11-1-2022

How to Get the Most Accurate Measurement-Based Estimates

Salvador Robles
The University of Texas at El Paso, sroblesher1@miners.utep.edu

Martine Ceberio
The University of Texas at El Paso, mceberio@utep.edu

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

Comments:
Technical Report: UTEP-CS-22-107

How to Get the Most Accurate Measurement-Based Estimates

Salvador Robles, Martine Ceberio, and Vladik Kreinovich

Abstract In many practical situations, we want to estimate a quantity y that is difficult – or even impossible – to measure directly. In such cases, often, there are easier-to-measure quantities $x_1, \ldots, x_n$ that are related to y by a known dependence $y = f(x_1, \ldots, x_n)$. So, to estimate y, we can measure these quantities x_i and use the measurement results to estimate y. The two natural questions are: (1) within limited resources, what is the best accuracy with which we can estimate y, and (2) to reach a given accuracy, what amount of resources do we need? In this paper, we provide answers to these two questions.

1 Introduction

Need for data processing and indirect measurements. One of the main objectives of science is to describe the current state of the world. The state of the world – and, in particular, the state of different objects and systems – is usually described by specifying the values of the corresponding quantities. For example, from the viewpoint of celestial mechanics, to describe the state of a planet, we need to know its location and its velocity.

Some quantities we can measure directly. For example, we can directly measure the size of an office or the weight of a person. Other quantities are difficult to measure directly. For example, at present, there is no way to directly measure the size or the weight of a planet. To estimate the value of each such quantity y, a natural idea is to find some easier-to-measure quantities $x_1, \ldots, x_n$ that are related to y by a known dependence $y = f(x_1, \ldots, x_n)$. Then, to estimate y:

- we measure these auxiliary quantities $x_1, \ldots, x_n$, and then

Salvador Robles, Martine Ceberio, and Vladik Kreinovich
Department of Computer Science, University of Texas at El Paso, 500 W. University
El Paso, Texas 79968, USA
e-mail: sroblesher1@miners.utep.edu, mceberio@utep.edu, vladik@utep.edu

- we use the results $\widetilde{x}_1, \ldots, \widetilde{x}_n$ of measuring these quantities to produce the estimate for y: $\widetilde{y} = f\,(\widetilde{x}_1, \ldots, \widetilde{x}_n)$.

Computing this estimate is an important case of *data processing*, and the whole measurement-based procedure for estimating y is known as *indirect measurement* (in contrast to *direct* measurements, when we simply measure the value y); see, e.g., [1].

Another objective of science is to predict the future state of the world – and this future state is also characterized by the future values of the corresponding quantities. For example, we may want to predict tomorrow's weather, i.e., temperature, wind speed, etc. To predict the future value y of each of these quantities, we can use the current values $x_1, \ldots, x_n$ of these quantities and of other related quantities, and we need to know how the future value y depends on these values x_i, i.e., we need to have an algorithm $y = f\,(x_1, \ldots, x_n)$ describing this prediction. In this case too, to estimate y:

- we measure the quantities $x_1, \ldots, x_n$, and then
- we use the results $\widetilde{x}_1, \ldots, \widetilde{x}_n$ of measuring these quantities to produce the estimate for y: $\widetilde{y} = f\,(\widetilde{x}_1, \ldots, \widetilde{x}_n)$.

This is another example of data processing and indirect measurement.

One of the main objectives of engineering is to describe how to make the world better. We may be looking for the best control strategy for a plant or for a car that will make it the most efficient and/or the least polluting, we may be looking for a gadget that leads to the best results. In all these cases, we need to find the optimal values of the quantities that characterize this control or this gadget. To find each some value y, we need to solve the corresponding optimization problem and find an algorithm $y = f\,(x_1, \ldots, x_n)$ that determines this value based on the known values of the quantities x_i that describe the corresponding environment. For example, the best control strategy for a petrochemical plant may depend on the amount of different chemical compounds in the incoming oil. In this case too, to find the desired value y, we measure the values of these quantities, and we use the results of these measurements to estimate y, i.e., we also apply data processing.

We want the most accurate data processing results, but our resources are limited. In all the above problems, we would like to get the estimate $\widetilde{y}$ which is as accurate as possible: we want the most accurate predictions of tomorrow's weather, we want to have the most accurate control of a petrochemical plant, etc. To get the more accurate estimates of y, we need to measure the auxiliary quantities $x_1, \ldots, x_n$ more accurately. However, high-accuracy measurements are more costly, and our resources are limited. So, we face the following problems:

- Within a limited budget, how accurately can we estimate the desired quantity y?
- When we want to estimate y with a given accuracy, how much money do we need?

2 Let Us Formulate Both Problems in Precise Terms

What do we need to do to formulate the problems in precise terms. In both problems, we want to find the relation between the cost of the corresponding measurements and the accuracy with which these measurements enable us to estimate y. So, to describe these problems in precise terms, we need to be able to describe accuracy of the measurements:

- we need to know how these accuracies affect the accuracy of estimating y, and
- we need to know how the cost of a measurement depends on its accuracy.

Let us analyze these questions one by one.

What we know about measurement accuracy. Measurements are never 100% accurate, so the result $\widetilde{x}_i$ of measuring the quantity x_i is, in general, different from the actual value x_i. The difference $\Delta x_i \stackrel{\text{def}}{=} \widetilde{x}_i - x_i$ is known as the *measurement error*.

In most cases, many different factors affect the measurement accuracy. As a result, the measurement error is the joint effect of many different independent factors. According to the Central Limit Theorem (see, e.g., [2]), when the number of these factors is large, the probability distribution of these errors is close to Gaussian (normal). In general, a normal distribution is uniquely determined by two parameters: mean and standard deviation.

Measuring instruments are usually *calibrated*: before the instrument is released, its readings are compared with some more accurate (*standard*) measuring instrument. The value $\widetilde{x}_{i,st}$ recorded by this more accurate measuring instrument is very close to the actual value x_i: $\widetilde{x}_{i,st} \approx x_i$. So, the difference $\widetilde{x}_i - \widetilde{x}_{i,st}$ between the measurement errors is very close to the measurement error $\Delta x_i = \widetilde{x}_i - x_i$:

$$\widetilde{x}_i - \widetilde{x}_{i,st} \approx \Delta x_i.$$

If the mean value of the measurement error is not 0, then, after a reasonable number of comparisons, the average value of the differences $\widetilde{x}_i - \widetilde{x}_{i,st}$ will be close to this mean. Once we realize that there is this mean difference, we can "recalibrate" the measuring instrument – namely, we can subtract this mean value (known as *bias*) from all the measurement results. For example, if a person knows that his/her watch is 5 minutes behind, this person can simply add 5 minutes to all the watch's reading – or, better yet, re-set the watch so that the watch will show the correct time.

Because of this calibration process that eliminates possible biases, we can safely assume that the mean value of the measurement error is 0. Thus, the only characteristic that describe the measurement error is the standard deviation σ_i.

Usually, for different measurements, measurement errors are caused by different factors. Thus, the random variables Δx_i describing different measurement errors are independent.

How measurement errors affect the accuracy with which we estimate y. We know that the desired quantity y is equal to $y = f(x_1, \ldots, x_n)$. So, if we knew the exact values of the quantities x_i, we would be able to apply the algorithm f and get the

exact value of y. As we have mentioned, in practice, we only know the measurement results $\widetilde{x}_i$ which are, in general, somewhat different from the actual values x_i. Based on the measurement results, we compute the estimate $\widetilde{y} = f\,(\widetilde{x}_1, \ldots, \widetilde{x}_n)$ for y. We want to find out how the estimation error $\Delta y \stackrel{\text{def}}{=} \widetilde{y} - y$ depends on the accuracies with which we measure x_i. Since we have argued that a reasonable description of each measurement accuracy is provided by the standard deviation σ_i, the question is: how the estimation accuracy depends on these standard deviations σ_i.

To answer this question, let us first describe the estimation error $\Delta y = \widetilde{y} - y$ in terms of the measurement errors Δx_i. Substituting the expressions for $\widetilde{y}$ and y into the formula that describes Δy, we conclude that

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

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, the above expression for Δy takes the following form:

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

Measurement errors Δx_i are usually small, so that terms which are quadratic (or of higher order) in terms of Δx_i can be safely ignored. Thus, we can expand the above expression in Taylor series in terms of Δx_i and keep only linear terms in this expansion. As a result, we get the following formula:

$$\Delta y = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\,(\widetilde{x}_1, \ldots \widetilde{x}_n)\,\Delta x_i,$$

i.e.,

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

where we denoted

$$c_i \stackrel{\text{def}}{=} \frac{\partial f}{\partial x_i}\,(\widetilde{x}_1, \ldots \widetilde{x}_n)\,.$$

According to the formula (1), the approximation error Δy is a linear combination of independent measurement errors Δx_i. As we have mentioned, each measurement error is normally distributed with mean 0 and standard deviation σ_i. It is known that a linear combination of several independent random variables is also normally distributed. Thus, we conclude that Δy is also normally distributed. As we have mentioned, the normal distribution is uniquely determined by its mean and its standard deviation σ.

The mean of a linear combination of random variables is equal to the linear combination of the means. Since the means of all the variables Δx_i are 0s, this implies that the mean of Δy is also 0.

The quantity Δy is the sum of n independent random variables $c_i \cdot \Delta x_i$. When we multiply a random variable by a constant c, its variance is multiplied by c^2. So, for each term $c_i \cdot \Delta x_i$, the variance is equal to $c_i^2 \cdot \sigma_i^2$. It is known that the variance

σ^2 of the sum of several independent random variables is equal to the sum of their variances. Thus, we conclude that

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

How the cost of measuring x_i depends on the measurement accuracy. A natural way to increase the accuracy is to repeat the measurement sereval times and take the average of the measurement results. How does this affect the accuracy?

Let us assume that we have a measuring instrument that measures the quantity x_i with mean 0 and standard deviation m_i. To increase the accuracy, we use this instrument n_i times, resulting in n_i measurement results $\widetilde{x}_{i,1}, \ldots, \widetilde{x}_{i,n_i}$, and then we return the average as an estimate for the desired value x_i:

$$\widetilde{x}_i = \frac{\widetilde{x}_{i,1} + \ldots + \widetilde{x}_{i,n_i}}{n_i}.$$

What is the accuracy of this estimate?

Subtracting the actual value x_i from both sides of this equality, we get the following expression for the measurement error $\Delta x_i = \widetilde{x}_i - x_i$:

$$\Delta x_i = \frac{\widetilde{x}_{i,1} + \ldots + \widetilde{x}_{i,n_i}}{n_i} - x_i.$$

To simplify this expression, we can add the two terms in the right-hand side and then re-order terms in the numerator:

$$\Delta x_i^{\text{av}} = \frac{\widetilde{x}_{i,1} + \ldots + \widetilde{x}_{i,n_i} - n_i \cdot x_i}{n_i} = \frac{\left(\widetilde{x}_{i,1} - x_i\right) + \ldots + \left(\widetilde{x}_{i,n_i} - x_i\right)}{n_i} =$$

$$\frac{\Delta x_{i,1} + \ldots + \Delta x_{i,n_i}}{n_i}, \tag{3}$$

where $\Delta x_{i,j} \stackrel{\text{def}}{=} \widetilde{x}_{i,j} - x_i$ is the measurement error of the j-th measurement.

The numerator of the formula (3) is the sum of n_i independent random variables, so its variance is equal to the sum of their variances. The variance of each measurement error $\Delta x_{i,j}$ is m_i^2, so the variance of the numerator is equal to $n_i \cdot m_i^2$. When we multiply a random variable by a constant c, its variance is multiplied by c^2. In our case,

$$c = \frac{1}{n_i},$$

thus, the variance σ_i^2 of the measurement error Δx_i is equal to

$$\sigma_i^2 = \left(\frac{1}{n_i}\right)^2 \cdot n_i \cdot m_i^2 = \frac{m_i^2}{n_i}. \tag{4}$$

What about the cost? Let d_i be the cost of each individual measurement. When we repeat the measurement n_i times, the overall cost of measuring the i-th quantity x_i is n_i times larger, i.e.,

$$D_i = n_i \cdot d_i \tag{5}$$

Let us analyze how the cost depends on accuracy. Once we are given the desired accuracy σ_i, we can find, from the formula (4), the value n_i corresponding to this accuracy:

$$n_i = \frac{m_i^2}{\sigma_i^2}.$$

Substituting this value n_i into the formula (5), we get

$$D_i = \frac{m_i^2}{\sigma_i^2} \cdot d_i,$$

i.e.,

$$D_i = \frac{C_i}{\sigma_i^2}, \tag{6}$$

where we denoted

$$C_i \stackrel{\text{def}}{=} m_i^2 \cdot d_i. \tag{7}$$

Comment. In some cases, to get a more accurate measurement, we explicitly repeat a measurement several times. An example is the way some devices for measuring blood pressure work: they measure blood pressure three times and take an average. Some measuring devices do it implicitly: e.g., super-precise clocks usually consist of several independent clocks, and the result is obtained by processing the reading of all these clocks.

In general, it is reasonable to assume that the dependence of measurement cost on measurement accuracy has the form (6)–(7). This allows us to answer a similar question about estimating y.

How the cost of estimating y depends on the estimation accuracy. If we measure each quantity x_i with accuracy σ_i, then the accuracy σ^2 of the resulting estimate for y is determined by the formula (2), and the overall measurement cost $D = D_1 + \ldots + D_n$ can be obtained by adding up the costs of all n measurements:

$$D = \sum_{i=1}^{n} \frac{C_i}{\sigma_i^2}. \tag{8}$$

So, we arrive at the following precise formulation of the above two problems.

What we want. In both problems, we want to find out the accuracies σ_i with which we need to measure the i-th quantity.

Formulating the first problem in precise terms. A limited budget means that our expenses D are bounded by some given value D_0. This means that we want to find

the values σ_i for which the estimation error (2) is the smallest possible under the constraint $D \le D_0$, i.e.,

$$\sum_{i=1}^{n} \frac{C_i}{\sigma_i^2} \le D_0. \tag{9}$$

Formulating the second problem in precise terms. In the second problem, we are given the accuracy σ^2 with which we want to estimate y. So, we want to find the values σ_i for which the cost (8) is the smallest possible under the constraint (2).

3 Within Limited Resources, What Is the Best Accuracy With Which We Can Estimate y?

Analysis of the problem. In this problem, to find the desired values σ_i, we minimize the expression (2) under the non-strict inequality constraint (9). The minimum of a function under non-strict inequality constraint is attained:

- either when the inequality is strict – in this case it is a local minimum of the objective function (2),
- or when the inequality becomes equality.

Let us show that the first case is not possible. Indeed, in the local minimum, all derivatives of the objective function (2) should be equal to 0. If we differentiate the expression (2) with respect to each unknown σ_i and equate the derivative to 0, we get $\sigma_i = 0$. In this case, the left-hand side of the inequality (9) is infinite, so this inequality is not satisfied.

Thus, the desired minimum occurs when the inequality (9) becomes an equality, i.e., when we have

$$\sum_{i=1}^{n} \frac{C_i}{\sigma_i^2} = D_0. \tag{10}$$

To minimize the objective function (2) under the equality constraint (10), we can use the Lagrange multiplier method, according to which this constraint optimization problem can be reduced, for an appropriate value λ, to the unconstrained optimization problem of minimizing the expression

$$\sum_{i=1}^{n} c_i^2 \cdot \sigma_i^2 + \lambda \cdot \left(\sum_{i=1}^{n} \frac{C_i}{\sigma_i^2} - D_0 \right), \tag{11}$$

where the value λ can be determined by the condition that the minimizing values σ_i satisfy the equality (10).

Differentiating the expression (11) with respect to σ_i and equating the derivative to 0, we conclude that

$$2c_i^2 \cdot \sigma_i - 2\lambda \cdot \frac{C_i}{\sigma_i^3} = 0,$$

i.e., equivalently,

$$2c_i^2 \cdot \sigma_i = 2\lambda \cdot \frac{C_i}{\sigma_i^3}.$$

To find an explicit expression for σ_i, we multiply both sides by σ_i^3 and divide both sides by $2c_i^2$. As a result, we get

$$\sigma_i^4 = \lambda \cdot \frac{C_i}{c_i^2},$$

thus

$$\sigma_i^2 = \sqrt{\lambda} \cdot \frac{\sqrt{C_i}}{|c_i|}. \tag{12}$$

To find λ, we substitute this expression into the equality (10) and get

$$D_0 = \sum_{j=1}^{n} \frac{C_j \cdot |c_j|}{\sqrt{\lambda} \cdot \sqrt{C_j}} = \frac{1}{\sqrt{\lambda}} \cdot \sum_{j=1}^{n} \left(\sqrt{C_j} \cdot |c_j|\right).$$

Thus,

$$\sqrt{\lambda} = \frac{1}{D_0} \cdot \sum_{j=1}^{n} \left(\sqrt{C_j} \cdot |c_j|\right).$$

Substituting this expression for $\sqrt{\lambda}$ into the formula (12), we get the following answer.

Optimal selection of measuring instruments. To get the best accuracy under limited resources, we need to select measuring instruments for which

$$\sigma_i^2 = \frac{1}{D_0} \cdot \sum_{j=1}^{n} \left(\sqrt{C_j} \cdot |c_j|\right) \cdot \frac{\sqrt{C_i}}{|c_i|}. \tag{13}$$

Substituting these expressions into the formula (2), we get the best accuracy σ^2 that we can achieve under this resource limitation:

$$\sigma^2 = \frac{1}{D_0} \cdot \sum_{j=1}^{n} \left(\sqrt{C_j} \cdot |c_j|\right) \cdot \sum_{i=1}^{n} \left(\sqrt{C_i} \cdot |c_i|\right),$$

i.e.,

$$\sigma^2 = \frac{1}{D_0} \cdot \left(\sum_{i=1}^{n} \left(\sqrt{C_i} \cdot |c_i|\right)\right)^2. \tag{14}$$

4 To Reach a Given Accuracy, What Amount of Resources Do We Need?

Analysis of the problem. In this problem, to find the desired values σ_i, we minimize the expression (8) under the constraint (2). To minimize the objective function (8) under the equality constraint (2), we can use the Lagrange multiplier method, according to which this constraint optimization problem can be reduced, for an appropriate value λ, to the unconstrained optimization problem of minimizing the expression

$$\sum_{i=1}^{n} \frac{C_i}{\sigma_i^2} + \lambda \cdot \left(\sum_{i=1}^{n} c_i^2 \cdot \sigma_i^2 - \sigma^2 \right), \tag{15}$$

where the value λ can be determined by the condition that the minimizing values σ_i satisfy the equality (2).

Differentiating the expression (15) with respect to σ_i and equating the derivative to 0, we conclude that

$$-2\frac{C_i}{\sigma_i^3} + 2\lambda \cdot c_i^2 \cdot \sigma_i = 0,$$

i.e., equivalently,

$$2\frac{C_i}{\sigma_i^3} = 2\lambda \cdot c_i^2 \cdot \sigma_i.$$

To find an explicit expression for σ_i, we multiply both sides by σ_i^3 and divide both sides by $2c_i^2$. As a result, we get

$$\sigma_i^4 = \frac{1}{\lambda} \cdot \frac{C_i}{c_i^2},$$

thus

$$\sigma_i^2 = \frac{1}{\sqrt{\lambda}} \cdot \frac{\sqrt{C_i}}{|c_i|}. \tag{16}$$

To find λ, we substitute this expression into the equality (2) and get

$$\sigma^2 = \sum_{j=1}^{n} c_j^2 \cdot \frac{1}{\sqrt{\lambda}} \cdot \frac{\sqrt{C_j}}{|c_j|} = \frac{1}{\sqrt{\lambda}} \cdot \sum_{j=1}^{n} \left(\sqrt{C_j} \cdot |c_j| \right).$$

Thus,

$$\frac{1}{\sqrt{\lambda}} = \frac{\sigma^2}{\sum\limits_{j=1}^{n} \left(\sqrt{C_j} \cdot |c_j| \right)}. \tag{17}$$

Substituting this expression into the formula (16), we get the following answer.

Optimal selection of measuring instruments. To get the smallest cost guaranteeing the given accuracy σ, we need to select measuring instruments for which

$$\sigma_i^2 = \frac{\sigma^2}{\sum\limits_{j=1}^{n}\left(\sqrt{C_j}\cdot|c_j|\right)}\cdot\frac{\sqrt{C_i}}{|c_i|}. \tag{18}$$

Substituting these expressions into the formula (8), we get the smallest cost that we can achieve to provide the desired accuracy:

$$D = \sum_{i=1}^{n}\frac{C_i\cdot|c_i|}{\sigma^2\cdot\sqrt{C_i}}\cdot\sum_{j=1}^{n}\left(\sqrt{C_j}\cdot|c_j|\right) = \frac{1}{\sigma^2}\cdot\sum_{i=1}^{n}\left(\sqrt{C_i}\cdot|c_i|\right)\cdot\sum_{j=1}^{n}\left(\sqrt{C_j}\cdot|c_j|\right),$$

i.e.,

$$D = \frac{1}{\sigma^2}\cdot\left(\sum_{i=1}^{n}\left(\sqrt{C_i}\cdot|c_i|\right)\right)^2. \tag{19}$$

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), and HRD-1834620 and HRD-2034030 (CAHSI Includes), and by the AT&T Fellowship in Information Technology.

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

References

1. S. G. Rabinovich, *Measurement Errors and Uncertainty: Theory and Practice*, Springer Verlag, New York, 2005.
2. D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman and Hall/CRC, Boca Raton, Florida, 2011.