measurements", In: F. Pavese, W. Bremser, A. Chunovkina, N. Fisher, and A.B. Forbes (eds.), *Advanced Mathematical and Computational Tools in Metrology and Testing AMTCM'X*, Singapore: World Scientific, Singapore, 2015, pp. 38-49.
- [3] V. Kreinovich and S. Ferson, "A new Cauchy-based black-box technique for uncertainty in risk analysis", *Reliability Engineering and Systems Safety*, vol. 85, no. 1-3, pp. 267-279, 2004.
- [4] H.T. Nguyen, V. Kreinovich, G.N. Solopchenko, C.-W. Tao, "Why two sigma? A theoretical justification", In: L. Reznik and V. Kreinovich (eds.), *Soft Computing in Measurements and Information Acquisition*, Berlin-Heidelberg: Springer-Verlag, 2003, pp. 10-22.
- [5] P.V. Novitskii and I.A. Zograph, *Estimating the Measurement Errors*, Energoatomizdat, Leningrad, 1991 (in Russian).
- [6] A.I. Orlov, "How often are the observations normal?", *Industrial Laboratory*, vol. 57, no. 7, pp. 770-772, 1991.
- [7] P. Pinheiro da Silva, A. Velasco, M. Ceberio, C. Servin, M.G. Averill, N. Del Rio, L. Longpre, and V. Kreinovich, "Propagation and provenance of probabilistic and interval uncertainty in cyberinfrastructure-related data processing and data fusion", In: R.L. Muhanna and R.L. Mullen (eds.), *Proceedings of the International Workshop on Reliable Engineering Computing REC'08*, Savannah, Georgia, February 20-22, 2008, pp. 199-234.
<http://www.cs.utep.edu/vladik/2007/tr07-56.pdf>
- [8] S. Rabinovich, *Measurement Errors and Uncertainties: Theory and Practice*, New York: Springer, 2005.
- [9] D.J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Boca Raton, Florida: Chapman and Hall/CRC, 2011.
- [10] R. Trejo and V. Kreinovich, "Error estimations for indirect measurements: randomized vs. deterministic algorithms for 'black-box' Programs", In: S. Rajasekaran, P. Pardalos, J. Reif, and J. Rolim (eds.), *Handbook on Randomized Computing*, Dordrecht: Kluwer, 2001, pp. 673-729.