

**ERROR ESTIMATION IN EXPERIMENTAL DATA AND
OPTIMAL FUNCTIONAL REPRESENTATIONS****Jay FRANKEL
Majid KEYHANI
Kunihiko TAIRA**Mechanical and Aerospace Engineering and Engineering Science Department
University of Tennessee, Knoxville, Tennessee (USA) 37996-2210
Email: frankel@titan.engr.utk.edu*Keywords: Least-squares, error estimation, radial basis functions***ABSTRACT**

The development of in-phase, a posteriori error estimates and the “best functional representation for the solution and its derivative” are examined in this paper. This presentation delivers several new and fundamental findings pertinent to both the experimentalist and analyst. Applications requiring a high degree of precision such as associated with the defense, aerospace and heat treatment arenas will greatly benefit from these new findings.

In the present context, noisy data are defined as data containing only a random error component. In error analysis, a single value for the global uncertainty is presented or assumed. In stark contrast, this study involves a local examination of the error and attaches an error estimation to each corresponding data value. If desired, a global uncertainty value can then be produced. Representing data by a function is advantageous in many situations. The “best solution” is developed in lieu of the “best curve fit”. This subtlety involves the development of a new methodology that minimizes $\|f - \bar{f}_N\|_2^2 = \|\epsilon_i - R_N\|_2^2$ where f is the ideal function truly being sought in the presence of errorless (continuous) data, \bar{f}_N is the approximate function being developed using contaminated data, ϵ_i is the local error for the i^{th} data point defined as $f(x_i) - f_i$, and R_N is the traditional local residual defined as $\bar{f}_N(x_i) - f_i$ where x_i is the i^{th} collection point and f_i is the corresponding data value to form the data set denoted as $\{x_i, f_i\}_{i=1}^M$ where M represents the total number of data collected from the experiment. Here, $\|\cdot\|_2$ represents the discrete 2-norm.

Numerical differentiation of noisy data is well known to be ill-posed in the sense of Hadamard since small perturbations in the function can lead to large variations in the derivative. This can quickly be understood by noting that the absolute error associated with forward difference approximation to $f'_\epsilon(x)$ in the presence of noisy data results in the error bound $f'_\epsilon < O(h) + 2\delta/h$ where δ is the maximum absolute error, i.e., $|f - f_i| < \delta$ and where h is the conventional step size. Therefore, for $\delta > 0$, as $h \rightarrow 0$ the resulting error bound blows up. Thus, $h \rightarrow 0$ can not be performed in an arbitrary manner.

The basic framework involves the classical least squares method which is re-examined with the intent of (i) approximating local measurement errors necessary for estimating uncertainty; and, (ii) predicting optimal solutions in both the primitive function and its derivative. With accurate local-error estimates, the initial data can be corrected by adding these estimates to the initial data to substantially reduce the root-mean-square error, i.e., uncertainty. Additionally, the clarification between “best curve fit” and “best prediction or solution” is elucidated in this paper in order to provide insight into the meaning of a desirable resolution. The “best prediction” permits the proper interpretation of the derivative which is essential to many contemporary studies involving preprocessing of data. The framework offered here presents an elegant but practical method for approximating local measurement errors and for developing a functional representation of the desired function that best represents the solution in the presence of errorless data.

Copyright ©2003 JSME