

Computational Geology 17

The Total Differential and Error Propagation

H.L. Vacher, Department of Geology, University of South Florida, 4202 E. Fowler Ave., Tampa FL 33620

Topics this issue –

- *Mathematics: calculus – partial derivatives, Taylor's series.*
- *Geology: Darcy's Law; strike and dip.*

Introduction

The previous column (CG-16, "The Taylor Series and Error Propagation," May 2001) discussed the propagation of uncertainties through a function of one variable. Suppose, for example, that x is a quantity measured to have a value a , and that there is an uncertainty (ε_x , for "error") associated with that measurement. Suppose, further, that your work calls for $\sin(x)$, or $\log(x)$, or $\tanh(x)$. What uncertainty should you associate with $\sin(a)$, or $\log(a)$, or $\tanh(a)$? That is, using the notation

$$\varepsilon_f = f(a \pm \varepsilon_x) - f(a), \quad (1)$$

what is ε_f , the propagated uncertainty (propagated error)?

We saw that the propagated error could be expressed by the derivative terms of a Taylor's series. In other words, ε_f depends on ε_x and the derivatives df/dx , d^2f/dx^2 , d^3f/dx^3 , and so on, evaluated at $x = a$.

The subject of this column is the logical next step. Suppose there are two measured quantities, each with its own uncertainty, and you need to evaluate a function that depends on both of them. How do two (or more) uncertainties carry through to produce an uncertainty in the calculated result? The answer is a straightforward extension of concepts covered in CG-16 – Taylor's series, differentials and derivatives. The extension is to functions of more than one variable.

Important applications of these concepts include what may be considered to be the most basic operation of mathematics: arithmetic combinations of two numbers. That is, if x and y are two numbers and each has an associated uncertainty, what are the uncertainties of their sum, difference, product and quotient?

The Topographic analogy

Function of one variable. The topographic analogy we used in CG-16 for the one-dimensional case is reviewed in Figure 1. You are located at Point O on a hillside, which is shown on both a cross-section and map. You want to know the elevation at P . The topography

varies only in the east-west direction; that is, elevation is given by the function $h(x)$, and the contours of h run north-south. The slope of the hillside at O is c .

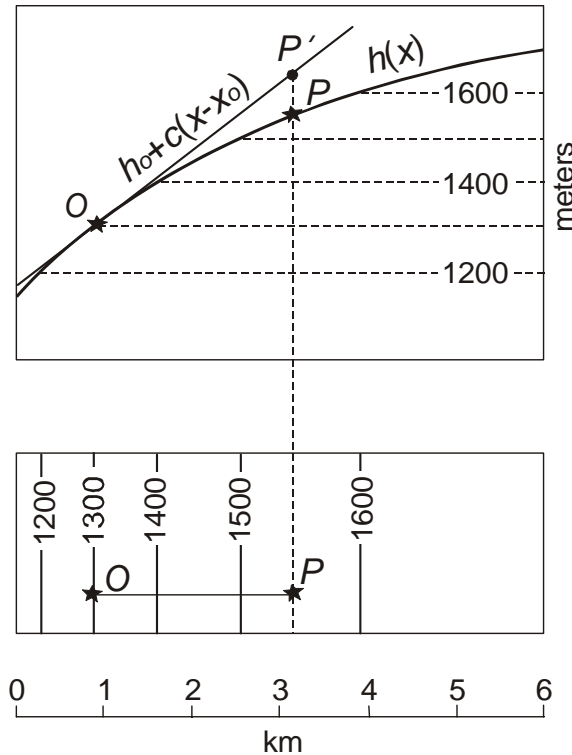


Figure 1. Cross section and map of a hillside function of one position variable, $h(x)$. The first-order term of Taylor's series produces the rise along a constant slope projected from O .

As discussed in CG-16, you can project that slope toward P and obtain an estimate for the elevation at P :

$$h(x_p) \approx h(x_o) + h'(x_o)(x_p - x_o), \quad (2)$$

where $h'(x_o)$ is c , the slope at Point O . Equation 2 says simply that the elevation at P is approximately the elevation at O plus the slope at O times the run (as in "slope equals rise over run") between O and P . The amount that this estimate misses the mark is illustrated by the vertical distance between P and P' in Figure 1. By just looking at the figure, one can see that this discrepancy goes away if the distance between O and P gets exceedingly small. When that happens, the tiny, tiny difference between $h(x_p)$ and $h(x_o)$, i.e., the differential dh , is

$$dh = h'(x)dx, \quad (3)$$

where the differential dx is the tiny, tiny difference in location between O and P . If the distance between O and P is not infinitesimally small, then we are left with finite differences rather than differentials:

$$\Delta h \approx h'(x)\Delta x . \tag{4}$$

Equation 4 is a rearranged version of Equation 2 with x_O and x_P of Equation 2 replaced by x and $x+\Delta x$, respectively

The discrepancy between P and P' , of course, is due to the fact that $h(x)$ is curved (Fig. 1) whereas Equation 2 assumes a straight line. The curve can be produced by the Taylor's Series (Equation 19 of CG-16),

$$h(x) = h(a) + (x - a)h'(a) + \frac{(x - a)^2}{2!}h''(a) + \dots + \frac{(x - a)^n}{n!}h^{(n)}(a) , \tag{5}$$

where x and a replace x_P and x_O , respectively, of Equation 2, and h'' and $h^{(n)}$ indicate the second and n^{th} derivatives, respectively. As discussed at length in CG-16, you get Equation 2 if you truncate Equation 5 before the second "plus" symbol. In other words, you can bring P' to P by adding terms to the Taylor's series.

Function of more than one variable. Now consider the difference in elevation between Points O and P on a hill in which the contours do not run north-south (Fig. 2). The elevation of the hill shown in Figure 2 varies with both x and y ; i.e., $h(x,y)$. The problem is the same: How do you express the difference in elevation between O and P in terms of slopes and horizontal distances?

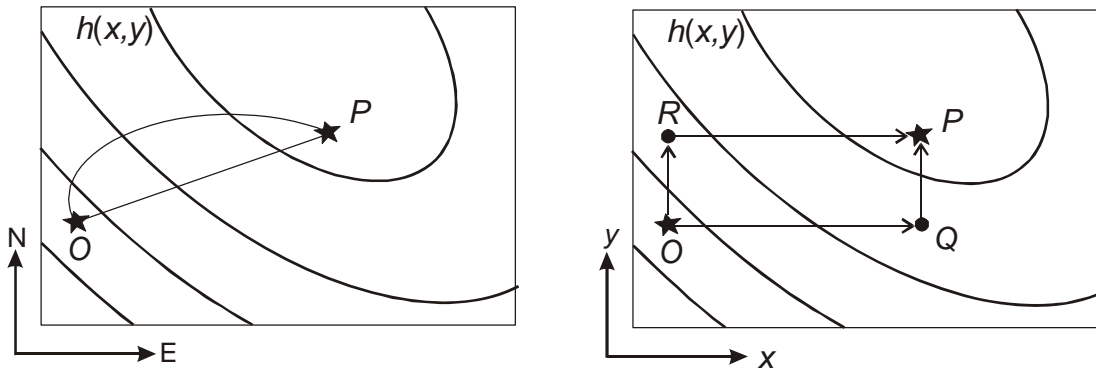


Figure 2. Routes between O and P on a hillside function of two position variables $h(x,y)$. Routes OQP and ORP are parallel to coordinate axes and can be used to give successive one-dimensional parts of the total rise from O and P .

Referring to the left panel of Figure 2, it is clear that the difference in elevation between the beginning point and the end point is Δh whether you walk from O to P along the straight line or along the curved line. The same is true if you walk either of the two routes OQP or ORP shown in the right panel of Figure 2. The advantage to these particular two routes is mathematical: you can apply Equations 3, 4 and 5 because the legs are parallel to coordinate axes.

Using OQP , the *total difference* in elevation, Δh , consists of two parts,

$$\Delta h = \Delta h_{OQ} + \Delta h_{QP}, \quad (6a)$$

where the first term is the difference in elevation between O and Q , and the second term is the difference in elevation between Q and P . From Equation 4, Δh_{OQ} can be estimated as the *rate of change of h in the east direction* at Point O (a slope) times the east-west distance between O and Q . In the same way, the route ORP produces

$$\Delta h = \Delta h_{OR} + \Delta h_{RP} \quad (6b)$$

for the total difference in elevation. From Equation 4, Δh_{OR} can be estimated as the *rate of change of h in the north direction* at Point O (a slope) times the south-to-north distance between O and R .

The italicized wording in the last paragraph referring to the rate of change of h in the eastward and northward directions is crucial. For functions of more than one variable, the slope (derivative) depends on direction, and therefore, one speaks of *directional derivatives* in multivariable calculus. In Figure 2, for example, the directional derivative to the NE at O is clearly larger than either the directional derivative to the E at O or the directional derivative to the N at O . These latter two directional derivatives parallel to the x and y -axes are particularly important and are given a special symbol and name. Thus $\partial h/\partial x$ is the *partial derivative* of h with respect to x holding y constant – the slope of the hillside in the east-to-west direction. Similarly, $\partial h/\partial y$ is the slope of the hillside in the south-to-north direction (holding x , or east, constant). Partial derivatives are the key elements of multivariable differential calculus.

Partial derivatives. The partial derivative $\partial f/\partial x$ at (x,y) which can also be written $f_x(x,y)$, is defined as

$$f_x(x,y) \equiv \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}. \quad (7)$$

Compare this to the definition of the derivative of single-variable calculus: df/dx at x , which can also be written $f'(x)$, is

$$f'(x) \equiv \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}. \quad (8)$$

In Equation 7, the numerator refers to the difference in the value of a function at two addresses on a *two-dimensional xy -grid* (lower left panel of Fig. 3): the base point, with address (x,y) , which is fixed in location; and a second point, which lies east of the base point by a distance Δx . This second point moves parallel to the x -axis, toward the base point, as the limit is taken. Therefore, the value of the of the partial derivative $\partial f/\partial x$ defined by Equation 7 is referred to the base point (x,y) . In the more-familiar definition of Equation 8, the numerator refers to the

difference in the value of a function at two addresses along a *one-dimensional line* (upper panel of Fig. 3). These points are

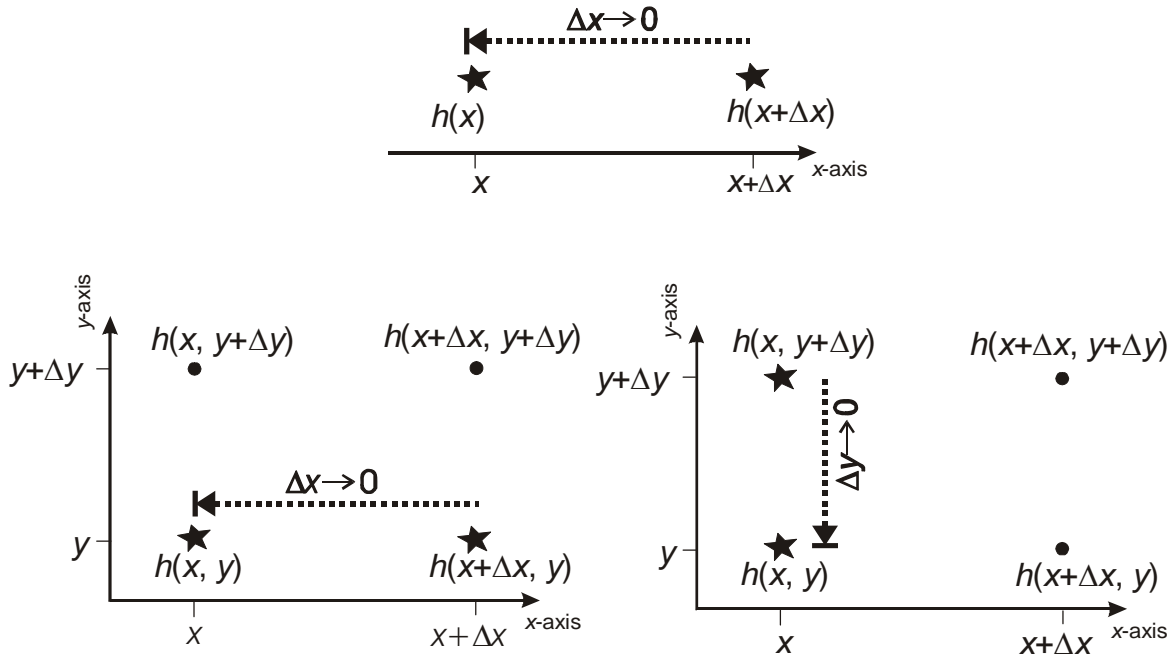


Figure 3. Comparison of the limit directions used in the derivative of a function of a single variable (upper panel) and the partial derivatives of a function of two variables (lower panels).

separated by the distance Δx , and the second point moves toward the base point (x) along the x -axis (the only way it can move) as the limit is taken. The derivative df/dx defined by Equation 8, therefore, is referred to the base point (x).

In the same way, the partial derivative with respect to y at (x,y) is

$$f_y(x, y) = \frac{\partial f}{\partial y} = \lim_{\Delta y \rightarrow 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y} . \quad (9)$$

The numerator refers to the value of f at two points, the base point (x,y) and a second, movable, point located Δy due *north* of the base point (lower right panel of Fig. 3). The second point moves southward toward the (x,y) as the limit is taken.

Finding partial derivatives of multivariable functions is no more difficult than finding derivatives of single-variable functions. For example, consider the function

$$f(x, \alpha) = x \sin \alpha . \quad (10a)$$

For the partial derivative $\partial f / \partial x$, hold the $\sin(\alpha)$ constant (e.g., in your mind, replace it with the number 2, say, so that the function would be $2x$) and then simply differentiate with respect to x . Thus

$$f_x(x, \alpha) = \sin \alpha . \quad (10b)$$

Similarly, for the partial derivative $\partial/\partial y$, hold the x constant, and differentiate $\sin(\alpha)$ with respect to α :

$$f_\alpha(x, \alpha) = x \cos \alpha . \quad (10c)$$

As a second example, consider the function

$$f(x, \alpha, \gamma) = x \sin \alpha \tan \gamma . \quad (11a)$$

The first partial derivatives are

$$f_x(x, \alpha, \gamma) = \sin \alpha \tan \gamma , \quad (11b)$$

$$f_\alpha(x, \alpha, \gamma) = x \cos \alpha \tan \lambda , \quad (11c)$$

and $f_\gamma(x, \alpha) = x \sin \alpha / \cos^2 \gamma . \quad (11d)$

We will use these partial derivatives (Equations 11b-d) later in a strike and dip problem. We will use even easier partial derivatives in discussing what happens to errors when they are taken through arithmetic operations.

The total differential. Returning now to Equations 6, the difference in elevation along each of the rectilinear legs between the points at O and P (Fig. 2) can be written down from analogy with Equation 4. Thus

$$\Delta h_{OQ} \approx h_x(x, y)\Delta x \quad . \quad (12a)$$

$$\Delta h_{QP} \approx h_y(x + \Delta x, y)\Delta y \quad (12b)$$

$$\Delta h_{OR} \approx h_y(x, y)\Delta y \quad (12c)$$

and $\Delta h_{RP} \approx h_x(x, y + \Delta y)\Delta x . \quad (12d)$

Substituting these values into Equations 6 and 7 produces

$$\Delta h = h_x(x, y)\Delta x + h_y(x + \Delta x, y)\Delta y \quad (13a)$$

and $\Delta h = h_y(x, y)\Delta y + h_x(x, y + \Delta y)\Delta x , \quad (13b)$

respectively, for the total difference in elevation between O and P . It should be noted (Fig. 2), that at the scale of these finite differences, the differences in elevation along parallel legs are not the same. That is

$$\Delta h_{OQ} \neq \Delta h_{RP} \quad \text{and} \quad \Delta h_{QP} \neq \Delta h_{OR} \quad (14)$$

If the distance between O and P is exceeding small, however, the finite differences represented by Δx , Δy , Δh become differentials represented by dx , dy , and dh . At the fine scale

of differentials, curves become straight lines and curved surfaces become sloping planes (Fig. 4). Because the surface is now a plane, Inequalities 14 become

$$\Delta h_{OQ} = \Delta h_{RP} \quad (15a)$$

and

$$\Delta h_{QP} = \Delta h_{OR} \quad (15b)$$

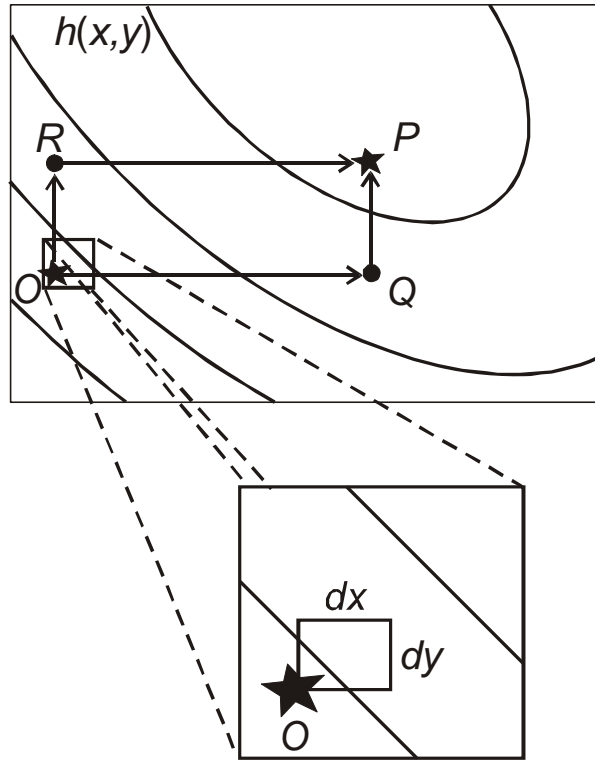


Figure 4. The finite differences OQ and OR become differentials dx and dy as the area of interest focuses more intensely on the vicinity of O . In the process, the curved surface in the field of view becomes a plane, as seen by the straight, as opposed to curved, contours.

In particular, Equation 6a, which uses the finite differences in Equations 12a and 12b, can be dealt with in terms of Equations 12a and 12c because of Equation 15b. Then, with differentials replacing differences, and using Equation 3 (because the curved cross profiles are now straight at this fine scale), Equation 6a becomes

$$dh = dh_{OQ} + dh_{OR} = h_x(x, y)dx + h_y(x, y)dy . \quad (16a)$$

Similarly, Equation 15a allows Equation 12d to be traded out for Equation 12a so that, with differentials replacing differences, Equation 6b becomes

$$dh = dh_{OR} + dh_{OQ} = h_y(x, y)dy + h_x(x, y)dx . \quad (16b)$$

Equations 16a and 16b are, of course , the same. The result, then, is that at the scale of differentials,

$$dh = h_x(x, y)dx + h_y(x, y)dy , \quad (17a)$$

or, as it is often written

$$dh = \frac{\partial h}{\partial x} dx + \frac{\partial h}{\partial y} dy . \quad (17b)$$

Equations 17 do for functions of two variables what Equation 3 does for functions of one variable. It is one of the most useful relationships in calculus. The quantity dh is the *total differential*. As shown by Equations 17, the total differential is the sum of partial derivatives weighted by the differentials of the corresponding independent variables.

Taylor's series. Equation 2 predicts the elevation of P from the elevation of O for the case where h depends only on x . The analogous expression for the case where h depends on both x and y is

$$h(x_P, y_P) \approx h(x_O, y_O) + f_x(x_O, y_O)(x_P - x_O) + f_y(x_O, y_O)(y_P - y_O) . \quad (18)$$

In language similar to that used in describing Equation 2, Equation 18 says that the elevation at P is approximately the elevation of O plus the west-to-east slope at O times the west-to-east distance between O and P , plus the south-to-north slope at O times the south-to-north distance between O and P . Thus Equation 18 estimates the elevation of P from the elevation at O by projecting h along a *plane* that has west-to-east and south-to-north slopes of $\partial h/\partial x$ and $\partial h/\partial y$, respectively.

Projecting the elevation along such a plane will miss the mark for the same reason that projecting along a line misses the mark in the single-variable case (Fig. 1) – the projection does not account for curvature. The curvature can be taken into account by adding terms with higher-order derivatives. Thus, with the next set of derivatives, Equation 18 becomes (Solkonikoff and Redheffer, 1966, p. 337)

$$h(x, y) \approx h(a, b) + (x - a)h_x(a, b) + (y - b)h_y(a, b) + \frac{(x - a)^2 h_{xx}(a, b) + 2(x - a)(y - b)h_{xy}(a, b) + (y - b)^2 h_{yy}(a, b)}{2!} \quad (19)$$

In this notation, x and y replace x_O and y_O respectively; a and b replace x_P and y_P , respectively; and the second-order derivatives are

$$h_{xx} \equiv \frac{\partial}{\partial x} \left(\frac{\partial h}{\partial x} \right) \equiv \frac{\partial^2 h}{\partial x^2} , \quad (20a)$$

$$h_{yy} \equiv \frac{\partial}{\partial y} \left(\frac{\partial h}{\partial y} \right) \equiv \frac{\partial^2 h}{\partial y^2}, \quad (20b)$$

$$h_{xy} \equiv \frac{\partial}{\partial y} \left(\frac{\partial h}{\partial x} \right) \equiv \frac{\partial^2 h}{\partial y \partial x}, \quad (20c)$$

and

$$h_{yx} \equiv \frac{\partial}{\partial x} \left(\frac{\partial h}{\partial y} \right) \equiv \frac{\partial^2 h}{\partial x \partial y}. \quad (20d)$$

Almost always (meaning whenever the derivatives are continuous at the point in question; Solkonikoff and Redheffer, 1966, p. 317), the mixed derivatives are equal

$$h_{xy} = h_{yx}. \quad (20e)$$

This fact is taken into account in Equation 19, which is the Taylor's series expansion for a function of two independent variables through the second-order terms.

Equation 19 should be compared to the Equation 5 through the $h''(a)$ term. The first two second-order derivatives, h_{xx} and h_{yy} (Equations 20a and 20b) treat the eastward change in the eastward slope and the northward change in the northward slope, respectively. The mixed second-order derivatives, h_{xy} and h_{yx} treat the northward change of the eastward slope, and the eastward change of the northward slope, respectively.

As discussed in CG-16, adding more terms to the Taylor's series improves the estimate of the function at points P far removed from the base point, O . On the other hand, the higher-order terms drop out as P approaches O . Thus for small projections, and using a general function $f(x,y)$, Equation 19 reduces to

$$\Delta f \approx \frac{\partial f}{\partial x} \Delta x + \frac{\partial f}{\partial y} \Delta y. \quad (21)$$

This is a form of the Taylor's series to first-order terms for functions of two variables. It is generally considered as the finite-difference version of Equation 17b.

Arithmetic and Error Propagation

Now consider four simple functions of two variables:

$$f(x,y) = x + y, \quad (22a)$$

$$f(x,y) = x - y, \quad (22b)$$

$$f(x,y) = xy, \quad (22c)$$

and

$$f(x,y) = x/y. \quad (22d)$$

Suppose x and y are two measurement variables, and suppose their measured values are a and b , respectively; that is, $x = a$ and $y = b$. The result of adding the two numbers is

$$f(x,y) = f(a,b) = a + b \quad (23a)$$

Similarly, the results of subtracting, multiplying and dividing the numbers are

$$f(x,y) = f(a,b) = a - b, \quad (23b)$$

$$f(x,y) = f(a,b) = ab, \quad (23c)$$

and $f(x,y) = f(a,b) = a/b. \quad (23d)$

Finally, suppose there is an uncertainty (error) associated with each of the measured values. What is the uncertainty propagated to the sum (Equation 23a), difference (23b), product (23c) and ratio (23d) of the two measured values.

Let $\pm \varepsilon_x$ and $\pm \varepsilon_y$ represent the error in a and b , respectively. The problem is to find ε_f , the propagated uncertainty. In other words, from

$$\varepsilon_f = f(a \pm \varepsilon_x, b \pm \varepsilon_y) - f(a, b), \quad (24)$$

what is ε_f ? We can use either Equation 19 or 21, depending on the size of the uncertainties. In either case, we need the partial derivatives.

Addition. The partial derivatives of the addition function (Equation 22a) are

$$\frac{\partial f}{\partial x} = 1 \quad \text{and} \quad \frac{\partial f}{\partial y} = 1. \quad (25a)$$

The second- and higher-order derivatives are all zero. Equation 19 for $f(x,y)$, therefore, becomes

$$f(x, y) = f(a, b) + (x - a) + (y - b). \quad (25b)$$

For addition, the maximum of the range of $f(a,b) \pm \varepsilon_f$ will occur at $x = a + \varepsilon_x$, and $y = b + \varepsilon_y$. Substituting these values for x and y into Equation 25b produces

$$f(a + \varepsilon_x, b + \varepsilon_y) = f(a, b) + \varepsilon_x + \varepsilon_y. \quad (25c)$$

The minimum of the range of $f(a,b) \pm \varepsilon_f$ will occur at $x = a - \varepsilon_x$, and $y = b - \varepsilon_y$. With these values of x and y , Equation 25b becomes

$$f(a - \varepsilon_x, b - \varepsilon_y) = f(a, b) - \varepsilon_x - \varepsilon_y. \quad (25d)$$

From Equations 24, 25c, and 25d,

$$\varepsilon_f = \pm(\varepsilon_x + \varepsilon_y). \quad (25e)$$

Subtraction. The partial derivatives of the subtraction function (Equation 22b) are

$$\frac{\partial f}{\partial x} = 1 \quad \text{and} \quad \frac{\partial f}{\partial y} = -1. \quad (26a)$$

The second- and higher-order derivatives are all zero. Equation 19 for $f(x,y)$, therefore, becomes

$$f(x, y) = f(a, b) + (x - a) - (y - b) \quad . \quad (26b)$$

as in addition (Equation 25b).

The maximum of the range $f(a,b) \pm \varepsilon_f$ will occur at $x = a + \varepsilon_x$, and $y = b - \varepsilon_y$. With these values of x and y , Equation 26b becomes

$$f(a + \varepsilon_x, b - \varepsilon_y) = f(a, b) + \varepsilon_x + \varepsilon_y \quad , \quad (26c)$$

as in addition (Equation 25c)

The minimum of the range $f(a,b) \pm \varepsilon_f$ will occur at $x = a - \varepsilon_x$, and $y = b + \varepsilon_y$, and so, Equation 26b becomes

$$f(a - \varepsilon_x, b + \varepsilon_y) = f(a, b) - \varepsilon_x - \varepsilon_y \quad (26d)$$

as in addition (Equation 25d)

From Equations 24, 26c, and 26d,

$$\varepsilon_f = \pm(\varepsilon_x + \varepsilon_y), \quad (26e)$$

as in addition (Equation 25e).

Multiplication. The first-order partial derivatives of the multiplication function (Equation 22c) are

$$\frac{\partial f}{\partial x} = y \quad \text{and} \quad \frac{\partial f}{\partial y} = x. \quad (27a)$$

The second-order partial derivatives are

$$\frac{\partial^2 f}{dx^2} = 0, \quad \frac{\partial^2 f}{dy^2} = 0, \quad \text{and} \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = 1. \quad (27b)$$

Third and higher-order derivatives are zero. Because the derivatives of Equations 27a are evaluated at $x = a$ and $y = b$, Equation 19 for $f(x,y)$ then becomes

$$f(x, y) = f(a, b) + (x - a)b + (y - b)a + (x - a)(y - b). \quad (27c)$$

The maximum of the range $f(a,b) \pm \varepsilon_f$ will occur at $x = a + \varepsilon_x$, and $y = b + \varepsilon_y$. With these values, Equation 27c becomes

$$f(a + \varepsilon_x, b + \varepsilon_y) = f(a, b) + \varepsilon_x b + \varepsilon_y a + \varepsilon_x \varepsilon_y. \quad (27d)$$

The minimum of the range $f(a,b) \pm \varepsilon_f$ will be at $x = a - \varepsilon_x$, and $y = b - \varepsilon_y$. With these values Equation 27c becomes

$$f(a - \varepsilon_x, b - \varepsilon_y) = f(a, b) - \varepsilon_x b - \varepsilon_y a + \varepsilon_x \varepsilon_y. \quad (27e)$$

From Equations 24, 27d, and 27e,

$$\varepsilon_f = \pm(\varepsilon_x b + \varepsilon_y a) + \varepsilon_x \varepsilon_y. \quad (27f)$$

Dividing both sides by ab , which is $f(a,b)$, gives the result

$$\frac{\varepsilon_f}{ab} = \pm \left(\frac{\varepsilon_x}{a} + \frac{\varepsilon_y}{b} \right) + \frac{\varepsilon_x \varepsilon_y}{ab}. \quad (27g)$$

Division. The first-order partial derivatives of the division function (Equation 22d) are

$$\frac{\partial f}{\partial x} = \frac{1}{y} \quad \text{and} \quad \frac{\partial f}{\partial y} = -\frac{x}{y^2}. \quad (28a)$$

The second-order partial derivatives are

$$\frac{\partial^2 f}{dx^2} = 0, \quad \frac{\partial^2 f}{\partial y^2} = \frac{2x}{y^3}, \quad \text{and} \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = -\frac{1}{y^2}. \quad (28b)$$

Not all third- and higher-order terms are zero. Stopping with the second-order terms, and evaluating the derivatives of Equations 28a and 28b at $x = a$ and $y = b$, Equation 19 for $f(x,y)$ becomes

$$f(x, y) \approx f(a, b) + (x - a)/b - (y - b)a/b^2 - (x - a)(y - b)/b^2 + (y - b)^2 a/b^2. \quad (28c)$$

The maximum of the range $f(a,b) \pm \varepsilon_f$ will occur at $x = a + \varepsilon_x$, and $y = b - \varepsilon_y$. With these values, Equation 28c becomes

$$f(a + \varepsilon_x, b - \varepsilon_y) = f(a, b) + \frac{\varepsilon_x}{b} + \frac{\varepsilon_y a}{b^2} + \frac{\varepsilon_x \varepsilon_y}{b^2} + \frac{\varepsilon_y^2 a}{b^3}. \quad (28d)$$

The minimum of the range $f(a,b) \pm \varepsilon_f$ will be at $x = a - \varepsilon_x$, and $y = b + \varepsilon_y$. With these values, Equation 28c becomes

$$f(a - \varepsilon_x, b + \varepsilon_y) \approx f(a, b) - \frac{\varepsilon_x}{b} - \frac{\varepsilon_y a}{b^2} + \frac{\varepsilon_x \varepsilon_y}{b^2} + \frac{\varepsilon_y^2 a}{b^3}. \quad (28e)$$

From Equations 24, 28d, and 28e,

$$\varepsilon_f \approx \pm \left(\frac{\varepsilon_x}{b} + \frac{\varepsilon_y a}{b^2} \right) + \left(\frac{\varepsilon_x \varepsilon_y}{b^2} + \frac{\varepsilon_y^2 a}{b^3} \right). \quad (28f)$$

Dividing both sides by a/b , which is $f(a,b)$, gives the result

$$\frac{\varepsilon_f}{a/b} \approx \pm \left(\frac{\varepsilon_x}{a} + \frac{\varepsilon_y}{b} \right) + \left(\frac{\varepsilon_x \varepsilon_y}{a b} + \frac{\varepsilon_y^2}{b^2} \right). \quad (28g)$$

Some Examples

Let $x = 40 \pm 3$ and $y = 10 \pm 1$ be two measured quantities. What are the results of combining these two measurements by addition, subtraction multiplication and division?

For $x = 40 \pm 3$, $a = 40$, $\varepsilon_x = 3$, and $\varepsilon_x/a = 0.075$. For $y = 10 \pm 1$, $\varepsilon_y = 1$, and $\varepsilon_y/b = 0.1$. With $a = 40$, and $b = 10$: $f(a,b)$ for addition is 40; $f(a,b)$ for subtraction is 30, $f(a,b)$ for multiplication is 400, and $f(a,b)$ for division is 4. What is ε_f for each of these calculated values?

Addition. By straightforward arithmetic, the range is given by

$$\begin{aligned} \text{maximum sum} &= (40 + 3) + (10 + 1) = 54; \\ \text{minimum sum} &= (40 - 3) + (10 - 1) = 46. \end{aligned}$$

From these results the range is obviously 50 ± 4 .

From application of Equation 25e,

$$\varepsilon_f = \pm(\varepsilon_x + \varepsilon_y) = \pm(3 + 1) = \pm 4$$

Subtraction. By simple arithmetic, the range is

$$\begin{aligned} \text{maximum difference} &= (40 + 3) - (10 - 1) = 34; \\ \text{minimum difference} &= (40 - 3) - (10 + 1) = 26. \end{aligned}$$

From these results the range is 30 ± 4 .

From Equation 26e,

$$\varepsilon_f = \pm(\varepsilon_x + \varepsilon_y) = \pm(3 + 1) = \pm 4$$

Multiplication. By arithmetic, the range is

$$\begin{aligned} \text{maximum product} &= (40 + 3) * (10 + 1) = 473; \\ \text{minimum product} &= (40 - 3) * (10 - 1) = 333. \end{aligned}$$

From these results the range is 400_{-67}^{+73} .

From Equation 27g,

$$\begin{aligned}\varepsilon_f &= \pm ab \left(\frac{\varepsilon_x}{a} + \frac{\varepsilon_y}{b} \right) + ab \left(\frac{\varepsilon_x \varepsilon_y}{a b} \right) \\ &= \pm 400 (0.175) + 400 (0.0075) = \pm 70 + 3.\end{aligned}$$

Division. By arithmetic, the range is

$$\begin{aligned}\text{maximum ratio} &= (40 + 3) \div (10 - 1) = 4.7778 \text{ (rounded)} \\ \text{minimum ratio} &= (40 - 3) \div (10 + 1) = 3.3636 \text{ (rounded)}\end{aligned}$$

From these results the range is $4^{+0.7778(\text{rounded})}_{-0.6363(\text{rounded})}$.

From Equation 28g,

$$\varepsilon_f \approx \pm \frac{a}{b} \left(\frac{\varepsilon_x}{a} + \frac{\varepsilon_y}{b} \right) + \frac{a}{b} \left(\frac{\varepsilon_x \varepsilon_y}{a b} + \frac{\varepsilon_y^2}{b^2} \right)$$

from which

$$\begin{aligned}\varepsilon_f &\approx \pm 4 (0.175) + 4 (0.0175) \\ \varepsilon_f &\approx \pm 0.7 + 0.07.\end{aligned}$$

This result duplicates the result we found from simple arithmetic to the first two figures.

If you enjoy this sort of thing, you might conjecture that the higher-order results for this example could be found from

$$\varepsilon_f = \pm 0.7 + 0.07 \pm 0.007 + 0.0007 \pm \dots$$

Then you might like to prove it by multiplying out the right hand side of

$$(40 \pm 3)(10 \pm 1)^{-1} = 4(1 \pm 0.075)(1 \pm 0.1)^{-1}$$

using the ever-useful binomial series for the last factor,

$$(1 \pm x)^{-1} = 1 \mp x + x^2 \mp x^3 + x^4 \mp x^5 + \dots$$

After that, you might like to go back to Equations 19 and 20 and conjecture what the third- and fourth-order terms might be and check yourself with this example.

Summary of Rules for Calculating Propagated Error

The basic rule of error propagation (Taylor, 1997) is

$$\varepsilon_f \approx \pm \left(\left| \frac{\partial f}{\partial x} \right| \varepsilon_x + \left| \frac{\partial f}{\partial y} \right| \varepsilon_y \right) \quad (29)$$

which is simply the first-order terms of the Taylor's series for two variables (Equation 21). The absolute values of the partial derivatives are used to assure that ε_f gives the largest range (as we did in the case of subtraction and division by using $+\varepsilon_x$ and $-\varepsilon_y$ for the maximum value and $-\varepsilon_x$ and $+\varepsilon_y$ for the minimum value). For addition and subtraction Equation 29 gives ε_f exactly. For multiplication and division, Equation 29 is close only if ε_x and ε_y are small relative to a and b – which seems reasonable if ε_x and ε_y are supposed to represent errors.

It is customary to distinguish between absolute and relative errors in discussing error propagation in arithmetic operations (Taylor, 1997). In the example of $x = 40 \pm 3$, and $y = 10 \pm 1$, the absolute errors are 3 and 1, respectively (i.e., ε_x and ε_y). The relative errors are 0.075 and 0.1, respectively (i.e., ε_x/x and ε_y/y). The relative errors can also be stated as a percentage. For example, the x -value in the example is $40 \pm 7.5\%$, and the y -value is $10 \pm 10\%$.

For addition and subtraction, the rule is that absolute errors add (Equations 25e and 26e, respectively). Thus, *the absolute propagated error for the sum and for the difference of two numbers is the sum of the absolute errors of those numbers.*

For multiplication and division, the rule is cast in terms of relative errors. With the assumption that the relative errors of the two numbers are small fractions, Equations 27g and 28g reduce to the same thing, namely

$$\frac{\varepsilon_f}{ab} = \frac{\varepsilon_f}{a/b} = \pm \left(\frac{\varepsilon_x}{a} + \frac{\varepsilon_y}{b} \right), \quad (30)$$

because higher-order terms drop out. Thus for multiplication and division involving small relative errors, *the relative propagated error for the product and for the quotient of two numbers is the sum of the relative errors for those numbers.*

Equation 30 is the basic rule for multiplication and division. It results in a symmetric error term (e.g., ± 70 for multiplication and ± 0.7 for division in the previous example). If the relative uncertainties are large – or if one wants the propagated uncertainty to more digits – then Equation 30 is not sufficient. Higher terms are required, and they result in an asymmetric error term. For multiplication, the propagated error is found from one more term (involving the product of the relative errors, Equation 27g). For division, the propagated error requires an infinite number of terms to pin it down precisely.

Beware Subtraction

One wants to assume that the error terms going into a calculation are small. After all, who wants to think of a number such as 10 ± 6 , or 60% relative error?

Such errors can easily arise from subtracting measured values. For an easy example, consider 80 ± 4 and 70 ± 2 . The difference is 10 ± 6 by the rule of summing absolute errors. But look at what happens to the relative error. The first number is $80 \pm 5\%$ of 80, and the second number is $70 \pm \sim 3\%$ of 70. In subtraction, these combine to produce a number with 60% relative error!

Now suppose you want to multiply or divide two such numbers. We will take multiplication, because it produces a correct result with one additional term. Suppose you multiply 10 ± 6 by 8 ± 4 . You know from straightforward arithmetic that the range is 16 (i.e., $4 \cdot 4$) to 192 (i.e., $16 \cdot 12$). But the basic rule of error propagation for multiplication (Equation 30) says the answer

is $80 \pm 110\%$ of 80 (where the 110% is the sum of 60% and 50%). Thus the basic rule says that the range is -8 to 168. The next term in the Taylor's series expansion (Equation 27g) adds 30% to the relative error (i.e., $\epsilon_x \epsilon_y / ab = 0.3$). This correction makes the range to be 80 minus 80% of 80 to 80 plus 140% of 80 – or 16 to 192. (The range is still huge of course, but the correction made by adding the next term pulls the lower part of the range out of negative territory. This can be important in some fields of geology – for example, it can turn water from running uphill.)

The bottom line here is that subtraction can produce large relative errors, and these may affect the applicability of the basic rules for multiplication and division.

Application: *Ye Olde Darcy's Law Problem.*

In CG-1 (Significant Figures!, May 1998), we looked at a plug-and-chug calculation based on an end-of-chapter problem from the hydrogeology textbook by Fetter (1994). The equation of choice is the one-dimensional Darcy's Law for velocity (v):

$$v = -\frac{K (h_2 - h_1)}{n_e \Delta s} \quad (31)$$

The numbers to plug in are 21 ft/day for K (hydraulic conductivity), 0.17 for n_e (effective porosity), 277.32 ft for h_2 (water level in distant well), 277.86 for h_1 (water level in the far-from-origin well), and 792 for Δs (distance between the wells). Carrying out the arithmetic on a calculator produces an answer of $v = 0.0842246$ on the display. The point of the CG-1 essay was that people should not record the answer with anywhere near that many digits. In fact, if you follow the rules of significant figures and write down 0.084, you are *overstating* the precision of the calculated result.

The column made the point by carrying the implied uncertainties through a direct-arithmetic calculation of the propagated error. For example, 21 for K implies that K is known as 21.0 ± 0.5 (Taylor, 1997). Thus we got an upper value for v by

$$v = -\frac{21.5 (277.315 - 277.865)}{0.165 \cdot 791.5} = 0.0905 \text{ ft/day} \quad (32a)$$

and a lower value for v by

$$v = -\frac{20.5 (277.325 - 277.855)}{0.175 \cdot 792.5} = 0.0783 \text{ ft/day}. \quad (32b)$$

These results indicate a range of 0.084 ± 0.006 , or $0.084 \pm 7\%$. In contrast, following the rules of significant figures and stating the answer as 0.084 communicates knowledge of the answer as 0.084 ± 0.0005 or $0.084 \pm 0.6\%$. Thus the precision is *overstated* by nearly a factor of 10 in this example, even when the rules of significant figures are followed.

We can now use the rules of error propagation to calculate the error in v . Writing out all the implied uncertainties, we get

$$v = -\frac{(21 \pm 0.5)}{(0.17 \pm 0.005)} \frac{[(277.32 \pm 0.005) - (277.86 \pm 0.005)]}{(792 \pm 0.5)} \text{ ft/day}. \quad (33a)$$

Using the rule of addition for absolute uncertainties during subtraction, the quantity in the brackets (Δh) is -0.54 ± 0.01 ft. Then, converting all the absolute uncertainties to relative uncertainties, we get

$$v = -\frac{(21 \pm 2.38\%) (-0.54 \pm 1.85\%)}{(0.17 \pm 2.94) (792 \pm 0.06\%)} \text{ ft/day.} \quad (33b)$$

Now, multiply the factors and add the relative errors:

$$v = -\frac{(21)(-0.54)}{(0.17)(792)} \pm (2.38\% + 1.85\% + 2.94\% + 0.06\%) \text{ ft/day.} \quad (33c)$$

This produces $v = 0.0842 \pm 7.23\%$ of 0.0842, or 0.0842 ± 0.006 ft/day as in Equation 32.

Further consideration. I would be remiss if I did not point out that this Darcy's Law problem is set up in the first place with unreasonable implied precision. To say that K for an actual site is known to two significant figures is not being realistic. I doubt that there are many cases where K is known to better than a half of an order of magnitude. To say that K is 21 ft/day more likely means that K is somewhere in the range of 10-30 ft/day (i.e., something like sand). Similarly, effective porosity is difficult to evaluate and apply over a reasonable space. A value of 0.17 may mean that n_e is in the range of 0.1 to 0.25 (i.e., large).

As for Δh , consider this. The original measurement might have gone something like: "I taped down to the water level a distance of 25.72 ft from a datum at the well of 303.58 ft, for a difference of 277.86 ft." Suppose, for the sake of argument, that quotation means "I taped down to the water level a distance of 25.72 ± 0.02 ft from a datum at the well of 303.58 ± 0.08 ft." The elevation of the water level then would be 277.86 ± 0.10 ft. Similarly, the other water level might be 277.32 ± 0.10 ft. Then Δh would be -0.54 ± 0.20 ft.

For $\Delta s = 790$, let's say that it is really 780 to 800 ft, or 790 ± 10 ft.

Then, instead of the values in Equation 33b, we have

$$v = -\frac{(20 \pm 50\%) (-0.54 \pm 37.0\%)}{(0.17 \pm 41.2\%) (792 \pm 01.3\%)} \quad (34)$$

This produces a result of $0.0802 \pm 129.5\%$ of 0.0802 from Equations 29 and 30. Thus, the range works out to be from -0.024 ft/day to 0.184 ft/day. Notice this range includes the possibility that the groundwater is flowing backwards.

Clearly, the uncertainties in the numbers going into the calculation are too large to use the first-order Taylor's series expansion. Box 1 goes through the calculation using the second-order terms. Step 1 expands the ratio of K/n_e . Step 2 expands the ratio $\Delta h/\Delta s$. Step 3 multiplies the two ratios together. The Taylor's series result compares favorably with the easy-arithmetic method (0.018 to 0.028), although adding the third-order terms would bring it closer.

In either case – via the second-order Taylor's series expansion or easy arithmetic – the result is worth thinking about. With these uncertainties going into the calculation, the propagated uncertainty in v ranges over an *entire order of magnitude* – from 0.02 ft/day to more than 0.2

ft/day. The central result (0.08) is not at the center of the range. It is much closer to the geometric mean (0.07) than to the arithmetic mean (0.15).

Box 1. Calculation of ν using second-order terms of Taylor's series

Step 1

$$\left(\frac{20 \pm 50\%}{0.17 \pm 41.2\%} \right) = \frac{20}{0.17} + \frac{20}{0.17} \left[\pm(0.50 + 0.412) + (0.50)(0.412) + (0.412)^2 \right]$$

$$= 117.647^{+151.5(\text{or } 128.8\%)}_{-63.1(\text{or } 53.6\%)}$$

Step 2

$$\left(\frac{0.54 \pm 37.0\%}{792 \pm 1.3\%} \right) = \frac{0.54}{792} + \frac{0.54}{792} \left[\pm(0.370 + 0.013) + (0.370)(0.013) + (0.013)^2 \right]$$

$$= 0.0006818^{+0.0002645(\text{or } 38.79\%)}_{-0.0002577(\text{or } 37.80\%)}$$

Step 3

$$\left(117.6^{+128.8\%}_{-53.6\%} \right) \left(0.0006818^{+38.79\%}_{-37.80\%} \right) = 0.0802^{+0.0802[+1.288+0.3879+(1.288)(0.3879)]}_{-0.0802[-0.536-0.378+(0.536)(0.378)]}$$

$$= 0.0802^{+0.17448(\text{or } 217.6\%)}_{-0.0571(\text{or } 71.1\%)}$$

or 0.023 to 0.255

Compare with

$$\text{minimum: } \nu = \left(\frac{10}{0.24} \right) \left(\frac{0.34}{800} \right) = 0.018 \quad \text{maximum: } \nu = \left(\frac{30}{0.1} \right) \left(\frac{0.74}{780} \right) = 0.28$$

Application: A Dip-and-Strike Problem

Applying Equation 29 to trigonometric functions involves a hazard: the differential of the angle must be in radians, *not* degrees. This is because a derivative such as $df/d\theta$, where θ is an angle, assumes that the angular measure is radians.

To illustrate an application involving trigonometric functions consider the strike and dip problem laid out in Figure 5. This example will also illustrate how Equation 29 expands to handle functions of three variables.

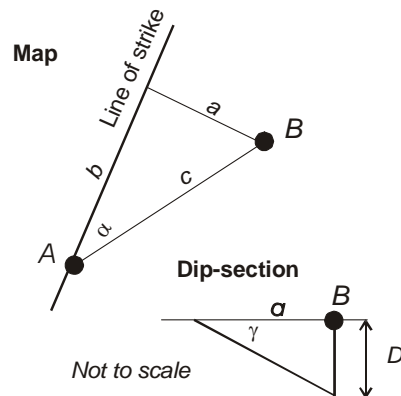


Figure 5. Map and cross-section showing the angles and distance for the strike-and-dip problem discussed in text.

Here is the problem. Suppose the top contact of a formation is exposed at A (Fig. 5), and you measure the strike as $N(32^\circ \pm 1^\circ)E$ and the dip as $22^\circ \pm 3^\circ$. Suppose you pace off a distance of 125 ± 3 m in a direction of $N(47^\circ \pm 1^\circ)E$ arriving at location B . Assuming the ground is level, how deep below you is the contact?

From the map view, the distance a is

$$a = c \sin \alpha \quad . \quad (35a)$$

From the dip section, the depth at B is

$$D = a \tan \gamma \quad (35b)$$

Combining the two equations produces

$$D = c \sin \alpha \tan \gamma , \quad (35c)$$

from which we can calculate an answer with $c = 125 \pm 3$ m, $\alpha = 15^\circ \pm 2^\circ$, and $\gamma = 22^\circ \pm 3^\circ$. The answer is $D = 13.1$ m.

To find the propagated error directly from Equation 35c, we use

$$\varepsilon_D = \left| \frac{\partial D}{\partial c} \right| \varepsilon_c + \left| \frac{\partial D}{\partial \alpha} \right| \varepsilon_\alpha + \left| \frac{\partial D}{\partial \gamma} \right| \varepsilon_\gamma . \quad (36)$$

Finding the partial derivatives and plugging in values of c , α and γ produces

$$\begin{aligned} \frac{\partial D}{\partial c} &= \sin \alpha \tan \gamma = 0.10457 \\ \frac{\partial D}{\partial \alpha} &= c \cos \alpha \tan \gamma = 48.78 \\ \frac{\partial D}{\partial \gamma} &= \frac{c \sin \alpha}{\cos^2 \gamma} = 37.63 . \end{aligned}$$

Using these values in Equation 36, with $\varepsilon_c = 2$ m, $\varepsilon_\alpha = 0.03491$ rad (or 2°), and $\varepsilon_\gamma = 0.05236$ rad (or 3°) produces the result $\varepsilon_D = 3.88$ m. Thus the solution to the problem using a first-order Taylor's series for the propagated error is 13.1 ± 3.9 m. This corresponds to a range of 9.2 to 17.0 m.

From simple arithmetic:

$$\begin{aligned} D_{max} &= (128) \sin(17^\circ) \tan(25^\circ) = 17.45 \text{ m} \\ D_{min} &= (122) \sin(13^\circ) \tan(19^\circ) = 9.45 \text{ m.} \end{aligned}$$

This problem can be varied in a number of ways. For example, D/c in Equation 35c is the tangent of the apparent dip from A to B , or the plunge of a linear element on the contact running from the surface at A to a depth D at B . In either case this angle (β) is given by

$$\tan \beta = \sin \alpha \tan \gamma \quad (37)$$

(e.g., Ragan, 1985, Equations 1.4 and 4.1). Then, given $\alpha = 15^\circ \pm 2^\circ$ and $\gamma = 22^\circ \pm 3^\circ$, what is $\beta \pm \varepsilon_\beta$? This problem is a little more computationally challenging: you need to know the derivative of the arctangent function, and you get to use the chain rule.

Final Remarks

Equation 30, which I am calling the basic equation of error propagation, gives the propagated uncertainty for a function of two variables, $f(x,y)$, if each of them has a *small* relative uncertainty. The first term in the parentheses is the part of the propagated uncertainty ε_f produced by the uncertainty in the first independent variable acting alone. Similarly, the second term is the part of the propagated uncertainty produced by the uncertainty in the second independent variable acting alone. As the equation shows, the propagated uncertainty is the sum of these parts. But, as we showed in the examples, there is more to this if the uncertainties in the independent variables are *not* small relative to their magnitude. In such cases, higher-order terms of the background Taylor's series are required. In cases where the mixed partial derivatives are not zero, there is a sort of synergy: the total propagated uncertainty is more than the sum of the parts (i.e., the partial uncertainties due to the uncertainties of each independent variable acting alone).

Equation 30 gives the *range* of the propagated uncertainty for cases where the uncertainties in x and y are small. The equation gives a *worst case scenario*: the maximum propagated error due to the first small uncertainty combined with the maximum propagated error due to second small uncertainty (the absolute values in the equation assuring the worst case). Equation 30 does not contemplate the likelihood that the effect of the first uncertainty partially offsets the effect of the second uncertainty. To take account of the probability of such partial compensation, the appropriate equation is (Taylor, 1997)

$$\varepsilon_f = \sqrt{\left(\frac{\partial f}{\partial x} \varepsilon_x\right)^2 + \left(\frac{\partial f}{\partial y} \varepsilon_y\right)^2}, \quad (38)$$

the quadratic sum of the effects of the two uncertainties considered individually. The propagated uncertainty due to Equation 38 is always smaller than the propagated uncertainty given by Equation 30.

Equation 38 contemplates that ε_x and ε_y reflect *measurement errors* (which are normally distributed) and that the measurements of x and y are *independent* (e.g., not measured with the same instrument). Thus, if ε_x and ε_y are both standard deviations from repeated measurements of x and y , respectively, then ε_f would represent the standard deviation for the variation in f .

Equation 38 is the choice of physicists making repeated laboratory measurements (Taylor, 1997). I believe, however, the equation is a bad idea for geologists. The reason is simple: Equation 38 assumes that the uncertainties in the measured variables are *small*, and in geology they usually are not. As illustrated in our examples, even the range given by Equation 30 is not

the maximum range resulting from geologically plausible uncertainties – because of the need for the second-order terms. Geologists, I believe, could do a better job in general of acknowledging the uncertainties in calculated values resulting from uncertainties in measurement. We don't need Equation 38 to make our uncertainties seem smaller.

If, however, the measurement uncertainties *seem* small (subjectively), and the measurements are *truly* independent, you may wish to have a test to see whether they are *small enough*. In that case, I would suggest using Equation 30 to calculate the propagated range, and calculate it also by direct arithmetic as we did in our examples. If the two answers agree (say to two significant digits), then you may be justified to think about Equation 38.

None of these considerations about uncertainties propagated through a calculation matter, of course, if the equations do not model the geological problem appropriately. As geologists we must not forget T.C. Chamberlain (1899, p. 224) (as quoted in Dalrymple, 1991, p. 44):

The fascinating impressiveness of rigorous mathematical analyses, with its atmosphere of precision and elegance, should not blind us to the defects of the premises that condition the whole process. There is perhaps no beguilement more insidious and dangerous than an elaborate and elegant mathematical process built upon unfortified premises.

References

- Chamberlain, T.C., 1899, On Lord Kelvin's address on the age of the Earth as an abode fitted for life: Annual report of the Smithsonian Institution, 1899, p 223-246.
- Dalrymple, G.B., 1991, The age of the Earth: Stanford CA, Stanford University Press, 474 p.
- Fetter, C.W., 1994, Applied hydrogeology (3rd edition): Upper Saddle River NJ, Prentice Hall, 691 p.
- Ragan, D.M., 1985, Structural geology, An introduction to geometrical techniques (3rd edition): New York, Wiley, 393 p.
- Taylor, J.R., 1997, An introduction to error analysis (2nd edition): Sausalito CA, University Science Books, 327 p.