PROBABILITY OF IMPROVEMENT METHODS FOR CONSTRAINED MULTI-OBJECTIVE OPTIMIZATION

G. I. Hawe, J. K. Sykulski

School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, U.K.
e-mail: gih@soton.co.uk, jks@soton.ac.uk, fax: +44 (0)23 8059 3709

Keywords: Constrained Multi-Objective Optimization, Kriging.

Abstract

This paper shows how the simultaneous consideration of multiple Kriging models can lead to useful metrics for the selection of design vectors in constrained multi-objective optimization. The savings in computational cost with such methods make them particularly useful for optimal electromagnetic design.

1 Introduction

The constrained multi-objective optimization problem (CMOOP) may be phrased as:

$$\begin{array}{lll} \text{Minimize} & f_i(\mathbf{x}) & i = 1, \ldots, M \\ \text{subject to} & g_i(\mathbf{x}) \le 0 & i = 1, \ldots, J \\ & h_i(\mathbf{x}) = 0 & i = 1, \ldots, K \\ \text{where} & x_i^l \le x_i \le x_i^u & i = 1, \ldots, d \end{array} \tag{1}$$

where the design vector $\mathbf{x} = [x_1, x_2, \ldots, x_d]^T$. If a design vector satisfies all the constraints, it is feasible; otherwise it is infeasible. For any two feasible design vectors $\mathbf{x}^a$ and $\mathbf{x}^b$, if $f_i(\mathbf{x}^a) \le f_i(\mathbf{x}^b)\ \forall i$, and $\exists j$ such that $f_j(\mathbf{x}^a) < f_j(\mathbf{x}^b)$, then $\mathbf{x}^a$ is said to dominate $\mathbf{x}^b$; otherwise $\mathbf{x}^a$ is said to be non-dominated by $\mathbf{x}^b$. If $\mathbf{x}^a$ is non-dominated by $\mathbf{x}^b$ and $\mathbf{x}^b$ is non-dominated by $\mathbf{x}^a$ then $\mathbf{x}^a$ and $\mathbf{x}^b$ are said to be equivalent. All infeasible design vectors are deemed to be equivalent, and dominated by each of the feasible design vectors. The solution to (1) is the set of feasible design vectors which are non-dominated over the entire search space, known as the Pareto-optimal set.

In single-objective optimization, a single Kriging surface [7] may be constructed modelling the behaviour (with respect to the design variables) of the objective function in question. In [3], a variety of selection criteria are proposed for utilizing statistical information from a *single* Kriging model, for the purpose of selecting a design vector (or multiple design vectors) to evaluate in the search for the minimum. Methods also exist for incorporating constraint handling into these selection criteria, e.g. [9].

In multi-objective optimization, the multiple objectives may be combined into a single objective using a scalarizing method [8], and a method from single-objective optimization used for selecting design vectors, e.g. [1, 6]. Alternatively, each objective may be modelled by its own individual Kriging surface; this allows the uncertainty in each objective to be modelled separately. The simultaneous consideration of these uncertainties then allows useful metrics for selecting design vectors to be constructed [2, 5], which do not suffer from the loss of information which inevitably occurs when using scalarizing methods. This paper proposes a method of extending the handling of non-linear constraints into such metrics.

2 Probability of Improvement with Constraints

Suppose that after sampling the design variable space (preferably using a space-filling experimental design, such as a Hammersley Sequence [4]), a set S of N_{par} non-dominated solutions exist (each of these solutions are both feasible, and non-dominated by the solutions not in S). Then it is desirable when sampling again to select a design vector which either dominates at least one (preferably more) solution in S *whilst being feasible*, or at the very least augments the set S (i.e. is equivalent to each solution in S *and is feasible*). In either case such a selection could be said to yield an improvement, as our solution set has improved. In *unconstrained* multi-objective optimization, by constructing Kriging surfaces for each objective individually, not only may the probability of an unevaluated design vector dominating at least one solution in S be calculated [5], but the probability of it dominating a particular number of solutions in S (representing a specific *level* of improvement), may also be determined [2]. This is illustrated schematically in Figure 1 for the case of a two-objective problem with $N_{\text{par}}=5$ non-dominated solutions. In the presence of constraints however, it is crucial to also ensure the feasibility of the design vectors being selected.

As can be seen from Figure 1, for an improvement level n (achieving a level n improvement means that a design vector dominates n solutions in S, with a level 0 improvement meaning that it is equivalent to each solution in S), a design vector may map to $N_{\text{par}}-n+1$ regions of objective function space. Denote by $P(I^n(\mathbf{x}))$ the probability that an unknown design vector $\mathbf{x}$ will yield a level of improvement n (i.e. it will dominate exactly n

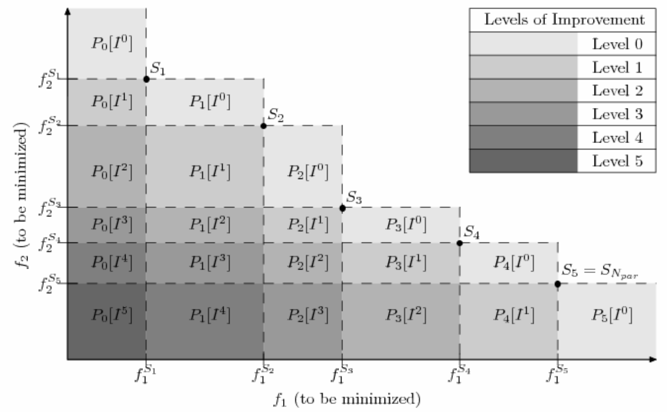

Figure 1: Levels of improvement for a two-objective problem with 5 non-dominated solutions.

existing non-dominated solutions *and* be feasible), and $P_i(I^n(\mathbf{x}))$ as the probability that design vector $\mathbf{x}$ will dominate the n non-dominated solutions S_{i+1}, S_{i+2}, … , S_{i+n} (these sub-regions are labelled in Figure 1) *and* be feasible. Define

$$\Phi_p^i(\mathbf{x}) = \Phi\left(\frac{f_p^{S_i} - \hat{f}_p(\mathbf{x})}{s_p(\mathbf{x})}\right) \quad (2)$$

for p=1,2, with

$$\Phi_1^0(\mathbf{x}) = 0, \; \Phi_1^{N_{par}+1}(\mathbf{x}) = 1 \quad (3), (4)$$

$$\Phi_2^0(\mathbf{x}) = 1, \; \Phi_2^{N_{par}+1}(\mathbf{x}) = 0 \quad (5), (6)$$

where $\hat{f}_p(\cdot)$ and $s_p(\cdot)$ are the Kriging prediction and standard error for each objective ($p=1,2$), and $f_p^{S_i}$ is the value of objective p for non-dominated solution S_i. Then

$$P(I^n(\mathbf{x})) = \prod_{i=1}^{J} \Phi\left(-\frac{\hat{g}_i(\mathbf{x})}{s_{g_i}(\mathbf{x})}\right) \times \prod_{j=1}^{K}\left[\Phi\left(\frac{\varepsilon_j - \hat{h}_j(\mathbf{x})}{s_{h_j}(\mathbf{x})}\right) + \Phi\left(\frac{\hat{h}_j(\mathbf{x}) - \varepsilon_j}{s_{h_j}(\mathbf{x})}\right)\right] \times \sum_{l=0}^{N_{par}-n}\left(\Phi_1^{l+1}(\mathbf{x}) - \Phi_1^{l}(\mathbf{x})\right)\left(\Phi_2^{n+l}(\mathbf{x}) - \Phi_2^{n+l+1}(\mathbf{x})\right) \quad (7)$$

where $\hat{g}_i(\cdot)$ ($i = 1,\ldots J$) and $\hat{h}_i(\cdot)$ ($i=1,\ldots,K$) are the Kriging predictions for each of the individual constraint functions, s_{g_i} and s_{h_i} are their standard errors, and ε_i ($i=1,\ldots,K$) are small tolerances chosen to transform each equality constraint into two inequality constraints.

In words, the design vector which maximizes the expression in (7) is that which is most likely to dominate n of the current non-dominated solutions, *and* be feasible. In total this gives $N_{\text{par}}+1$ different levels of improvement to maximize at each iteration; by grouping the design vectors which maximize each of these levels of improvement into clusters (using, e.g. the method proposed in [3]), and selecting a representative design vector from each cluster, a robust method (which is easily parallelized) – using the probability of improvement method – is made available for constrained multi-objective optimization. Although the description given here is for two-objective problems, the method is extensible to higher numbers of objectives.

3 Conclusions

A novel utility function, which is easily parallelized and which does not require normalization of the objective functions, has been proposed for computationally expensive constrained multi-objective optimization. It is ideally suited for applications such as electromagnetic design.

References

[1] G. I. Hawe, J. K. Sykulski. "A scalarizing one-stage algorithm for efficient multi-objective optimization", *To appear in IEEE Transactions in Magnetics*, (2008).

[2] G. I. Hawe, J. K. Sykulski. "An enhanced probability of improvement utility function for locating Pareto-optimal solutions", *Proceedings of the 16th Conference on the Computation of Electromagnetic Fields, COMPUMAG 2007, pp. 965 – 966,* (2007).

[3] D. R. Jones. "A taxonomy of global optimization methods based on response surfaces", *Journal of Global Optimization*, **21**, pp. 345-383, (2001).

[4] J. R. Kalagnanam, U. M. Diwekar. "An efficient sampling technique for off-line quality control", *Technometrics*, **39 (2)**, pp. 131-148, (1997).

[5] A. J. Keane. "Statistical improvement criteria for use in multiobjective design optimization", *AIAA Journal*, **44 (4)**, pp. 879-891, (2006).

[6] J. Knowles. "ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multi-objective optimization problems", *IEEE Transactions on Evolutionary Computation Journal*, **10 (1)**, pp. 50-66, (2006).

[7] L. Lebensztajn, C. A. R. Marretto, M. C. Costa, J-L. Coulomb. "Kriging: a useful tool for electromagnetic devices optimization", *IEEE Transactions on Magnetics*, **40 (2)**, pp. 1196-1199, (2004).

[8] K. M. Miettinen. *Nonlinear Multiobjective Optimization*, Kluwer Academic Publishers, (1999).

[9] M. J. Sasena, P. Y. Papalambros, P. Goovaerts. "Exploration of metamodeling sampling criteria for constrained global optimization", *Engineering Optimization*, **34 (3)**, pp. 263-278, (2002).