BUILDING ADAPTIVE STOCHASTIC MODELS USING LOW-COST OUTPUT ERROR ESTIMATES

Isaac M. Asher[1] and Krzysztof J. Fidkowski[2]

[1] Graduate Student, University of Michigan, 1320 Beal Ave., Ann Arbor, MI 48109,
isaaca@umich.edu
[2] Assistant Professor, University of Michigan, 1320 Beal Ave., Ann Arbor, MI 48109,
kfid@umich.edu, www.umich.edu/~kfid

Key words: *Stochastic Collocation, Adaptation, Error Estimation, Uncertainty Quantification*

Uncertainty modeling (or quantification) is a valuable technique for analyzing the reliability of designs and physical models in computer simulations. The process of uncertainty modeling can be reduced to interpolating a high-dimensional function from samples. The interpolation (or stochastic) error can be large in regions far from sample points. The extremely large size of the high-dimensional space implies that an exponential number of samples is needed (this is the "curse of dimensionality"). In practice, the function of interest may have a low inherent dimensionality, that is it only varies in a small (perhaps two- or three-dimensional) subspace. In addition, since the choice of dimensions can be ad-hoc (e.g. when they are parameters of a simulation) it is possible that the "interesting" directions may not aligned with the originally specified dimensions, but instead are linear combinations of them.

It is hypothesized that by utilizing an extra low-cost piece of information, the curse can be avoided for such functions. The proposed information is an error estimate which can explore the high-dimensional space cheaply. At the same time, an uncertainty model based on ideas similar to support vector machines and ridge-regression is adaptively built within the "interesting" space. The model allows for variation along any direction:

$$u(\mathbf{x}) = \sum_{i=1}^{N} \overline{u}_i \phi_i(\mathbf{x}^T \mathbf{d}_i) \tag{1}$$

Here, the function of interest is $u(\mathbf{x})$. The $\overline{u}_i$ are coefficients of the one-dimensional basis functions ϕ_i. The $\mathbf{d}_i$ are directions along which the model varies. Adding more and higher-order ϕ_i results in arbitrarily high accuracy. The coefficients can be found via least-squares fitting to samples $u(\mathbf{x}_j)$. The goal is to choose the $\mathbf{d}_i$ to be only within the "interesting" subspace, thereby restricting the number of samples required.

Exploring the high-dimensional space is accomplished by examining errors associated with a *direction* $\mathbf{d}$. Using the low-cost error estimate, a model is built for the direction-error $\epsilon(\mathbf{d})$. Then, a particle swarm method does the exploration by maximizing $\epsilon(\mathbf{d})$. If the model has large errors along a particular direction, then that direction should be added as a new term in Eq 1. New samples along the direction are then taken to find the new $\overline{u}$. If the true function only varies along a small number of $\mathbf{d}$, that is, if the inherent dimensionality is small, then samples will be restricted to this low-dimensional space.

Figure 1 shows an example of the adaptation results for the simple case of interpolating a 10-dimensional function in which only two dimensions are actually active. The algorithm detects these dimensions and quickly builds an accurate model. The plot shows the error in the average value of the function as samples are added. The method proposed here compares favorably to the same method but with randomly chosen $\mathbf{d}$ and standard Gaussian and Monte Carlo schemes, which suffer from the curse of dimensionality.
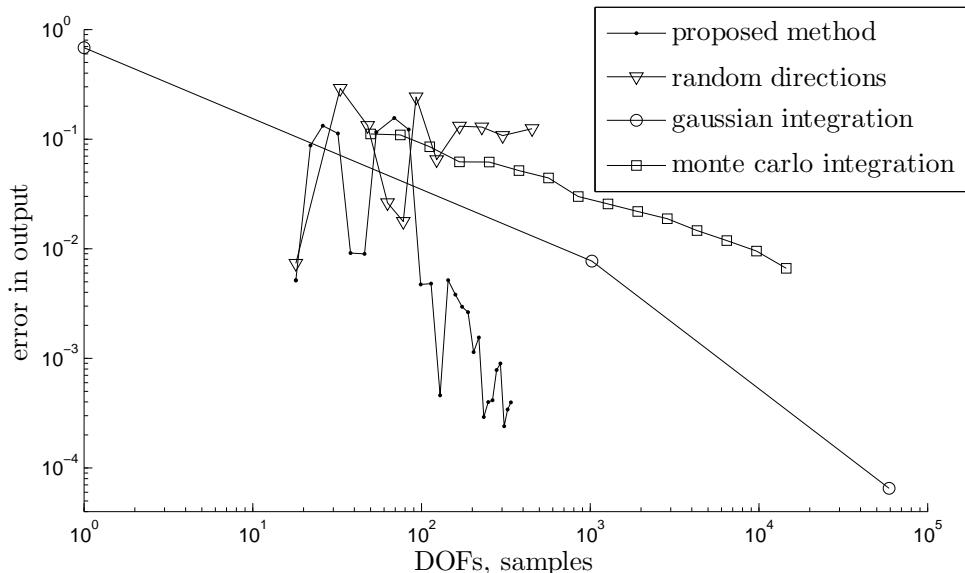

Figure 1: Comparison of four methods for finding the average value of a 10-dimensional function.

An important consideration is that samples may have inaccuracy themselves, often called deterministic error. In general, the stochastic error and deterministic error should be balanced. A method for reducing the stochastic error is presented, with the understanding that in the future it will be extended to balance the two sources of error.

The method can be applied to any simulation that provides error estimates for non-converged solutions, for example adjoint-based error estimates. The adjoint error can be used to target the model to a statistical output, for example the mean or variance of some quantity over the uncertain parameters, further enhancing efficiency. Either way, the method effectively reduces the dimensionality of the problem when possible, even if the directions of variation do not lie along originally specified dimensions.