Efficient Construction of Surrogates for Computer Experiments with Gradient Information

Multiscale Math and Optimization for Complex Systems

Fred J. Hickernell
Illinois Institute of Technology
Room E1-208, 10 W. 32nd St., Chicago, IL 60616

Abstract

Computations of a model that accurately represent a complex system may be expensive. Specifically, for any multivariate input to the model, $\boldsymbol{x}$, the evaluation of the output, $y = f(\boldsymbol{x}) + \varepsilon$, may take a lot of time. Here $f : \mathbb{R}^d \to \mathbb{R}$ denotes the functional relationship between the input and output, and ε denotes numerical or stochastic noise. A surrogate, $\tilde{f}$, is an accurate approximation to f that can be constructed quickly given data from the computer experiment, $(\boldsymbol{x}_1, y_1), \ldots, (\boldsymbol{x}_n, y_n)$.

In constructing a surrogate one must make several choices:

- the *design*, i.e., the positions of the data sites, $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$,
- the choice of output data, e.g., whether *gradient* values in addition to function values facilitate the surrogate construction,
- the family of surrogate *models* to consider, e.g., regression (parametric) or kriging (nonparametric),
- the selection of model *attributes*, e.g., regression basis, kernel shape parameters, etc.

This presentation provides some answers to these questions. Using a simplified model of heat transport in a nuclear reactor core, it is demonstrated that the addition of gradient information, whose marginal cost is small, allows one to construct accurate surrogates with much fewer experimental runs. Including gradient information in the regression model using the traditional polynomial basis, however, can lead to an ill-conditioned information matrix. This problem is shown to be circumvented by a new polynomial basis that is orthogonal with respect to an inner product whose definition includes the gradient. Low discrepancy designs are proven to provide acceptable integrated root mean square prediction errors for surrogates based on regression models. Kriging, while having attractive optimality properties, may suffer from a curse of dimensionality. On the other hand, this difficulty may be overcome if the shape parameters decay as the dimension increases. This means that the nominally high dimensional problem is essentially a low dimensional one.