Proceedings of the 2018 Winter Simulation Conference
M. Rabe, A. A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, eds.

GRADIENT BASED CRITERIA FOR SEQUENTIAL DESIGN

Collin Erickson

Northwestern University
2145 Sheridan Road
Evanston, IL 60208, USA

ABSTRACT

Computer simulation experiments are commonly used as an inexpensive alternative to real-world experiments to form a metamodel that approximates the input-output relationship of the real-world experiment. While a user may want to understand the entire response surface, they may also want to focus on interesting regions of the design space, such as where the gradient is large. In this paper we present an algorithm that adaptively runs a simulation experiment that focuses on finding areas of the response surface with a large gradient while also gathering an understanding of the entire surface. We consider the scenario where small batches of points can be run simultaneously, such as with multi-core processors.

1 INTRODUCTION

In the general setup for design and analysis of computer experiments (Sacks et al. 1989), there is a black box function or simulation that can be evaluated at any point, but each evaluation is expensive. After the experiment, a metamodel (also called an emulator or surrogate model) is created that approximates the true response function and is much faster to evaluate. Traditional experiment designs, or one-shot designs, select an entire list of inputs where the computer model will be evaluated before any data is collected. When the data is collected iteratively, the experiment can be run adaptively. This allows the experimenter to use the data collected up to the present time to make an informed decision for which point(s) should be evaluated next. Existing methods for adaptive experiments generally either are trying to find an optimal response, e.g., minimization, or are trying to fit the entire surface well, e.g., minimizing the mean squared error over the surface.

In our setting, we posit that practitioners often try to find a general understanding of the entire response surface with some focus on interesting areas. Instead of defining interesting to be an optimum, we consider situations where steep areas are the primary focus. That is, a user would like to understand the function in areas where the derivative of the computer model is large (in absolute value). Identifying where the derivative is large allows the user to identify regions where small changes in the inputs can result in relatively large changes in the output. Our goal is that at the end of our adaptive approach, we can build a metamodel that behaves well throughout the surface, but does a bit better where the derivative is large. We consider the experimental setting where design points can be evaluated in small batches. This is practical in many settings, e.g., most experimenters can run computer simulations on multiple cores simultaneously.

2 METHODOLOGY

Informally, our objective is to have higher prediction accuracy in areas where the gradient is large. Mathematically, our objective function that the experiment attempts to minimize is the weighted mean squared error of the metamodel, where the weight is the gradient norm squared. However, when running an experiment, the true function is unknown, so estimates of the gradient norm squared of the function and the squared error must be used from the metamodel. Both of these are available when using a Gaussian process model for the metamodel, as is common in design and analysis of computer experiments (Santner

et al. 2003). The squared norm of the metamodel's estimate of the gradient, which we call the plug-in estimator, is suboptimal since it does not account for the uncertainty of the metamodel. A better, unbiased estimator that we propose is to use the expected value of the gradient norm squared, which accounts for the uncertainty in the metamodel. This quantity can be calculated using the Gaussian process metamodel. Using these quantities from the metamodel, we create a criterion that, given a set of data already evaluated, determines which point(s) should be evaluated next to best minimize our objective.

In the batch setting we are considering, this optimization is difficult. To reduce the computational cost, we only consider selecting points from a set of candidate points. This set should be space filling, and needs to grow faster than the set of evaluated points to ensure that there are sufficient options for selection. Even when using candidate points, heuristics must be used to pick the batch of points to be evaluated, such as with a greedy algorithm. We use an sFFLHD (Duan et al. 2017) to generate candidate points since it can generate points as long as needed while maintaining good properties for the whole design.

3 RESULTS

In our numerical results, We compare our algorithm and proposed criterion to four other criteria. The first is the same as ours, except using the plug-in estimate of the gradient. The second is the traditional method of minimizing the mean squared error over the entire surface. The final two are nonadaptive methods that just use points from an sFFLHD or a Sobol sequence (Sobol 1967). We compared these methods on three known functions that each have two input dimensions. The metric we use for comparison is the actual gradient norm squared weighted mean squared error, which can be calculated since we use known equations. The results show that the our proposed method and the plug-in method give the best performance by a large margin. The nonadaptive methods that simply use points from an sFFLHD or Sobol sequence, perform much worse, as expected.

4 CONCLUSION

We propose a new method that allows adaptive computer experiments to have a higher prediction accuracy where the gradient is large. This is practical since many situations where practitioners are interested in finding areas of the response surface where a small change in inputs can have a large effect on the output. Our algorithm adaptively selects a batch of points from a set of space filling candidate points to be evaluated in each iteration. The information from data already collected in the experiment is incorporated into a Gaussian process metamodel that allows for efficient selection of points. In future work we will perform more extensive numerical comparisons and expand the method so it can be applied in a more general setting.

REFERENCES

Duan, W., B. E. Ankenman, S. M. Sanchez, and P. J. Sanchez. 2017. "Sliced Full Factorial-Based Latin Hypercube Designs as a Framework for a Batch Sequential Design Algorithm". *Technometrics* 59(1):11–22.

Sacks, J., W. J. Welch, T. J. Mitchell, and H. P. Wynn. 1989. "Design and Analysis of Computer Experiments". *Statistical Science* 4(4):409–423.

Santner, T. J., B. J. Williams, and W. Notz. 2003. *The Design and Analysis of Computer Experiments*. New York: Springer-Verlag.

Sobol, I. M. 1967. "On the Distribution of Points in a Cube and the Approximate Evaluation of Integrals". *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki* 7(4):784–802.