Fast Parameter Optimization Using Kriging Metamodeling

[1]E.S. Siah, [1]T. Ozdemir, [1,2]J.L. Volakis, [1]P. Papalambros and [3]R.Wiese
[1]The University of Michigan, Ann Arbor, Michigan 48109-2122
[2]ElectroScience Lab, Ohio State University, Columbus, OH 43212
[3]General Motors Corporation, 3300 GM Rd, Milford, MI 48380, USA
{esiah, tayfun, volakis}@eecs.umich.edu

Introduction

The last ten years has seen significant progress in fast numerical EM methods for electromagnetic modeling and simulation [1-3]. Coupled with the equally impressive progress recorded in the area of optimization techniques, true EM design and optimization is now within our reach. However, robustness of the optimization algorithms continues to be a formidable challenge. Despite its popularity in EM related design problems, evolution type algorithms suffer from slow convergence rate. In this paper, we propose a hybrid optimizer, which combines the Kriging macro-modeling [4-5] and the divided rectangles method (DIRECT) [6-7] to perform global optimization. The latter yields a deterministic answer with fast convergence rate and possesses local as well as global optimization properties. Two examples are presented, where DIRECT optimizer is combined with Kriging metamodeling. The first example involves optimization of the shape of a slot array Frequency Selective Surface (FSS). In the second example, the electromagnetic coupling between the harness and the FM antenna on a Deville model automobile is minimized. Both examples demonstrate the exceptional rate of convergence and the flexibility of the proposed optimizer.

Kriging Metamodeling

Polynomial fitting is often the choice in interpolating sampled data. However, for rapidly changing data, polynomial fits fail due to their oscillatory behavior. On the other hand, Kriging interpolation functions, first reported in 1951 in the analysis of mining data [4], exhibit much less oscillations and have been shown to provide better fitting in multi-dimensional domains. Kriging fitting utilizes the correlation between neighboring points to determine the overall function at an arbitrary point. Consider the one dimensional case:

$$Y(x) = f(x) + e(x) \tag{1}$$

where $Y(x)$ is the interpolated point, $f(x)$ is the true data and $e(x)$ denotes the error in the prediction. Polynomial interpolation function regards $e(x)$ as independent process whereas Kriging metamodel does not. Instead, $e(x)$ is modeled as zero mean Gaussian process. With this in mind, for a k dimensional problem, (1) can be written as

$$Y(\bar{x}) = \sum_{j=1}^{k} \beta_j f_j(\bar{x}) + Z(\bar{x}) \tag{2}$$

where $f_j(\bar{x})$ are the basis functions, β_j are the corresponding coefficients and $Z(\bar{x})$ is the zero mean, Gaussian-distributed error function that models the deviation from $Y(\bar{x})$. The covariance of the error function is in turn modeled as

$$Cov(Z(\bar{w}), Z(\bar{x})) = \sigma_z^2 R(\bar{w}, \bar{x}) \tag{3}$$

$$R(\bar{w}, \bar{x}) = \prod_{d=1}^{k} e^{-\theta^d \left| w^d - x^d \right|^{p^d}} \tag{4}$$

in which, σ_z^2 is a scaling factor known as the process variance that can be tuned to fit the given data and $R(\bar{w}, \bar{x})$ is the spatial correlation function (SCF). The vector $\bar{w}$ refers to the vector of given neighboring data points with respect to the vector $\bar{x}$, which refers to the stationary data point in the k^{th} dimension. θ^d is a measure of the influence of the surrounding data points on the

0-7803-7846-6/03/$17.00 ©2003 IEEE

predicted point, with larger values indicating a smaller degree of influence and thus a weaker covariance value. Finally, the parameter p determines the continuity of the function. The covariance and spatial correlation function R increase in cómplexity with the number of design variables. Before the application of the Kriging algorithm, the values of σ_z^2, θ and p are determined from an auxiliary optimization problem where the difference between the function values of the predicted and the given data points is minimized (referred to as Maximum Likelihood Estimation.) Figure 1 shows an example of a badly fit as well as a well fit Kriging metamodel for one-dimensional case. Either better fitted Kriging parameters or more data samples are needed to improve the fit in Figure 1(a).

DIRECT Global Optimizer

The DIRECT optimization algorithm is a derivative free, global algorithm that yields a deterministic and unique solution. Its attribute of possessing both local and global properties makes it ideal for fast convergence. An essential aspect of the DIRECT algorithm is the subdivision of the entire design space into hyper-rectangles, or hyper-cubes for multi-dimensional problems. The iteration starts by choosing the center of the design space as the starting point. Subsequently, at each iteration step, DIRECT selects and subdivides the set of hyper-rectangles (hyper-cubes) that are mostly likely to produce the smallest value of the objective function. This decision is based upon the Lipschitizian optimization theory, specifically the manipulation of the Lipschitizian constant. Mathematically, the Lipschitizian constant K satisfies the relation

$$|f(x_1)-f(x_2)| \le K\|x_1 - x_2\| \qquad x_1, x_2 \in domainR \tag{5}$$

where x_1 and x_2 lie within the entire design space and $f(x)$ is the objective function. The Lipschitzian process finds the global minimum point provided the constant K is specified to be greater than the largest rate of change of the objective function within the design space and that the objective function value is continuous. Within DIRECT, all possible values of the Lipschitizian constant K are used with the larger values of K chosen for global optimization (to find the basin of convergence) followed by smaller values of K for local optimizations within the basin of convergence.

An illustration of two-dimensional optimization by DIRECT is shown in Figure 2. First, the design space is divided into three rectangles. Next, the centers of new rectangles are evaluated. Then, the Lipschitz constant is used to select which rectangles will be further divided. The process is repeated until the termination criterion is met.

Example 1: Shape Optimization of Frequency Selective Surface (FSS)

The hybrid DIRECT optimizer (with Kriging metamodeling) is used to optimize the dimensions of the unit cell of a slot array FSS to achieve a prescribed pass-band reflection coefficient from 10.7 to 11.3GHz (see Figure 3). For this optimization problem, there are four design variables and four inequality constraints. These constraints are imposed so that the unit cell is always less than 1.2 cm and the slot is smaller than the unit cell by at least the length of one mesh element. A design of experiments is carried out over the entire design space to create the Kriging metamodel, which is in turn used by the DIRECT optimizer. The analyzer code for this example is a hybrid FE-BI solver.

The overall objective function F is defined as

$$F = \sum_{i=1}^{N} w_i |\Gamma_{in}|^2 + \sum_{j=1}^{M} w_j |\Gamma_{in} - 1|^2 \tag{6}$$

which is a sum of the 10dB reflection coefficients of the FSS both within ($|\Gamma_{in}|$) and outside ($|\Gamma_{in} - 1|$) the pass-band region, where w_i and w_j refer to the weights for the i^{th} in-band and j^{th} out of band frequency components, respectively. DIRECT optimizer converged after 112 iterations yielding following optimized dimensions: $x = 0.0995cm, y = 0.9243cm$, *unitx* = 1.1cm and *unity* = 1.1 cm.

Example 2: Reduction of Harness-FM Antenna Coupling

The coupling from a wire harness to a printed FM antenna located at the rear of an automobile is optimized to be within a range (Figure 4a). For demonstration purposes, a collection of metallic strips were used to represent the harness, which were placed just above the floor of the car. Sensors located at four ends of the harness are independently driven by complex voltage amplitudes, which are optimized such that the average field magnitude along the length of the FM antenna is between 9.73 and 9.77 μV/m. The problem has four variables (the sensor voltages) and four inequality constraints: $0 = |V_{1,2}| = 150$ μV and $100 = |V_{3,4}| = 200$ μV. The objective function for this problem is given by

$$(F) = \left| \frac{\sum_{j=1}^{4} \sum_{i=1}^{M} E_{ij}^{ANT}}{M} \right| - 9.75x10 \qquad (7)$$

where E_{ij}^{ANT} is the complex electric field amplitude at the i^{th} sample point along the antenna ($M=16$) due to the j^{th} sensor. The convergence history is shown in Figure 4(b). Eight iterations sufficed for the optimization to converge yielding the optimal sensor voltages of $V_1 = 20.5e^{j0}$ μV, $V_2 = 49.1e^{-j35.9}$μV, $V_3 = 97.6e^{-j33.3}$ μV, and $V_4 = 115.549.1e^{-j32.9}$μV.

Conclusions

We demonstrated that the proposed hybrid optimizer, which is a hybridization of the DIRECT global optimizer and the Kriging metamodeling, is capable of performing rapidly converging global optimizations. This puts the proposed optimizer at an advantage compared to those based on the genetic algorithms, which tend to suffer from slow convergence rates. Examples have been given where the optimizer consistently found the global minimum making it a robust optimizer.

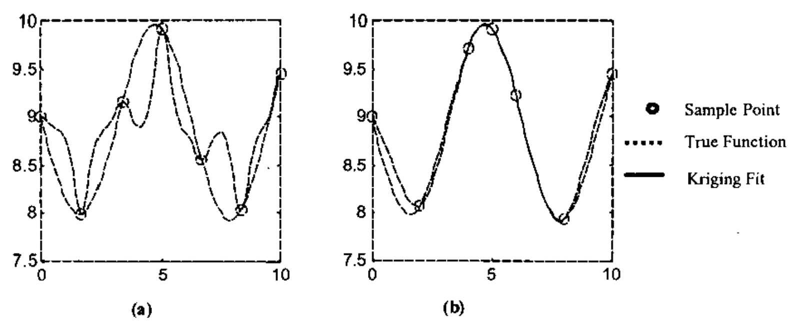


Figure 1: Comparison between the Kriging-approximation and the true function for a given number of sample points, (a) badly fitted Kriging metamodel, (b) well fitted Kriging metamodel.

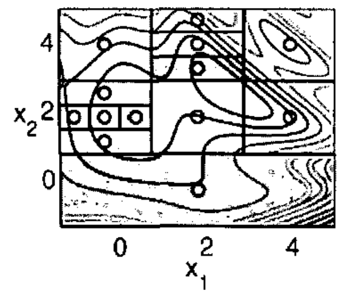


Figure 2: Two dimensional application of the DIRECT optimizer.

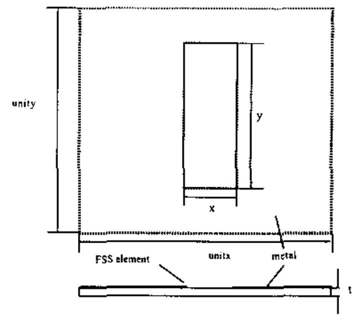
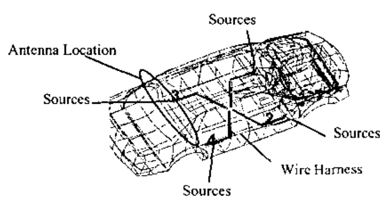
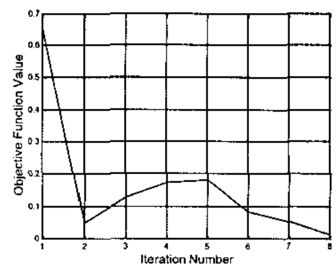


Figure 3: Unit cell of a slot array FSS.



(a)



(b)

Figure 4: (a) Deville CAD model shown together with the harness and the sensor locations, (b) Convergence of the optimizer.

References

[1] L.R. Hamilton, P.A. Macdonald, M.A. Stalzer, R.S. Turley, J.L. Visher and S.M. Wandzura, "3D method of moments scattering computations using the fast multipole method," *Antennas Propagat. Soc. Intl. Symp., Digest*, vol.1, pp. 435-438, 1994.

[2] J. Song, C.C. Lu and W.C. Chew, "Multilevel fast multipole algorithm for electromagnetic scattering by large complexobjects," *IEEE Trans. Antennas Propagat.*, Vol. 45, No. 10, pp.1488-1493, Oct. 1997.

[3] K. Sertel, J.L. Volakis, "Multilevel fast multipole method implementation using parametric surface modeling," *Antennas Propagat. Soc. Intl. Symp., Digest*, vol.4, pp. 1852-1855, 2000.

[4] D.G. Krige, A statistical approach to some mine problems and allied problems at the Witwatersrand, master's thesis, University of Witwatersrand, 1951.

[5] N. Cressie, "Spatial prediction and ordinary Kriging", *Mathematical Geology*, Vol. 40, No.4, pp.405-421, 1997.

[6] D.R. Jones, C.D. Perttunen and B.E. Stuckman, "Lipschitzian optimization without the Lipschitz constant", *J. Opt. Theory and Appl.*, Vol.79, No.1, pp.157-181, Oct 1993.

[7] D.R. Jones, The DIRECT global optimization algorithm, *Encyclopedia of Optimization*, 1999.