UCRL-16172

University of California

Ernest O. Lawrence Radiation Laboratory

ALLOCATION OF ERRORS TO THEORY AND EXPERIMENT

Berkeley, California

DISCLAIMER

Submitted to the Journal of
Theoretical Biology

UCRL-16172

UNIVERSITY OF CALIFORNIA

Lawrence Radiation Laboratory
Berkeley, California

AEC Contract No. W-7405-eng-48

ALLOCATION OF ERRORS TO THEORY AND EXPERIMENT

Grove C. Nooney

June 3, 1965

ALLOCATION OF ERRORS TO THEORY AND EXPERIMENT*†

Grove C. Nooney

Lawrence Radiation Laboratory
University of California
Berkeley, California

June 3, 1965

SUMMARY

Errors in theory and experiment are each taken into account explicitly in the methodology here presented. The theory is assumed to be represented by a mathematical model and the experiment by a measured function. A suitably defined composite deviation permits a proper allocation to model and to measurement of any discrepancy between them. It is shown how quantitative limits may be obtained for the trustworthiness of the theory relative to the experiment. An appendix contains appropriate mathematical formalism, based on the calculus of variations, as well as examples of application of the methods described.

I. Introduction

The principal method for judging the validity of a theory is comparison with experiment. This testing of the theory is often accomplished by taking into account experimental error and estimating how likely could be the experimental results, if the theory were perfectly valid. Since the theory is but a conjecture it may contain error also, but the described procedure cannot consider error in theory directly. In the following discussion, we show how to take explicit account of error in theory, and we describe a method for allocation to theory and to experiment of any discrepancy between them. This leads to a quantitative measure of the trustworthiness of the theory relative to the experiment.

We confine our attention to the quantitative theory represented by a mathematical model and to the quantitative experiment represented by measurements. The mathematical model is assumed to define a function that is measured experimentally. By considering a suitably defined deviation of the measurement from the function given by the model, we determine the appropriate allocation of that deviation to model error and to measurement error. Independent information about either error may then be used in judging both the validity of the model and the reliance to be placed in the model relative to the experiment.

Associated mathematical formalism is relegated to the Appendix, where several examples are given. This formalism most readily allows the analysis of linear models, in which the important errors often occur in the common use of the linear model as an acknowledged approximation.

II. Strong and Weak Solutions

Consider a mathematical model of the form

$$M(y', y, x) = 0, \tag{1}$$

relating the independent variable x to the dependent variable y and its derivative y'. We assume M to be a continuously differentiable function of its arguments.

The relation (1) may contain a first-order differential equation, and there may be an associated initial condition; but more general views of the relation (1) may be accommodated. For example, the variable y may be considered as a vector, with the relation (1) then containing a system of first-order ordinary differential equations. We permit the absence of initial conditions, considering them as parameters of the model, and we permit the model to contain other parameters as well. The variable x may also be considered a vector, with the relation (1) then interpreted to contain a system of partial differential equations. Even the restriction to first-order differential equations is not necessary, although that is the case we illustrate.

A function y^* satisfying equation (1) is called a strong solution of the model. If the model is given by equation (1), we assume that the experiment results in a measurement $\tilde{y}$ of a strong solution over some interval. If $\tilde{y} = y^*$, then the model is taken to be perfectly valid, and the measurement to be exact.

Since the model that is only nearly valid may provide a useful description of reality and since, in fact, we cannot expect absolute validity (Nooney, 1965),a satisfactory model may need satisfy only the inequality

$$\left| M(z', z, x) \right| < \epsilon \tag{2}$$

for some function z and small positive ϵ. Any function satisfying the inequality (2) for specified ϵ is called a weak solution of the model. A more precise designation is ϵ-weak solution, and we term the model ϵ-valid.

III. Deviations

A common measure of deviation D between a strong solution y^* and its measurement $\check{y}$ is

$$D(y^*) = \int [y^*(x) - \check{y}(x)]^2 \, dx, \tag{3}$$

where the range of integration is the domain of $\check{y}$. We assume that the model defines y^* over that domain also. If the model and consequently its strong solution contain parameters, then the definition is modified to make $D(y^*)$ the minimum of the right-hand side of equation (3) with respect to the parameters. The resulting $D(y^*)$ is the usual least-squares deviation. The y^* specified by the parameter values yielding the minimum is called the best-fitting strong solution. A common method of judging the validity of the model given in relation (1) has been to estimate the probability distribution of measurement error, assumed random, and to calculate the probability of a deviation's exceeding $D(y^*)$ by applying the errors to y^*. If that probability is reasonably large, then the model is said to have a large degree of validity; if that probability is small, then the model is said to have a small degree of validity.

Now, suppose we do not request absolute validity of the model, but only ϵ-validity. Then we may define a composite deviation that gives a measure of how well satisfied is the model and how well approximated is the measurement. For any function z we set

$$\bar{M}(z) = \int [M(z', z, x)]^2 \, dx,$$

where the range of integration is again the domain of $\check{y}$, and we define the composite deviation

$$C(z) = w \bar{M}(z) + D(z), \tag{4}$$

where w is a weight reflecting the relative importance of model and measurement. The parameter w may also be considered to specify the trust placed in the theory relative to the experiment.

IV. The Function z* and the Weight w

It is natural to seek the minimum composite deviation, and we set

$$C(z^*) = \min_{z} C(z), \tag{5}$$

where the functions competing to achieve the minimum are the functions continuously differentiable over the domain of $\tilde{y}$. We call the function z* the best-fitting weak solution. In general, z* is not a strong solution. Equation (A2) of the Appendix shows that $D(z^*) \leq D(y^*)$, and therefore z* better approximates the measurement $\tilde{y}$ than does y*. If the model contains parameters, then the minimization of the composite deviation must be carried out simultaneously with respect to the function z and all parameters of the model. The resulting best-fitting weak solution is again a better approximation to $\tilde{y}$ in the sense of D than is the best-fitting strong solution. Note that the two sets of parameter values obtained by minimizing the two deviations need not be identical. The expression $\overline{M}(z^*)$ is a measure of the error in the model; any discrepancy between z* and $\tilde{y}$, given for instance by $D(z^*)$, may be attributed to measurement error. Both errors are influenced by the choice of w. The measurement error, we realize, is actually due entirely to experimental inaccuracies only if the value of w employed is the correct value and only if the model truly has error $M(z^{*\prime}, z^*, x)$. The deviation $D(z^*)$ may be investigated by the same statistical methods used with $D(y^*)$.

Regarding w as variable in equation (4), we see that as w becomes small, the best-fitting weak solution z* tends to the measurement $\tilde{y}$, and as w becomes large, z^* tends to a strong solution y^*. We may expect, then, that a given maximum allowable value for $D(z^*)$ will define an upper bound for w. We may expect also that a requirement for the model to be ϵ-valid, with a given maximum allowable value of ϵ, will define a lower bound for w. Should this upper bound exceed the lower bound, there would be a weak solution satisfying both conditions given in advance. That is, there would be a function that sufficiently well satisfies the model while sufficiently well approximating the measurement. Should the lower bound for w exceed the upper bound, there would be no function simultaneously satisfying the advance conditions. Each bound separately limits the trustworthiness of the model relative to the measurement.

V. Calculations

The Appendix demonstrates the application of the calculus of variations in the determination of the best-fitting weak solution z^*. Although the expressions $\overline{M}(z^*)$ and $D(z^*)$ are of major interest, it is expedient for us to make use of available formalism in first obtaining z^*, then calculating those expressions. With the linear model, an explicit formula may be obtained for z^* as a function of the variable x and the weight-parameter w. Such a formula lends itself to the relatively easy determination of the aforementioned bounds for w.

Throughout this discussion the measurement $\tilde{y}$ was assumed given only at each point of some interval. This assumption is violated in many experiments, which may give $\tilde{y}$ only on a sparsely distributed, discrete set of points. For this reason, the Appendix concludes with a definition of the composite deviation for discrete measurements and with an indication of formalism appropriate to that definition.

APPENDIX

Formalism

If the measurement $\tilde{y}$ is given over the interval (a, b), then the mathematical problem set in the main text of this paper is the minimization of

$$C(z) = \int_a^b \left\{ w[\, M(z', z, x)]^2 + [\, z(x) - \tilde{y}(x)]^2 \right\} dx \tag{A1}$$

with respect to the class of functions z continuously differentiable on (a, b). Depending on the formulation of the model M, the functions z may be subjected also to an initial condition, $z(a) = z_0$. We assume there is a unique function z^* of the class mentioned for which

$$C(z^*) = \min_z C(z).$$

We have already defined the deviation D and the (continuously differentiable) strong solution y^* by the relations

$$D(y) = \int_a^b [\, y(x) - \tilde{y}(x)]^2 \, dx$$

and

$$M(y^{*\prime}, y^*, x) = 0.$$

From the definition (A1) and these last three relations follow the inequalities

$$D(z^*) \leq C(z^*) \leq C(y^*) = D(y^*)\,. \tag{A2}$$

Therefore, in the sense of the deviation D, z^* more closely approximates $\tilde{y}$ than does y^*.

The minimization of the integral in equation (A1) is a standard problem of the calculus of variations (e.g., Bliss, 1946), and leads to the Euler-Lagrange equation as well as supplementary boundary conditions for z^*. In general, the function z^* will not be a strong solution of the model. If the model M contains but a single first-order ordinary differential equation, the corresponding Euler-Lagrange equation will be a second-order ordinary differential equation. Systems of such equations and systems of partial differential equations arise as the Euler-Lagrange equations corresponding to our other interpretations of M.

For illustration, let us assume for the model the form

$$M(z', z, x) = z' - f(z, x),$$

with the function f having continuous partial derivatives. This is a common form for dynamical models, as in the study of tracer kinetics, for instance. For this model, the function z^* yielding the minimum composite deviation is given by the Euler-Lagrange equation satisfied by $z = z^*$,

$$w z'' + w(f_x - ff_z) - (z - \tilde{y}) = 0 , \tag{A3}$$

where the subscripts denote partial derivatives. To that differential equation must be adjoined the natural boundary conditions

$$z' - f(z, x) = 0, \quad x = a, b,$$

obtained also through the formalism of the calculus of variations. It is remarkable that, although the solution z^* of equation (A3) need not be a strong solution of the model, the natural boundary conditions require exact satisfaction of the model by z^* at the points a and b. If an initial condition has been specified, it replaces the natural boundary condition for $x = a$. If the function f is linear in z, then we may write the explicit solution of (A3) in terms of integrals (Ince, 1944). Of course, that explicit solution contains w and thus offers the possibility of bounding w as discussed in Sec. IV.

Example 1. Suppose the model is given by an expression containing no derivatives,

$$M(z', z, x) = z - g(x) ,$$

for some function g. Then we obtain for the strong solution $y^*(x) = g(x)$. The composite deviation takes the form

$$C(z) = \int [w(z-g)^2 + (z - \tilde{y})^2] \, dx ,$$

and the associated Euler-Lagrange equation for the minimizing function is

$$w(z^* - g) + z^* - \tilde{y} = 0 .$$

This yields for the best-fitting weak solution

$$z^* = \frac{1}{1+w} (w f + \tilde{y}) .$$

The error in the model is

$$z^* - g = \frac{1}{1 + w} (\tilde{y} - g) ,$$

and the difference between the best-fitting weak solution and the measurement is

$$z^* - \tilde{y} = \frac{w}{1 + w} (g - \tilde{y}) .$$

Since $y^* = g$, we may write

$$z^* - \tilde{y} = \frac{w}{1 + w} (y^* - \tilde{y}) ,$$

and conclude that z^* is closer to y at every point than is y^*. We find also that

$$D(z^*) = \left(\frac{w}{1 + w} \right)^2 D(y^*) .$$

It is clear that specifying upper bounds for model error and deviation D furnishes lower and upper bounds, respectively, for w.

Example 2. Suppose the model is given by

$$M(z', z, x) = z' - z ,$$

with the initial condition $z(0) = 1$. Suppose further the experimental measurement is given by

$$y(x) = e^x + x, \quad 0 \leq x \leq 1 .$$

Then we find $y^*(x) = e^x$ and $D(y^*) = 1/3$. The composite deviation becomes

$$C(z) = \int_0^1 [w(z' - z)^2 + (z - \tilde{y})^2] \, dx ,$$

with the associated Euler-Lagrange equation

$$w z'' - (1 + w) z + \tilde{y} = 0 .$$

The initial condition and the natural boundary condition $z'(1) - z(1) = 0$ must be satisfied for $z = z^*$. We find

$$z^*(x) = e^x + \frac{1}{1 + w} x .$$

The error in the model is given by

$$M(z^{*\prime}, z^*, x) = \frac{1}{1 + w} (1 - x)$$

or

$$\overline{M}(z^*) = \frac{1}{3} \left(\frac{1}{1 + w} \right)^2 .$$

Again, we see that z^* is closer to $\tilde{y}$ at every point than is y^*. Further,

$$D(z^*) = \left(\frac{w}{1 + w} \right)^2 D(y^*) .$$

As before, a priori specifications of maximum allowable errors result in upper and lower bounds for w.

Example 3. Let us take the conditions of Example 2, deleting the initial condition. The value $y^*(0) = y_0$ then becomes a parameter of the model to be determined by considering $y^*(x) = y_0 e^x$ and

$$D(y^*) = \min_{y_0} \int_0^1 (y^* - \tilde{y})\, dx .$$

The minimum is given by the best-fitting strong solution

$$y^*(x) = (1 - e^{-1})^{-1} e^x .$$

The composite deviation and the Euler-Lagrange equation for the best-fitting weak solution z^* remain unchanged. The absence of initial condition requires the imposition on z^* of both natural boundary conditions $z'(x) - z(x) = 0$ for $x = 0, 1$. The resulting solution of the Euler-Lagrange equation is

$$z^*(x) = e^x + \frac{1}{1 + w}\; x + A\, e^{rx} + B\, e^{-rx} ,$$

where

$$r = (1 + 1/w)^{1/2}$$

$$A^{-1} = (1 + w)\,(1 - r)\,(1 - e^{2r}) ,$$

$$B^{-1} = (1 + w)\,(1 + r)\,(1 - e^{-2r}) .$$

Here algebraic difficulties are annoying, but again the imposition of a priori limits on the errors result in bounds for w.

Discrete Measurement

When the measured function $\tilde{y}$ is known at a relatively few points x_i (including the end points) of some interval (a, b), we may employ another definition of the composite deviation

$$C(z) = w\, \overline{M}(z) + \sum_i \left[z(x_i) - y(x_i) \right]^2 .$$

In this case, existing formalism of the calculus of variations is unfortunately of little use. For the calculation of the best-fitting weak solution we may turn to the theory of dynamic programming, where constructive methods are available (Bellman, 1957).

FOOTNOTES AND REFERENCES

*This work was done under the auspices of the U.S. Atomic Energy Commission.

†A preliminary report of this work was presented under the title "Remarks on the Determination of Parametric Models" at the Third Annual Symposium on Biomathematics and Computer Science in the Life Sciences, Houston, 1965.

1. Bellman, Richard E. (1957). "Dynamic Programming". Princeton: Princeton University Press.
2. Bliss, Gilbert A. (1946). "Lectures on the Calculus of Variations". Chicago: University of Chicago Press.
3. Ince, Edward Lindsay (1944). "Ordinary Differential Equations". New York: Dover.
4. Nooney, Grove C. (1965). Mathematical Models in Medicine: A Diagnosis, J. Chronic Diseases, Symposium Issue.