Propagation of Error or Uncertainty

Marcel Oliver

December 4, 2015

1 Introduction

1.1 Motivation

All measurements are subject to error or uncertainty. Common causes are noise or external disturbances, imperfections in the experimental setup and the measuring devices, coarseness or discreteness of instrument scales, unknown parameters, and model errors due to simplifying assumptions in the mathematical description of an experiment. An essential aspect of scientific work, therefore, is quantifying and tracking uncertainties from setup, measurements, all the way to derived quantities and resulting conclusions.

In the following, we can only address some of the most common and simple methods of error analysis. We shall do this mainly from a calculus perspective, with some comments on statistical aspects later on.

1.2 Absolute and relative error

When measuring a quantity with true value x_{true}, the measured value x may differ by a small amount Δx. We speak of Δx as the *absolute error* or *absolute uncertainty* of x. Often, the magnitude of error or uncertainty is most naturally expressed as a fraction of the true or the measured value x. Since the true value x_{true} is typically not known, we shall define the *relative error* or *relative uncertainty* as $\Delta x/x$.

In scientific writing, you will frequently encounter measurements reported in the form $x = 3.3 \pm 0.05$. We read this as $x = 3.3$ and $\Delta x = 0.05$.

1.3 Interpretation

Depending on context, Δx can have any of the following three interpretations which are different, but in simple settings lead to the same conclusions.

1. An exact specification of a deviation in one particular trial (instance of performing an experiment), so that

$$x_{\text{true}} = x + \Delta x \, .$$

Of course, typically there is no way to know what this true value actually is, but for the analysis of error propagation, this interpretation is useful.

2. An error bound (which may or may not be strict). In this interpretation, we understand that x_{true} may lie anywhere in the interval $[x - \Delta x, x + \Delta x]$.

3. Uncertainty in the statistical sense. If repeated measurements of x are expected to follow a probability distribution, then Δx is usually taken as the *sample standard deviation* of a set of measurements.

Our focus is on analyzing how the error or uncertainty in the input of a calculation propagates through the calculation, leading to error or uncertainty in the output.

2 Error propagation in one variable

Suppose that x is the result of a measurement and we are calculating a dependent quantity

$$y = f(x)\,. \tag{1}$$

Knowing Δx, we must derive Δy, the associated error or uncertainty of y.

Let us recall the equation for the tangent line to f at point x,

$$\ell(x + \Delta x) = f(x) + f'(x)\,\Delta x\,. \tag{2}$$

We use the tangent line equation as a (linear) approximation to f; when Δx is not too large, we expect this approximation to be good. Thus,

$$\Delta y = f(x + \Delta x) - f(x) \approx \ell(x + \Delta x) - f(x) = f'(x)\,\Delta x\,. \tag{3}$$

If we think of Δx as a positive number which specifies the (expected) magnitude of error, we shall write the error propagation formula in the form

$$\Delta y \approx |f'(x)|\,\Delta x\,. \tag{4}$$

Example 1. If you measure $x = 49 \pm 4$, what should you report for $y = \sqrt{x}$ together with its uncertainty? To answer this question, we use (4) to compute

$$\Delta y \approx \frac{1}{2\sqrt{x}}\,\Delta x = \frac{1}{2\sqrt{49}} \cdot 4 = \frac{2}{7} \approx 0.3\,. \tag{5}$$

Thus, $y = 7 \pm 0.3$. (Note that it does not make sense to compute an estimate for Δy to a large number of significant digits as there is no basis for reporting the resulting error with great accuracy.)

Example 2 (Area of circle)**.** Suppose you can measure the radius of a circle with relative uncertainty $\Delta r/r = 5\%$. We ask for the associated relative uncertainty of the computed area

$$A = \pi\, r^2\,. \tag{6}$$

By (4), $\Delta A \approx \pi\, 2r\, \Delta r$. Thus, the relative uncertainty of the area is given by

$$\frac{\Delta A}{A} \approx \frac{\pi\, 2r\, \Delta r}{\pi\, r^2} = 2\,\frac{\Delta r}{r} = 10\%\,. \tag{7}$$

Example 3 (Allometric relation)**.** More generally, when y is allometrically related to x, i.e.,

$$y = k\, x^\alpha\,, \tag{8}$$

the same type of computation (fill in the details!) gives

$$\frac{\Delta y}{y} \approx \alpha\,\frac{\Delta x}{x}\,. \tag{9}$$

Example 4 (Exponential relation)**.** When y is exponentially related to x, i.e.,

$$y = c\,\mathrm{e}^{\lambda x}\,, \tag{10}$$

we find $\Delta y \approx |c\lambda|\, \mathrm{e}^{\lambda x}\, \Delta x$, so that

$$\frac{\Delta y}{y} \approx \frac{|c\lambda|\, \mathrm{e}^{\lambda x}\, \Delta x}{c\,\mathrm{e}^{\lambda x}} = |\lambda|\, \Delta x\,. \tag{11}$$

(We take $c > 0$ for simplicity.) Thus, in an exponential relation it is most natural to express the relative error of the output in terms of the absolute error of the input data!

3 Error propagation in several variables

3.1 Linear approximation and simple error propagation

Suppose, for simplicity, that $z = f(x,y)$ is a function of two variables. Once again, we may use linear approximation, where

$$\ell(x+\Delta x, y+\Delta y) = f(x,y) + \frac{\partial f}{\partial x}(x,y)\,\Delta x + \frac{\partial f}{\partial y}(x,y)\,\Delta y \tag{12}$$

so that

$$\Delta z \approx \ell(x+\Delta x, y+\Delta y) - \ell(x,y) = \frac{\partial f}{\partial x}(x,y)\,\Delta x + \frac{\partial f}{\partial y}(x,y)\,\Delta y\,. \tag{13}$$

As in Section 2, this leads to the error propagation formula

$$\Delta z \approx \left|\frac{\partial f}{\partial x}(x,y)\right| \Delta x + \left|\frac{\partial f}{\partial y}(x,y)\right| \Delta y\,. \tag{14}$$

This formula is in fact the best we can do when Δx and Δy represent true interval constraints on the error. It is, however, unnecessarily pessimistic in the case when the true error can be though of as lying in a "ball of uncertainty" with a known radius in the x-y plane. This is the case, for example, when measuring distances in the plane where the estimated uncertainty should not depend on the particular choice of coordinates. Formula (14), however, is not invariant under rotation of the coordinate axes.

3.2 A geometric view on error propagation[1]

Suppose you know that the true values x_{true} and y_{true} lie in an ellipse with center (x, y) and semi-axes Δx and Δy.[2] Then the maximum error in (13) is the maximum of

$$F(\xi, \eta) = a\,\xi + b\,\eta \tag{15}$$

with

$$a = \frac{\partial f}{\partial x}(x, y)\,\Delta x \qquad \text{and} \qquad b = \frac{\partial f}{\partial y}(x, y)\,\Delta y\,. \tag{16}$$

under the constraint that (ξ, η) is a point on the unit circle, i.e.,

$$g(\xi, \eta) \equiv \xi^2 + \eta^2 = 1\,. \tag{17}$$

This problem is easily solved using Lagrange multipliers. Setting up the Lagrange equations,

$$\frac{\partial F}{\partial \xi} = \lambda\,\frac{\partial g}{\partial \xi} \qquad \text{and} \qquad \frac{\partial F}{\partial \eta} = \lambda\,\frac{\partial g}{\partial \eta}\,, \tag{18}$$

we find $a = \lambda\,2\xi$ and $b = \lambda\,2\eta$, so that

$$\frac{a}{b} = \frac{\xi}{\eta}\,. \tag{19}$$

Using this relation to eliminate η from the constraint $\xi^2 + \eta^2 = 1$, we find

$$\xi^2 + \xi^2\,\frac{b^2}{a^2} = 1 \qquad \text{so that} \qquad \xi^2 = \frac{a^2}{a^2 + b^2}\,. \tag{20}$$

A similar expression is obtained by eliminating ξ and solving for η so that, taking square roots, we have

$$\xi = \frac{\pm a}{\sqrt{a^2 + b^2}} \qquad \text{and} \qquad \eta = \frac{\pm b}{\sqrt{a^2 + b^2}}\,. \tag{21}$$

Inserting this expression into (15) and maximizing the value of F by choosing the positive sign on each root, we obtain

$$F_{\max} = a\,\frac{a}{\sqrt{a^2 + b^2}} + b\,\frac{b}{\sqrt{a^2 + b^2}} = \frac{a^2 + b^2}{\sqrt{a^2 + b^2}} = \sqrt{a^2 + b^2}\,. \tag{22}$$

[1] Supplementary material, skip on first reading.

[2] The analogous geometric picture behind (14) is that x_{true} and y_{true} lie in a box with center (x, y) and semi-widths Δx and Δy.

Substituting the definitions of a and b back into this expression yields a formula for Δz which we write out in the next section.[3]

3.3 Propagation of independent uncertainties

The geometrically motivated computation in the previous section leads to the following error propagation formula for $z = f(x, y)$:

$$\Delta z \approx \sqrt{\left(\frac{\partial f}{\partial x}(x,y)\right)^2 \Delta x^2 + \left(\frac{\partial f}{\partial y}(x,y)\right)^2 \Delta y^2}\,. \tag{23}$$

The formula extends to more than two variables in the obvious way. Moreover, probability theory shows that this formula is also correct in the case where x and y are *independent random variables* with standard deviations Δx and Δy. A first-principles discussion is beyond the scope of this class, but the concept is very frequently used in practice. We will discuss some examples and implications.

Example 5. If you measure $x = 6.0 \pm 0.5$ and $y = 2.0 \pm 0.1$ independently, what should you report for $z = x/y$ together with its uncertainty? We compute

$$\frac{\partial z}{\partial x} = \frac{1}{y} \qquad \text{and} \qquad \frac{\partial z}{\partial y} = -\frac{x}{y^2} \tag{24}$$

and substitute these into (23) to find

$$\Delta z \approx \sqrt{\left(\frac{1}{2.0}\,0.5\right)^2 + \left(\frac{6.0}{2.0^2}\,0.1\right)^2} \approx 0.3\,. \tag{25}$$

Thus, $z = 3.0 \pm 0.3$.

Example 6 (Sum of independently measured quantities)**.** Suppose you measure x and y independently with uncertainties Δx and Δy. What is the uncertainty of their sum $z = x+y$? Here, the partial derivatives $\partial z/\partial x = \partial z/\partial y = 1$, so that

$$\Delta z \approx \sqrt{\Delta x^2 + \Delta y^2}\,. \tag{26}$$

This formula suggests that, geometrically, we may think of the absolute uncertainty Δz as the Euclidean length of the absolute uncertainty vector $(\Delta x, \Delta y)$.

[3]Using vector calculus, there is an alternative, much shorter argument. The estimate for Δz is the maximal rate of ascend of $\phi(\xi, \eta) = f(x + \xi\Delta x, y + \eta\Delta y)$, so it is given by the length of the gradient of ϕ. This directly yields (23). From this point of view, the argument in this section can be seen as a proof that the gradient is the vector pointing in the direction of steepest ascend, its magnitude giving the rate of change in this direction. This goes beyond what is covered in this class, but can be found in any standard text on vector calculus, e.g. [1, Chapter 13].

Example 7 (Product of independently measured quantities)**.** In the setting of Example 6, let us look at the product $z = xy$. Here, $\partial z/\partial x = y$ and $\partial z/\partial y = x$, so that

$$\Delta z \approx \sqrt{y^2\,\Delta x^2 + x^2\,\Delta y^2}\,. \tag{27}$$

Dividing through by z on both sides, we find that

$$\frac{\Delta z}{z} \approx \sqrt{\frac{\Delta x^2}{x^2} + \frac{\Delta y^2}{y^2}}\,. \tag{28}$$

In other words, when taking products, it is the *relative* uncertainty $\Delta z/z$ which is the Euclidean length of the *relative* uncertainty vector $(\Delta x/x, \Delta y/y)$.

Example 8 (Atwood machine)**.** The acceleration of two masses m_1 and m_2 in an Atwood machine[3] is given by the formula

$$a = g\,\frac{m_1 - m_2}{m_1 + m_2}\,, \tag{29}$$

where $g = 9.81\,\mathrm{m/s^2}$ is the constant of gravity. When $m_1 = 100 \pm 1\,\mathrm{g}$ and $m_2 = 50 \pm 1\,\mathrm{g}$, estimate the uncertainty Δa. After a short computation (fill in the details!), we find

$$\frac{\partial a}{\partial m_1} = g\,\frac{2\,m_2}{(m_1+m_2)^2} \qquad \text{and} \qquad \frac{\partial a}{\partial m_2} = g\,\frac{-2\,m_1}{(m_1+m_2)^2} \tag{30}$$

so that

$$\Delta z \approx \frac{2\,g}{(m_1+m_2)^2}\sqrt{(m_2\,\Delta m_1)^2 + (m_1\,\Delta m_2)^2}\,. \tag{31}$$

Inserting the numbers gives $a = 3.27\,\mathrm{m/s^2}$ and $\Delta a = 0.097\,\mathrm{m/s^2}$. We might want to report this result as $a = 3.27 \pm 0.1\mathrm{m/s^2}$.[4]

Example 9 (Volume of cylinder with independent errors)**.** Suppose that you measure the radius r and the height h of a cylinder each with an independent relative error of 1%. What is the relative error of the computed volume $V = \pi h r^2$? You could apply (23) directly and go from there (try it!), but at this point it is quicker to refer to Example 3 which shows that r^2 has a relative error of 2%. Further, the relative error of h and the relative error of r^2 add according to Example 7, so the relative error of V is $\sqrt{5}\%$. Thus, the relative error increases by a factor $\sqrt{5} \approx 2.24$.

Example 10 (Volume of cylinder with dependent errors)**.** Now suppose that the errors in the measurements are actually dependent. Such a situation might arise, for example, due to temperature sensitivity of the equipment which affects each measurement in exactly the same way. In this case, we need to refer to (14). Leaving out the details of the computation (try it!), we arrive at an increase of the relative error by a factor of 3.

[4] Example adapted from `https://courses.washington.edu/phys431/propagation_errors_UCh.pdf`. There, it is suggested to report the final result as $a = 3.3 \pm 0.1\mathrm{m/s^2}$, but then the reported confidence interval is clearly outside of the computed confidence interval, so I do not recommend rounding too aggressively.

Warning. The last two examples show that assuming independence of uncertainty when this is not actually guaranteed leads to an underestimate of the uncertainty of the result. This is dangerous as it may lead to conclusions which are not supported by the measured data.

Remark. When the quantities involved can be described in statistical terms, then (23) may be corrected by covariance terms which represent correlations between the different errors, see [4]. This, however, is beyond the scope of this class and also rarely used in practice as the required covariances are typically difficult to estimate.

3.4 Repeated experiments

Sometimes, an experiment is repeated many times and the values recorded are averaged over all trials. Suppose we have collected N readings $x_1, \ldots, x_N$ and we compute their mean

$$\overline{x} = \frac{x_1 + \cdots + x_N}{N} \,. \tag{32}$$

When an individual reading has uncertainty $\Delta x_1 = \cdots = \Delta x_N \equiv \Delta x$, what is the uncertainty of $\overline{x}$? Extending (14) to N variables, we find

$$\Delta\overline{x} \approx \frac{\Delta x_1}{N} + \cdots + \frac{\Delta x_N}{N} = \Delta x \,, \tag{33}$$

so there is no advantage in repeating the experiment. However, when the uncertainties are known to be independent, we may use (23), extended to N variables, instead, so that

$$\Delta\overline{x} \approx \sqrt{\left(\frac{\Delta x_1}{N}\right)^2 + \cdots + \left(\frac{\Delta x_N}{N}\right)^2} = \frac{\Delta x}{\sqrt{N}} \,. \tag{34}$$

Here, by taking the average over many measurements, errors partially cancel with high probability and the uncertainty of $\overline{x}$ converges to zero at a rate of $1/\sqrt{N}$ as the number of measurements goes to infinity.

Note that the benefit of repeated trials may be destroyed by unnoticed dependencies between successive measurements—it is not helpful to make the same error twice!

4 Monte-Carlo testing

Monte-Carlo testing is a computational technique for estimating the propagation of uncertainties. The idea is to perform the computation many times with randomly perturbed input and use the statistics of the output to determine its uncertainty. With Scientific Python/Pylab or similar mathematical software, Monte-Carlo testing is extremely easy and quick so that it should be part of the tool set of every working scientist. We describe the procedure in the setting of Example 5.

Rather than assigning a single value for each input variable, we create a vector of values that follow a probability distribution whose mean equals the measured value and whose standard deviation equals the corresponding uncertainty. Typically, we choose a normal distribution via the `normal` function; its first argument is the mean, the second argument is the standard deviation, and the third argument the number of samples which we set to a moderately large integer value:

```
In [1]: x = normal(6,0.5,10000)
In [2]: y = normal(2,0.1,10000)
```

Now we can perform the computation *with all samples at once*:

```
In [3]: z = x/y
```

The uncertainty of the output is then given by the standard deviation of z,

```
In [4]: std(z)
Out[4]: 0.29531173914816444
```

This corresponds well with the result suggested by our error propagation formula.

Monte-Carlo testing has a number of advantages: it can be done with ease in a routine manner, it does not use linearization, thus will also work as expected in situations when linear approximation is not appropriate, and it is possible to include dependencies between the different errors in the input statistics if necessary.

The drawbacks are that it requires many repeated computations. This is not an issue when testing simple function evaluations on contemporary computers as in the example above. However, when dealing with much more complex models, computational cost may become an issue. Another drawback is that the procedure is purely numerical. Thus, it is not possible to see how the output uncertainty depends on parameters of the setup, which is sometimes of interest. For this, the calculus-based error propagation formulas are better suited.

References

[1] H.J. Keisler, "Elementary Calculus," 2nd edition, Prindle, Weber & Schmidt, 1986. http://www.math.wisc.edu/~keisler/calc.html

[2] J.R. Taylor, "An Introduction to Error Analysis," 2nd edition, University Science Books, 1997. http://people.westminstercollege.edu/faculty/kphilippi/An_Introduction_To_Error_Analysis_The_Study_Of_Uncertainties_In_Physical_Measurements_Taylor_John.pdf

[3] "Atwood machine." Wikipedia, The Free Encyclopedia. December 6, 2015, 16:52 UTC. http://en.wikipedia.org/w/index.php?title=Atwood_machine&oldid=670698954

[4] "Propagation of uncertainty." Wikipedia, The Free Encyclopedia. December 5, 2015, 14:35 UTC. http://en.wikipedia.org/w/index.php?title=Propagation_of_uncertainty&oldid=693873085