Journal of Geoscience Education, v. 49, n.3, p. 305-313, May 2001 (edits, June 2005)

Computational Geology 16

The Taylor Series and Error Propagation

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

Topics this issue –

- *Mathematics: algebra -- polynomial and transcendental functions; calculus -- differential, derivative, finite difference.*
- *Geology: velocity equation; Stokes' law; dip angle; free-air gravity correction.*

Introduction

One of the most important concepts in first-year calculus is the distinction between differentials and derivatives. A *differential* is an infinitesimally small *amount*. A *derivative* is a *rate* of change. Given a function $f(x)$, an infinitesimally small change in the dependent variable (f) is brought about by an infinitesimally small change in the independent variable (x). The derivative $f'(x)$ can be viewed as a ratio of those infinitesimal changes.

This is the calculus of Leibniz (1646-1716), John Bernoulli (1667-1748), and Euler (1707-1783). Thus, the derivative of f with respect to x can be written

$$f'(x) = \frac{df}{dx}, \tag{1a}$$

where the right hand side is the ratio of the two infinitesimals, df and dx. In the calculus of the Leibnizian tradition, this ratio is called the differential quotient (e.g., Thompson, 1987), and the two differentials can be separated. Thus, multiplying Equation (1a) by dx,

$$df = f'(x)dx\ . \tag{1b}$$

Equation 1b articulates the conceptual difference between a differential and a derivative. The df is the infinitesimal change in h brought about by the infinitesimal change dx in x. The derivative $f'(x)$, or df/dx, is the rate of change of f with respect to x.

The trouble with this intuitive notion is that differentials and infinitesimal increments are "evanescent" quantities – as they were called in a withering 1734 essay by George Berkeley (1685-1753), an Irish bishop and philosopher who exposed logical inconsistencies underpinning intuitive calculus (Bell, 1945, Chap.13; Edwards 1979, p.292-299). With d'Alembert (1717-1783), Lagrange (1736-1813) and finally Cauchy (1789-1857), a logically consistent calculus was developed, resulting (1821) in Cauchy's

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}, \qquad (2)$$

which everyone learns as the "official" definition of derivative

Equations 1, therefore, hold only in the limit. They generally do not hold for measurable changes in x or f. In other words, if the limit is removed from Equation 2, the equality applies only "approximately." Taking away the limit and multiplying by Δx produces

$$\Delta f \approx f'(x)\Delta x, \qquad (3a)$$

where

$$\Delta f = f(x+\Delta x) - f(x). \qquad (3b)$$

Equation (3a) is sometimes called the finite-difference form of Equation (2). The delta-quantities (finite differences) are tangible quantities, whereas the *d*-quantities (differentials) are "evanescent."

The subject of this essay is the word "approximately" in the previous paragraph. How close are Equations 1b and 3a to each other? The Taylor series addresses the question directly. Brooke Taylor (1685-1731), a protégé of Newton (1642-1727) and contemporary of Berkeley, is sometimes referred to as the inventor of finite differences (Bell, 1945, p. 285).

With the Taylor series we can address a question that should be of great interest to anyone who makes calculations from measured quantities: given that there is some uncertainty (error) in the measured quantity (x), how much uncertainty (error) is there in the calculated value $f(x)$? In this column, we will focus on functions of one variable. In the next column, we will consider functions of more than one variable: given uncertainties in both x and y, how much uncertainty is propagated into $g(x,y)$?

The Problem, Topographically Posed

Imagine that you are on a hillside. To limit ourselves to a one-dimensional problem, suppose that the hillside slopes upward to the east (x-direction) and that there is no slope in the north-south direction (i.e., suppose the topographic contours run north-south). We are interested in your elevation, h. In this situation, h depends only on x – thus, $h(x)$. Suppose you are located a distance of 1200 m from a benchmark taken as the origin of your coordinate system; thus, x = 1200 m. Suppose your elevation is 1125 m, and the upward slope is 0.2 (i.e., 20% grade, or 11°). The question is, What is your elevation if you walk eastward to x = 1300 m?

In more mathematical language, the question can be stated as follows: If a = 1200 and $h(a)$ = 1125, what is $h(x)$ if x = 1300 and $h'(x) = 0.2$? Or, more generally, what is $h(x)$, if you know $h(a)$ and $h'(a)$? Or, in English, if you know the elevation and slope at one point (a), how can you calculate your elevation at a nearby point (x)?

Possible answers. Because we are given $h'(a)$, it is appropriate to recast Equation 2 as:

$$h'(a) = \lim_{(x-a)\to 0} \frac{h(x)-h(a)}{x-a} . \qquad (4)$$

Then, removing the limit and rearranging, the expression equivalent to Equation 3 is:

$$h(x) - h(a) \approx (x-a)h'(a) , \qquad (5a)$$

Geologists, familiar with the expression "slope is rise over run," will recognize that Equation 5a has the rise [($h(x)$-$h(a)$] on the left and the run (x–a) times the slope [h'(a)] on the other.

From Equation 5a, we can easily get an equation explicitly for $h(x)$:

$$h(x) \approx h(a) + (x-a)h'(a) \qquad (5b)$$

Substituting in values, we get an answer:

$$h(1300) \approx 1125 + 100 * 0.2 = 1145 \text{ (m)} \qquad (6)$$

Equation 5b amounts to *projecting* the hillside upward at a constant slope from a = 1200 m to x = 1300 m. Clearly, the answer (1145 m) in Equation 6 is correct if the slope (0.2) does not change. In that case, the profile of the hillside is a straight line with the equation:

$$h(x) = h(a) + (x-a)h'(a) \qquad (7)$$

With a = 1200, $h(a)$= 1125, and $h'(a)$ = 0.2, Equation 7 is

$$h = 885 + 0.2x . \qquad (8)$$

Equation 8 can be used to find the elevation at x = 1300 or at any other location.

If, however, the slope does change as one moves away from x = 1200, then other answers for h(1300) are clearly possible. For example, here are three other functions for which h(1200) = 1125 and h'(1200) = 0.2:

$$h = 309 + 1.16x - 0.0004x^2 , \qquad (9a)$$
$$h = -555 + 2.6x - 0.001x^2 , \qquad (9b)$$
and
$$h = -3435 + 7.4x - 0.003x^2 . \qquad (9c)$$

It is a worthwhile exercise to verify that these functions do satisfy the stipulation that $h'(1200) = 0.2$.

In the functions of Equations 9, the slope decreases as one goes uphill (increasing x) as shown in Figure 1A. The elevations h(1300) calculated from these equations, therefore, are smaller than the value (1145 m) projected from Equation 8 (Table 1). The departure from the projected value increases with the magnitude (absolute value) of the coefficient of x^2 (Table 1). This coefficient is related to the *rate* that the slope changes as x increases.

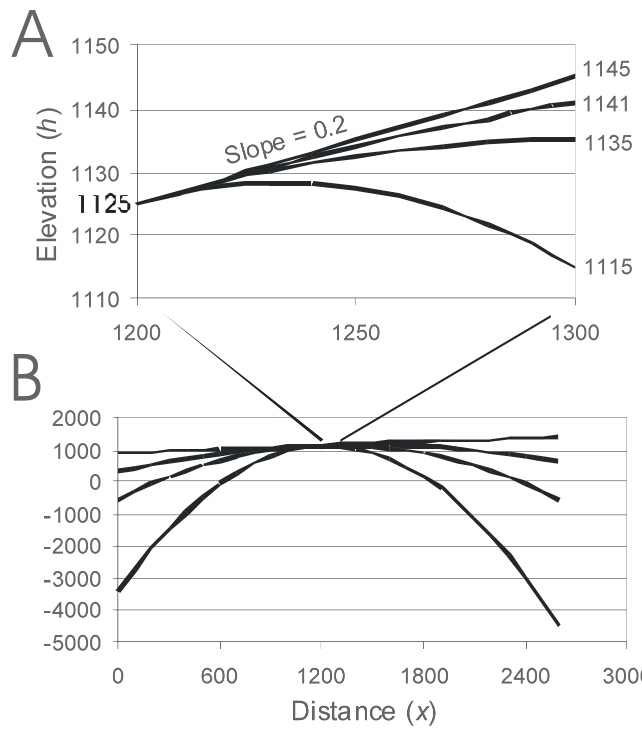

Figure 1. Plots of Equations 8 (straight line) and 9 (three parabolas) at two scales. All four lines pass through the same point (1200, 1115) with the same slope (0.2).

	Straight Line	parabolas		
	Eqn 8	Eqn 9a	Eqn 9b	Eqn 9c
c_0	885	309	−555	−3435
c_1	0.2	1.16	2.6	7.4
c_2		−0.0004	−0.001	−0.003
$h(1300)$	1145	1141	1135	1115
$R(1300)$		−4	−10	−30
$R(1210)$		− 0.04	− 0.1	−0.3
$R(1201)$		−0.0004	− 0.001	−0.003

Table 1. The four lines in Figure 1.

An infinitude of parabolas. The curves of Equations 9 and Figure 1 are all convex-upward parabolas. For the parabola of Equation 9a, the maximum (the top of the hill in this analogy) is at x>1300. For the second parabola (Equation 9b), the maximum is at x =1 300. For the third (Equation 9c), the maximum is between x = 1200 and x =1 300, and the parabola is on its way down at x = 1300.

An infinite number of parabolas can be drawn through the point x = 1200 and $h(1200) = 1125$ with a slope $h'(1200)$ of 0.2. They all have the form of

$$h(x) = c_0 + c_1 x + c_2 x^2 \,, \tag{10}$$

which is a second-order polynomial function ("second order" refers to the highest power of the independent variable, x).

The coefficients (c_i) of Equation 10 determine the shape and location of the curve as can be seen by looking at the three functions of Equations 9 at a smaller scale (Fig. 1B). The first coefficient, c_0, positions the parabola vertically, because it gives the y-intercept (where the parabola crosses the y-axis). The second coefficient, c_1, gives the slope of the parabola at the y-intercept. The ratio of two of the coefficients (c_1 and c_2) tells the x-position of the maximum, because, from setting the first derivative equal to zero and solving for x,

$$x_{summit} = -\frac{c_1}{2c_2} \,. \tag{11}$$

Meanwhile, c_2 is proportional to the second derivative, $h''(a)$ and determines how curved the parabola is.

With the parabolas all passing through the same spot on the hillside [$h(1200) = 1125$] with the same slope [$h'(1200) = 0.2$], there is only one additional degree of freedom in stipulating the values of the three coefficients. For example, if we specify the value of c_2, then the values of c_0 and c_1 are completely determined – and thus so is the location and shape of the parabola. With the progressive increase of c_2 (Table 1), the y-intercept gets deeper (c_0), the slope at $x = 0$ gets steeper (c_1), and the summit comes closer to $x = 1200$ (Equation 11).

The relationship between the coefficients of parabolas passing through a point (at $x = a$) with a stipulated slope can be found easily by differentiating Equation 10 and back-substituting. Thus, from Equation 10,

$$h'(x) = c_1 + 2c_2 x \,, \tag{12a}$$

from which,

$$h''(x) = 2c_2 \,. \tag{12b}$$

But we know that, at $x = a$, the value of the function is $h(a)$ and the value of its derivative is $h'(a)$. Then, Equations (10) and (12a) become

$$c_0 + c_1 a + c_2 a^2 = h(a) \tag{12c}$$

and

$$c_1 + 2c_2 a = h'(a) \,, \tag{12d}$$

respectively. Solving Equation 12d for c_1 produces

$$c_1 = h'(a) - 2ac_2 . \qquad (12e)$$

Combining Equations 12c and 12e produces

$$c_0 = h(a) - ah'(a) + a^2 c_2 . \qquad (12f)$$

Meanwhile, from Equation 12b, c_2 is simply

$$c_2 = h''(a)/2 . \qquad (12g)$$

From Equations 12e and 12f one can write down the equation for any parabola given a point that it passes through, the value of the slope at that point, and an assumed value for c_2. Thus with $h(1200) = 1125$, $h'(1200) = 0.2$, and $c_2 = -0.001$, Equation 12e becomes

$$c_1 = 0.2 - (2)(1200)(-0.001) = 2.6 , \qquad (13a)$$

and Equation 12f becomes

$$c_0 = 885 + (1200^2)(-0.001) = -555 . \qquad (13b)$$

Substituting these values and $c_2 = 0.001$ into the general form of the parabola (Equation 10) produces the equation of the sought-for parabola – which is Equation 9b.

Although it is interesting to know how to fit a parabola through a point given its slope at the point, the chief purpose of this discussion has been to show that the coefficients are interrelated. If one wishes to change the coefficient (c_2) representing the departure from a straight line, then the other coefficients must adjust to keep the parabola passing through the stipulated point with the stipulated slope.

The *R*-term. Let's return to the original question, How approximate are Equations 3a and 5b? We can answer this question for parabolas moving through the identified spot on the hillside. Equations 3a and 5b predict the uphill elevation based on an assumption that the hillside has a straight-line profile. The approximation comes in because the profile is parabolic (in these examples), not linear. As shown in Figure 1A and Table 1, the inaccuracy of the prediction at $x = 1300$ increases with the magnitude of c_2. This inaccuracy is typically referred to in problems such as this as the *R*-term. The "*R*" stands for remainder. We can solve for it explicitly.

We can obtain an expression for R by substituting the equations for the coefficients (Equations 12d and 12e) into the general equation of the parabola (Equation 10). This produces

$$h(x) = [h(a) - ah'(a) + a^2 c_2] + [h'(a) - 2ac_2]x + c_2 x^2 , \qquad (14a)$$

which, with a little algebra, rearranges to

$$h(x) = h(a) + (x-a)h'(a) + (x-a)^2 c_2 . \qquad (14b)$$

Substituting now for c_2 from Equation 12g, our expression for $h(x)$ becomes

$$h(x) = h(a) + (x-a)h'(a) + \frac{(x-a)^2}{2} h''(a) . \qquad (15)$$

Here is the crucial step: compare Equation 15 to Equation 5b. Because R is the difference between the left and right sides of Equation 5b,

$$R = h(x) - [h(a) - (x-a)h'(a)] , \qquad (16)$$

R then is

$$R = \frac{(x-a)^2}{2} h''(a) \qquad (17)$$

from Equations 15 and 16.

Equation 17 says, then, that the "approximately" in Equations 3a and 5b is determined in part by the magnitude of the second derivative. Larger second derivatives -- larger curvatures of these parabolas -- mean larger departures between the left and right sides of those equations. This point is obvious from Figure 1A.

Also evident in Figure 1A is the fact that the "approximately" in Equations 3a and 5b is determined partly too by the distance $x-a$. The approximation gets better with decreasing $x-a$, no matter how large the departure from the straight line. This is seen in Equation 17 by the fact that $R \to 0$ as $x \to a$. At the scale of Figure 1A, the three parabolas are indistinguishable from the straight line close to $x = 1200$. At $x = 1210$, for example, the elevation from the straight line is 1127.00 m, whereas from the parabolas it is 1126.96, 1126.90 and 1126.70 m. The R-term decreases by a factor of 100 for every 10-fold decrease in $x-a$ (Equation 17 and Table 1).

Taylor's Series

The Taylor series is the representation of a function by a polynomial function, where the coefficients of the polynomial are related to the values of the function and its derivatives at a particular point, a. In essence, the Taylor series says that knowledge of the values of h and all of its derivatives at one spot on the hillside can be used to predict, as accurately as you want, the value of h at any other spot on the hillside.

Polynomial functions. The Taylor series representation of a straight line is

$$h(x) = h(a) + (x-a)h'(a) ,$$

which is Equation 7. The Taylor series representation of a parabola is

$$h(x) = h(a) + (x-a)h'(a) + \frac{(x-a)^2}{2} h''(a) ,$$

which is Equation 15. The Taylor series representation of a cubic function (a third-order polynomial) is

$$h(x) = h(a) + (x-a)h'(a) + \frac{(x-a)^2}{2!}h''(a) + \frac{(x-a)^3}{3!}h'''(a) \qquad (18)$$

where 2! and 3! refer to 2-factorial (2∗1) and 3-factorial (3∗2∗1), respectively, and h''' is the third derivative. By comparing these three equations it is not difficult to guess that the Taylor series representation of an n^{th}-order polynomial function would be:

$$h(x) = h(a) + (x-a)h'(a) + \frac{(x-a)^2}{2!}h''(a) + \ldots + \frac{(x-a)^n}{n!}h^{(n)}(a), \quad (19)$$

where $h^{(n)}$ refers to the n^{th} derivative.

To show that Equations 18 and 19 are correct, it is useful to look at another way of getting Equations 7 and 15. This other approach shows conceptually what is involved in projecting $h(a)$ to $h(x)$.

Start back at the beginning (Equation 1b):

$$dh = h'(x)dx.$$

Suppose we know $h(a)$. Then to find h at some other value of x, we can simply integrate between a and x,

$$\int_{h(a)}^{h(x)} dh = \int_a^x h'(x)dx, \qquad (20a)$$

from which,

$$h(x) = h(a) + \int_a^x h'(x)dx. \qquad (20b)$$

If $h(x)$ is a straight line, then $h'(x)$ is *constant*. Suppose we know its value at $x = a$. We can then substitute $h'(a)$ for $h'(x)$ in Equation 20b and easily carry out the integration, because $h'(a)$ is a constant. The result is Equation 7.

If $h(x)$ is *not* a straight line, then $h'(x)$ is not a constant and we need to find an expression for it. If we know the second derivative $h''(x)$, we can derive an expression for $h'(x)$ from

$$dh' = h''(x)dx, \qquad (21a)$$

which is analogous to Equation 1b. Integrating Equation 21a between a and x produces:

$$h'(x) = h'(a) + \int_a^x h''(x)dx . \tag{21b}$$

Substituting Equation 21b into 20b and performing the integration, we get:

$$h(x) = h(a) + (x-a)h'(a) + \int_a^x\int_a^x h''(x)dxdx . \tag{21c}$$

The approximation in Equation 5a – the *R*-term of Equation 16 – is clearly the repeated integral in Equation 21c. We now can refer to it as the R_2-term, meaning that it is the *second* derivative integrated *twice* and represents the remainder between the function and its representation by a *first*-order polynomial.

If $h(x)$ is a *parabola*, then $h''(x)$ is a *constant*. Suppose we know its value at $x = a$. Then we can substitute $h''(a)$ for $h''(x)$ in Equation 21c and perform the integration. Thus

$$\int_a^x\int_a^x h''(a)dxdx = \int_a^x (x-a)h''(a)dx = \frac{(x-a)^2}{2}h''(a) . \tag{21d}$$

Substituting this result into Equation 21c produces Equation 15, the Taylor's series for the second-order polynomial (parabola).

If $h(x)$ is *not* a first- or second-order polynomial, the second derivative is not constant and so we need to find an expression for it. As with Equations 20b and 21b, we can get an expression from the next higher derivative:

$$h''(x) = h''(a) + \int_a^x h'''(x)dx \tag{22a}$$

Substituting Equation 22a into Equation 21c produces

$$h(x) = h(a) + (x-a)h'(a) + \int_a^x\int_a^x h''(a)dx^2 + \int_a^x\int_a^x\int_a^x h'''(x)dx^3 , \tag{22b}$$

and now we see that the approximation in Equation 5a is the *sum* of the two repeated integrals. We have already evaluated the first one (Equation 21d), which, substituted in Equation 21b produces

$$h(x) = h(a) + (x-a)h'(a) + \frac{(x-a)^2}{2}h''(a) + \int_a^x\int_a^x\int_a^x h'''(x)dx^3 . \tag{22c}$$

Equation 22c shows that if a *parabola* is used to project $h(a)$ to $h(x)$, then the difference between the function $h(x)$ and the parabola (second-order polynomial) is the *third* integral form *a* to *x* of the *third* derivative. Call this difference R_3.

If $h(x)$ is a *third*-order polynomial, then $h'''(x)$ is a constant, and equal to $h'''(a)$. Integrating $h'''(a)$ three times between a and x gives

$$\int_a^x\int_a^x\int_a^x h'''(a)dx^3 = \frac{(x-a)^3}{3\cdot 2}h'''(a)\,. \qquad (22d)$$

Substituting Equation 22d into Equation 22c produces Equation 18.

If $h(x)$ is a *fourth*-order polynomial, the right-hand side of Equation 22d becomes the penultimate term in the series. The last term is the *fourth* integral of the (constant) fourth derivative:

$$\int_a^x\int_a^x\int_a^x\int_a^x h^{iv}(a)dx^4 = \frac{(x-a)^4}{4\cdot 3\cdot 2}h'''(a)\,, \qquad (23)$$

and so we can write down the Taylor series of the fourth-order polynomial. Obviously, this kind of thing can be carried on as long as one would like for an nth-order polynomial function (Equation 19).

These high-order terms, however, are not significant if x stays very close to a. As $x-a$ goes to zero, the higher-order terms drop out. The third-order polynomial, for example, becomes a good approximation to the fourth-order polynomial; then the parabola approximates the third-order polynomial; then the straight line approximates the parabola (as in Fig. 1A). When that happens, Equation 7 is good enough.

The ubiquity of polynomial functions. Equation 19 applies to a polynomial function of order n. Its importance is not limited to polynomial functions.

Derivations of Equation 19 typically start with a statement like "Let $f(x)$ be a function with a continuous nth derivative throughout the (closed) interval $[a,b]$" ((Sokolnikoff and Redheffer, 1966, p. 36). Then the nth derivative is integrated n times between a and x, where x is any point on $[a,b]$. This produces a polynomial like that in Equation 19 up through the $(n\text{-}1)^{\text{th}}$ derivative. The last term, then, is left as

$$R_n = \int_a^x \ldots \int_a^x f^{(n)}(x)(dx)^n\;, \qquad (24a)$$

which represents the difference between the function and an (n-1)-order polynomial. This R_n-term is rewritten in a form due to Lagrange (who followed Taylor by several decades),

$$R_n = \frac{(x-a)^n}{n!}f^{(n)}(\xi) \qquad (24b)$$

for some ξ in [a,x].

If $f(x)$ has continuous derivatives of all orders and R_n (Equation 24b) goes to zero as $n \to \infty$ for each x on $[a,b]$, then $f(x)$ can be represented by the power series

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)(x-a)^n}{n!}\,. \tag{24c}$$

This is Equation 19 for $n \to \infty$, and the implication is that the function can be treated as a polynomial.

Nearly all functions that we are apt to run into have continuous derivatives of all orders and can be expanded into Taylor series. Examples include transcendental functions such as the exponential function, the sine and cosine functions, the tangent function (except at places like $\pm\pi/2$), and the logarithmic function (for $x>0$), and binomial expansions for negative and fractional exponents.

Geological examples that are *not* included are topographic profiles that have cliffs and overhangs, and stratal surfaces that are faulted or deformed into kinks or overturned folds. By definition, faults are discontinuous. Kink folds have a discontinuous first derivative. Cliffs, overhangs and overturned folds are not even functions because a function, by definition, is single-valued [has only one value $f(x)$ for each x].

An important exercise in first-year calculus is to represent transcendental functions and binomial expansions by Taylor series about $a = 0$. This produces well-known Maclauren series, such as

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + ... \tag{25a}$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - + ... \tag{25b}$$

$$(1-x)^{-1} = 1 - x + x^2 - x^3 + ... \qquad |x| < 1. \tag{25c}$$

Such series make the point that, given information about the function and its derivatives at a point (i.e., zero, in the case of Maclauren series), one can calculate the value of the function at any other x, if one is willing to carry enough terms.

In the following application of Taylor series, the value of a used in the expansion is not zero, but rather some measured value. In general, we will stick close to a (i.e., $x-a$ will be small), so we will not need to carry nearly so many terms.

Error Propagation

Suppose we measure a variable x and that its value is a. Suppose the measurement has an uncertainty ("error"), $\pm\varepsilon_x$, and we use the measured value in the function, $f(x)$. The uncertainty ε_x results in a propagated uncertainty ("error"), ε_f, in the calculated value, $f(a)$. In other words,

$$f(x) = f(a) \pm \varepsilon_f\,. \tag{26}$$

Because ε_x is small and ε_f is $f(x)-f(a)$, ε_f can be equated to the first-derivative term of the Taylor series (e.g., Taylor, 1997):

$$\varepsilon_f = \pm f'(a)\varepsilon_x \,. \tag{27}$$

Certainly, it is reasonable to assume that ε_x is small. After all, it is an uncertainty term. Who wants to think about a measured quantity such as 35 ± 25 m?

We will now consider some examples from the exercises and problems discussed in Computational Geology 6, "Solving Problems" (May 1999).

Example 1. The first exercise of CG-6 asked for the velocity (v) of the Pacific plate given the age of the shield volcanoes at Oahu (2.6 Ma) and western Hawaii (0.43 Ma) and a separation between them of "about 250 km". The answer is obtained from

$$v = s/t\,, \tag{28}$$

where s is the displacement (2.5×10^7 cm), and t is the difference of the two ages (2.17×10^6 years). The velocity works out to 11.52 cm/yr (with excess digits).

Now suppose "about 250 km" means 250 ± 20 km. What ± term should we attach to the calculated velocity? Applying Equation 27,

$$\varepsilon_v = \pm v'(s)\varepsilon_s \,. \tag{29a}$$

Differentiating Equation 28,

$$v'(s) = 1/t\,, \tag{29b}$$

[t is a constant in this one-dimensional problem (i.e., v depends only on s)]. Combining Equations 29a and 29b produces

$$\varepsilon_v = \pm\varepsilon_s / t\,. \tag{29c}$$

With $\varepsilon_s = \pm 2.0\times10^7$ cm and $t = 2.17\times10^6$ yr, then $\varepsilon_v = \pm 0.92$ cm/year. Thus,

$$v = 11.52 \pm 0.92 \text{ cm/yr} \tag{30a}$$

This is exactly the result we get by substituting the upper and lower limits of the range of s directly into Equation 28:

$$v_{upper} = (250+20)\text{ km} / 2.17\text{ m.y.} = 12.44\text{ cm/yr} \tag{30b}$$

and

$$v_{lower} = (250-20)\text{ km} / 2.17\text{ m.y.} = 10.60\text{ cm/yr}\,. \tag{30c}$$

Example 2. Another exercise in CG-6 involved the settling velocity (v) of a foraminifer using Stokes Law for a falling sphere:

$$v = \frac{\Delta\rho g}{18\mu} D^2\,, \tag{31}$$

where $\Delta\rho$ is the difference in density between the falling particle and the fluid ($\rho_s - \rho_f$); g = 981 cm/sec^2; μ is the fluid viscosity; and D is the grain diameter.

Now let's consider a spherical silt particle with diameter 0.005256 cm (4.25ϕ) and density 2.67 g/cm^3 falling through a fluid with density 1.025 g/cm^3 and viscosity 0.015 g/cm-sec. From Equation 31, v = 0.165088 cm/sec (with excess digits).

Suppose there is an uncertainty in D. Specifically, let $\varepsilon_D = \pm 0.0002$ cm, so that the sphere diameter is D = 0.005256±0.000200 cm. What is ε_v?

From Equation 27,

$$\varepsilon_v = \pm v'(D)\varepsilon_D \,. \tag{32a}$$

Rewriting Equation 31 as

$$v = CD^2\text{, where } C = \frac{\Delta\rho g}{18\mu}, \tag{32b}$$

a constant. Equation 32a becomes

$$\varepsilon_v = 2CD\varepsilon_D \,. \tag{32c}$$

Plugging in values for C (5976.833 cm^{-1}sec^{-1}), D and ε_D, then ε_v is ±0.012565 cm/sec. Thus the sphere's settling velocity is

$$v = 0.165088 \pm 0.012565 \text{ cm/sec} \tag{33a}$$

to six digits.

Using Equation 31 with D+ε_D and D−ε_D directly, v = 0.177892 cm/sec for D+ε_D and v = 0.152763 cm/sec for D−ε_D. These results are

$$v = 0.165088^{+0.012804}_{-0.012325} \text{ cm/sec,} \tag{33b}$$

which is not the same as Equation 33a. Thus, the direct method gives an ε_v term that isn't even symmetric. The ε_v term of Equation 33a is an intermediate value that differs in the fourth decimal place from the uncertainties in Equation 32b.

The discrepancy between the two ways of calculating ε_v is larger with larger ε_D. For example, let $\varepsilon_D = \pm 0.001$ cm. Then

$$v = 0.165088^{+0.068801}_{-0.056847} \tag{34a}$$

by the direct method. Meanwhile, Equation 32 produces

$$v = 0.165088 \pm 0.062824 \text{ cm/sec.} \tag{34b}$$

The difference is in the third decimal place.

The reason for the discrepancy is that Equation 27 truncates the Taylor series at the first-derivative term. Adding the second-derivative term from Equation 19, and carrying the sign of ε_x through the arithmetic within each term, Equation 27 should be replaced by

$$\varepsilon_f = \pm f'(a)\varepsilon_x + \frac{1}{2} f''(a)\varepsilon_x^2 . \tag{35}$$

The second term is "+" because ε_x is squared. As a result, ε_f is not symmetric. The "+" error exceeds the "−" error.

From Equation 32b, Equation 35 for Stokes' Law becomes

$$\varepsilon_v = \pm 2CD\varepsilon_D + C\varepsilon_D^2 \tag{36}$$

For $\varepsilon_D = \pm 0.0002$ cm, the first term is ±0.012565 (Equation 33a), and the second term is +0.000239; the sum of the two terms reproduces the results in Equation 33b exactly. For $\varepsilon_D = \pm 0.001$ cm, the first term is ±0.062824 (Equation 34b), and the second term is +0.005977. Their sum reproduces the results in Equation 34a exactly.

Example 3. In the first example, the function [$v(s)$] is a first-order polynomial, and the direct method produces a calculated uncertainty that is duplicated exactly by the first-derivative term of a Taylor's series expansion of the function (Equation 29a). In the second example, the function [$v(D)$] is a second-order polynomial, and a two-term expression (Equation 35) corresponding to the first- and second-derivative terms of a Taylor's series expansion exactly produces the propagated uncertainty found by the direct method. Other functions require more terms of the Taylor's series. Functions such as those in Equations 25, for example, lead to an infinite number of terms to calculate the propagated uncertainty exactly.

For illustration, consider again Example 1 with $s = 25\times10^6$ cm and $t = 2.17\times10^6$ yr. Now, rather than having the uncertainty in s, we will consider an uncertainty in t. Thus the function is $v = s/t$ again, but now t is the variable, and we want ε_v due to ε_t. The expressions of Equations 27 and 35 need to be extended to

$$\varepsilon_f = \pm f'(a)\varepsilon_x + \frac{1}{2} f''(a)\varepsilon_x^2 \pm \frac{1}{2\cdot 3} f'''(a)\varepsilon_x^3 + \frac{1}{4!} f^{iv}(a)\varepsilon_x^4 \pm .... \tag{37}$$

from Equation 19. Differentiating $v = s/t$, and substituting derivatives and ε_t into Equation 37 produces

$$\varepsilon_v = \mp \frac{s}{t^2}\varepsilon_t + \frac{s}{t^3}\varepsilon_t^2 \mp \frac{s}{t^4}\varepsilon_t^3 + \frac{s}{t^5}\varepsilon_t^4 \mp .... \tag{38}$$

Suppose $\varepsilon_t = \pm 0.01\times10^6$ yr. Then, calculating v for both $t = 2.16\times10^6$ yr and $t = 2.18\times10^6$ yr,

$$v = 11.52074^{+0.05333675}_{-0.05284742}\ \text{cm/yr} \quad (39a)$$

Meanwhile, Equation 38 produces

$$\varepsilon_v = \mp 0.05309554 + 0.00024466 \mp 0.00000113, \quad (39b)$$

which reproduces the uncertainty in Equation 39a to the eighth decimal place.

In contrast, $\varepsilon_t = \pm 0.1\times 10^6$ yr produces

$$v = 11.52074^{+0.5565574}_{-0.5075215}\ \text{cm/yr}, \quad (40a)$$

exactly, and the first three terms of Equation 38 are

$$\varepsilon_v = \mp 0.530909554 + 0.0244658781 \mp 0.0112745982. \quad (40b)$$

The sum of these three terms reproduces the uncertainty in Equation 40a to the fourth decimal place.

Finally $\varepsilon_t = 1\times 10^6$ yr produces the strongly asymmetric result

$$v = 11.5207^{+9.8468}_{-3.6343}\ \text{cm/yr} \quad (41a)$$

to four places. The first eight terms of the Taylor expansion are

$$\varepsilon_v = \mp 5.3091 + 2.4466 \mp 1.1275 + 0.5196 \mp 0.2394 + 0.1103 \mp 0.0508 + 0.0234, \quad (41b)$$

which gives the correct uncertainty term to only two places.

Discussion of results. These examples underscore the point that uncertainties in the independent variable (x) are supposed to be small if one intends to use Equation 27 to find the propagated uncertainty (ε_v) in $f(x)$. Equation 27 is an approximation for exactly the same reason that Equation 3b is an approximation. As the uncertainty ε_x increases, it becomes less and less like a "true differential," and so Equation 27 becomes ever more approximate. The larger the uncertainty ε_x is, the more terms that must be carried in the Taylor series to accurately produce ε_f to the desired number of places.

One way to reduce the need for additional terms is to recognize that ε_f is produced only approximately by Equation 27 and, therefore, we should be satisfied with only one or, at most, two significant digits in ε_f. Thus for $\varepsilon_D = \pm 0.0002$ cm in Example 2, stating the result as $v = 0.165 \pm 0.012$ cm/sec represents the range in Equations 33 (using two digits for ε_v, because the last digit of v is "5"). Similarly for $\varepsilon_D = \pm\ 0.001$ cm, stating $v = 0.165 \pm 0.063$ cm/sec is appropriate for Equations 34. For Example 3 and $\varepsilon_t = \pm 0.01\times 10^6$ yr, use $v = 11.52 \pm\ 0.05$ cm/yr (for Equation 39), and, for $\varepsilon_t = \pm 0.1\times 10^6$ yr, use 11.5 ± 0.5 cm/yr (for Equation 40). Finally the $\varepsilon_t = \pm 1\times 10^6$ yr in Example 3 is 46% of the value of t itself, which is beyond the pale for a differential. With such large uncertainties, one needs to calculate the range from the end points.

If one can always find the propagated uncertainty by calculating it directly from the endpoints of the range, why even know Equation 27? There are at least two answers. The first is that the result from the one-term Taylor series can be easily folded into more complicated calculations involving uncertainties in more than one variable. This point will be discussed in the next Computational Geology.

The second answer is that the one-term Taylor series result is easily stated formally in the equation to be used in the calculation. For example, to calculate the plate velocity from displacement (*s*) and travel time (*t*), don't use Equation 28, but rather

$$v = \frac{s}{t} \pm \frac{\varepsilon_s}{t} = \frac{s}{t}\left(1 \pm \frac{\varepsilon_s}{s}\right) \tag{42}$$

from Equation 29c if there is an uncertainty in *s* (Example 1). On the other hand, if there is an uncertainty in *t* (Example 3), use

$$v = \frac{s}{t} \pm \left|\frac{\varepsilon}{t^2}\right| = \frac{s}{t}\left(1 \pm \left|\frac{\varepsilon_t}{t}\right|\right) \tag{43}$$

from Equation 38. For a Stokes' Law problem and an uncertainty in *D* (Example 2), use

$$v = \frac{\Delta\rho g D^2}{18\mu}\left(1 \pm \frac{2\varepsilon_D}{D}\right) \tag{44}$$

from Equation 32c.

By stating the equations in this way, one immediately sees the effect of the uncertainty in the independent variable. Thus a 1% error in measuring the distance produces a 1% error in the velocity (Equation 42). The *same* is true for a 1% error in measuring the travel time (Equation 43): it produces a 1% error in the velocity. On the other hand, a 1% variation in the grain diameter produces a 2% variation in the settling velocity calculated by Stokes' Law (Equation 44).

A final problem: Consequences of an error in dip angle. To illustrate the convenience of incorporating the one-term formula for propagated error into the equation to calculate the target unknown, we will consider one last problem. Suppose you are on level ground at an outcrop of sandstone overlying limestone. Suppose the contact strikes north-south and dips east at an angle β. How deep below ground is the contact at a distance a = 100 m to the east (i.e., in the direction of dip) given that there is a possible 1° error in the measurement of β? On what does the answer depend?

Call the depth b. Then, remembering that $\tan(\beta) = b/a$, and that the derivative of the tangent is the secant squared, then, from Equation 27, the general relationship is

$$b = a\left(\tan\alpha \pm \frac{\varepsilon_\beta}{\cos^2\beta}\right) \tag{45}$$

where ε_β is in radians. Equation 45 clearly shows that the consequence of the 1° error increases substantially for large dips. Thus for 10°, 30°, 45°, 60° and 80° dips, the answers are 17.6 ± 1.8 m, 58.0 ± 2.3 m, 100 ± 3.5 m, 173 ± 7 m, and 567 ± 58 m, respectively.

But, then, the depth (β) to the contact also increases substantially with depth. The *relative uncertainty* for the depth (ε_β/β) is

$$\frac{\varepsilon_\beta}{\beta} = \pm\frac{1}{\tan\beta\cos^2\beta} \tag{46}$$

from Equation 45. Thus the relative error goes through a minimum at $\beta = 45°$. For dip angles of 10°, 30°, 45°, 60° and 80°, the relative errors for the depth are 10%, 4.0%, 3.5%, 4.0%, and 10%, respectively, for the 1° error in dip. This result -- unlike the result for the absolute uncertainty (ε_β) in Equation 46 – is independent of the horizontal distance *a*.

Final Remarks

The Taylor series is one of the truly useful tools in first-year calculus. Its usefulness is not surprising given the early history of calculus. As Stillwell (1989) states at the beginning of his chapter on calculus, calculus emerged in the seventeenth century as a system of shortcuts to results obtained by more tedious methods for finding areas and volumes. Calculus "is about calculation, after all" (p. 101). [Stillwell's book aims "to give a unified view of undergraduate mathematics by approaching the subject through its history" (p. vii).]

Newton understood calculus "as an algebra of infinite series" (Stillwell, 1989, p. 107). Newton, and another founder of calculus, James Gregory (1638-1675), started their work with interpolation (Stillwell, 1989, p. 123). They developed the method now known as Gregory-Newton interpolation – a standard topic in courses on numerical methods. In that context, they independently discovered the binomial theorem (Newton in 1665, Gregory in 1670), which produces a power series. Taylor derived his series expansion as a limiting case of the Gregory-Newton interpolation formula in 1715.

Euler was "probably the greatest virtuoso of series manipulation" (Stillwell, 1989, p. 124). His textbooks on *calculi differentialis* (1755) and *calculi integralis* (1768-1774) included Taylors' theorem "with many applications" (Struik, 1987).

Speaking of applications, anyone who has had a course in geophysics has heard the words "And dropping higher-order terms…" many times. Typically the setting is a derivation in which a derived, but inconvenient, expression is simplified to a more workable one. For example (Fowler, 1990, p. 170), the free-air variation in the acceleration due to gravity (g) with elevation (z) is, from Newton's Law of Gravitation,

$$g(z) = g_0\left(\frac{R}{R+z}\right)^2,$$

where g_0 is gravity at sea level, and R is the radius of the Earth. The next line is

$$g(z) = g_0\left(1 - \frac{2z}{R}\right),$$

because $z \ll R$. Thus the change in g due to the increase in radius from R to $R+z$ is $-2g_0z/R$ (the free-air correction). Such simplifications apply Taylor's theorem.

The one-term formula for propagated error (Equation 27) is a similar simplification. So is the statement that the differential of the dependent variable (df) can be calculated as the product of the derivative [$f'(x)$] and the differential of the independent variable (dx) (Equation 1b).

References

Bell, E.T., 1945, The development of mathematics (2nd edition): New York, McGraw Hill, 636 p (New York, Dover edition, 1992).

Edwards, C.H., Jr., 1979, The historical development of the calculus: New York, Springer-Verlag, 351 p.

Fowler, C.M.R., 1990, The solid Earth: Cambridge University Press, 472 p.

Sokolnikoff, I.S., and Redheffer, R.M., 1966, Mathematics of physics and modern engineering (2nd edition): New York, McGraw-Hill, 752 p.

Stillwell, 1989, Mathematics and its history: New York, Springer-Verlag, 371 p.

Struik, D.J., 1987, A concise history of mathematics (4th revised edition): New York, Dover Publications, 228 p.

Taylor. J.R., 1997, An introduction to error analysis (2nd edition): Sausalito CA, University Science Books, 327 p.

Thompson, S.P., 1946, Calculus made easy, being a very-simplest intorduction to those beautiful methods of reckoning which are generally called by the terrifying names of the differential calculus and the integral calculus (3rd edition): New York, St. Martins Press, 250 p.