BRIEF NOTE

Sequential sampling for contour estimation with concurrent function evaluations

**Felipe A. C. Viana · Raphael T. Haftka ·
Layne T. Watson**

Abstract We demonstrate the use of multiple surrogates and kriging believer for parallelizing surrogate-based contour estimation. For the demonstration example, we reduce wall clock time with minimal penalty in number of simulations.

Keywords Design and analysis of computer experiments · Metamodeling · Gaussian process

1 Introduction

Surrogate use for constrained optimization and design under uncertainty has become popular (Basudhar and Missoum 2010; Ranjan et al. 2008; Kuczera et al. 2010; Valdebenito and Schuëller 2010; Picheny et al. 2010; Lee et al. 2011; Dubourg et al. 2011). Surrogates fitted to constraints must be most accurate near the contour demarcating the boundary of the feasible domain (limit state). Thus, algorithms

F. A. C. Viana (✉) · R. T. Haftka
Department of Mechanical and Aerospace Engineering,
University of Florida, Gainesville, FL 32611-6250, USA
e-mail: felipeacviana@gmail.com

R. T. Haftka
e-mail: haftka@ufl.edu

L. T. Watson
Departments of Computer Science and Mathematics,
Virginia Polytechnic Institute & State University, Blacksburg,
VA 24061, USA
e-mail: ltw@cs.vt.edu

that prefentially sample near the limit state have been developed—e.g., the efficient global reliability analysis (EGRA) algorithm (Bichon et al. 2008). These algorithms add one point at a time, not taking advantage of parallel computing for concurrent constraint evaluations.

We propose two strategies for parallelizing the sampling, both based on the efficient global optimization (EGO) algorithm (Jones et al. 1998). The first is to fit multiple surrogates to the data (Viana et al. 2010), each contributing one sample point. The second is based on the "kriging believer" approach (Ginsbourger et al. 2010), where after selecting a new sample, kriging is updated using the estimated value as if it were data.

2 Sequential sampling for contour estimation with the Efficient Global Reliability Analysis (EGRA) algorithm

Given the scalar constraint $g(\mathbf{x}) \leq \bar{g}$ defining the limit state $\bar{g}$, EGRA (Bichon et al. 2008) defines the feasibility[1] at $\mathbf{x}$ as

$$F(\mathbf{x}) = \max\big(\epsilon(\mathbf{x}) - |\bar{g} - G(\mathbf{x})|, 0\big), \tag{1}$$

where $G(\mathbf{x})$ is the random variable representing the unknown $g(\mathbf{x})$, and $\epsilon(\mathbf{x})$ measures the uncertainty in $G(\mathbf{x})$.

$F(\mathbf{x})$ is nonzero in a band near the limit state and maximal where the uncertainty $\epsilon(\mathbf{x})$ is largest. If $G(\mathbf{x})$ is

---

[1] An alternative definition is found in Ranjan et al. (2008) and Ranjan et al. (2011)

modeled by a kriging surrogate (Gaussian process), then the expectation of $F(\mathbf{x})$ is easy to calculate[2] as

$$
\begin{aligned}
E_G[F(\mathbf{x})] = {}& (\hat{g}(\mathbf{x}) - \bar{g}) \left[2\Phi(u) - \Phi(u^+) - \Phi(u^-) \right] \\
& - s(\mathbf{x}) \left[2\phi(u) - \phi(u^+) - \phi(u^-) \right] \\
& + \epsilon \left[\Phi(u^+) - \Phi(u^-) \right] ,
\end{aligned}
\tag{2}
$$

where Φ and ϕ are the standard Gaussian CDF and PDF, respectively, $u = \dfrac{\bar{g} - \hat{g}(\mathbf{x})}{s(\mathbf{x})}$, $u^+ = \dfrac{\bar{g} + \epsilon - \hat{g}(\mathbf{x})}{s(\mathbf{x})}$, $u^+ = \dfrac{\bar{g} - \epsilon - \hat{g}(\mathbf{x})}{s(\mathbf{x})}$, $\hat{g}(\mathbf{x})$ is the kriging prediction, $s(\mathbf{x})$ is the square root of the kriging prediction variance, and $\epsilon = \alpha s(\mathbf{x})$ (Bichon et al. 2008 suggested $\alpha = 2$).

The efficient global reliability analysis (EGRA) algorithm sequentially adds points to the data set in an effort to improve the surrogate accuracy near the limit state. As an adaptive sampling strategy, EGRA involves solving the optimization problem of finding the best point to be sampled. In each cycle, the next point to be sampled is the one that maximizes[3] $E_G[F(\mathbf{x})]$. Kriging is updated after adding the new point to the existing data set. EGRA iterates until a convergence criterion is met (e.g., computational budget or $E_G[F(\mathbf{x})]$ history).

3 Strategies for multiple point EGRA

The first strategy we propose to provide multiple points per cycle is the multiple surrogate efficient global reliability analysis (MSEGRA) algorithm. MSEGRA is inspired by the multiple surrogate efficient global optimization algorithm (Viana et al. 2010) and uses n surrogates to provide up to n points per cycle. In each cycle, we maximize $E_G[F(\mathbf{x})]$ of each surrogate. MSEGRA could use different instances of kriging (with different correlation functions). However, we illustrate the algorithm with surrogates based on different techniques (radial basis function, linear Shepard, and support vector regression).[4]

The second strategy uses the kriging believer heuristic (Ginsbourger et al. 2010). In each cycle, n points are created by sequentially maximizing $E_G[F(\mathbf{x})]$ (n times). When

---

[2] That is $E_G[F(\mathbf{x})] = \int_{\bar{g} - \epsilon(\mathbf{x})}^{\bar{g} + \epsilon(\mathbf{x})} \left[\epsilon(\mathbf{x}) - |\bar{g} - G(\mathbf{x})| \right] f_G(g) dg$. Further discussion and derivation can be found in Bichon (2010).

[3] Here, we used differential evolution as implemented in the companion software of Price et al. (2005) to solve this optimization problem.

[4] This version runs EGRA with surrogates that might not furnish uncertainty estimates. These estimates can certainly be provided by sophisticated schemes, e.g. the Bayesian approach (Seok et al. 2002). Here, we use the kriging uncertainty estimates with all other surrogates (Viana and Haftka 2009). Although theoretically less attractive, this heuristic avoids the overhead of estimating the uncertainty for each surrogate.

a point is obtained, via maximization of $E_G[F(\mathbf{x})]$, the kriging prediction is "believed" (taken as new data) and kriging is updated. In other words, the kriging believer replaces the actual response at the chosen sites by the kriging prediction. In practice, this makes the kriging uncertainty vanish at that point—and thus also $E_G[F(\mathbf{x})]$. At the end of the cycle, actual simulations are conducted at the n new points (which are added to the data set). Kriging is then updated with the augmented data set.

While kriging believer find samples in sequence, in MSEGRA they can be found in parallel. Therefore, when search for samples (and surrogate updates) dominates the cost of function evaluations, MSEGRA may have some gains of parallelization. When using multiple surrogates, selection is done after sampling is finished (e.g., using cross-validation (Viana et al. 2009)).

4 Numerical experiments

As test problem, we used the Branin-Hoo function (Dixon and Szegö 1978)

$$
\begin{aligned}
g(\mathbf{x}) = {}& \left(x_2 - \frac{5.1 x_1^2}{4\pi^2} + \frac{5 x_1}{\pi} - 6 \right)^2 \\
& + 10 \left(1 - \frac{1}{8\pi} \right) \cos(x_1) + 10 \le 50, \\
& -5 \le x_1 \le 10 , \text{ and } 0 \le x_2 \le 15.
\end{aligned}
\tag{3}
$$

Table 1 Setup for the surrogates used

Surrogates	Details
(1) krg	Kriging: constant trend and Gaussian correlation. Kriging parameters $\boldsymbol{\theta}$ are optimized within $10^{-3} \le \theta_i \le 2 p^{1/d}$, $i = 1, \ldots, d$ (p is the number of points in the d-dimensional space) using $\theta_i = p^{1/d}$ as initial guess.
(2) rbf-mq	Radial basis functions: 'mq' and 'g' stand for
(3) rbf-g	multiquadric and Gaussian basis function, respectively (both using $spread = 2$).
(4) shep	Linear Shepard model: Subroutine LSHEP from SHEPPACK (Thacker et al. 2010).
(5) svr	Support vector regression: with Exponential kernel function and ϵ-insensitive loss function ($\epsilon = 0$ and $C = \infty$).

The DACE (Lophaven et al. 2002), RBF (Jekabsons 2009), SURRO-GATES (Viana 2011), and SVM (Gunn 1997) toolboxes were used to run the kriging, radial basis function, linear Shepard, and support vector regression algorithms, respectively. In these implementations, only kriging furnishes an uncertainty estimate. All other surrogates import the kriging uncertainty

Table 2 Setup for the sequential sampling

Strategy	Points per cycle	Points added
EGRA	1	20
MSEGRA	5	100
EGRA-believer	5	100

The SURROGATES toolbox (Viana 2011) is used to run EGRA, MSEGRA (with all surrogates of Table 1) and EGRA-believer for 20 cycles

The surrogates of Table 1 are initially fitted with ten points. To average out the influence of the initial data, we repeat the experiments with 100 different Latin hypercube designs created with the MATLAB® function *lhsdesign* (Mathworks contributors 2004) (set with "maxmin" and 1,000 iterations).

We run EGRA, MSEGRA and EGRA-believer for 20 cycles (details in Table 2) tracking the performance with test points and the misclassification fraction (Basudhar and Missoum 2010)

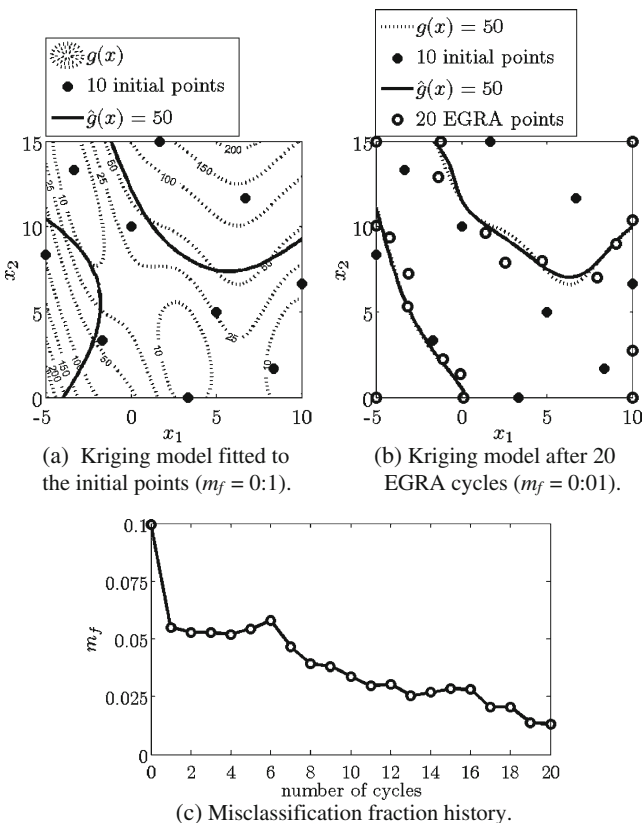$$m_f = \frac{\sum_i I(g(\mathbf{x}_i) \leq \bar{g}) \oplus I(\hat{g}(\mathbf{x}_i) \leq \bar{g})}{p_{\text{test}}} , \quad (4)$$

where $I(.)$ is the indicator function, $\oplus$ represents the exclusive disjunction (which is 1 if only one of the operands is true and 0 otherwise), and $p_{\text{test}} = 10,000$ (created with the *lhsdesign* set with "maxmin" and 50 iterations). m_f indicates the fraction of the design space the surrogate misclassifies as either feasible or infeasible.

5 Results and discussion

We illustrate the EGRA using a single data set. Figure 1a shows kriging fitted to ten points (misclassification fraction $m_f = 0.1$). Figure 1b illustrates results after 20 cycles. The misclassification fraction is reduced to $m_f = 0.01$. Figure 1c shows changes in misclassification fraction as points are added one at a time.

Figure 2 shows the median of m_f with 100 experimental designs for the three EGRA variants. Figure 2a shows that while MSEGRA and EGRA-believer have approximately $m_f = 0.007$ by the third cycle, EGRA only achieves that mark by the 12th cycle. If function evaluations can be done in parallel (concurrently), MSEGRA and EGRA-believer save about ten cycles (many days for some applications).
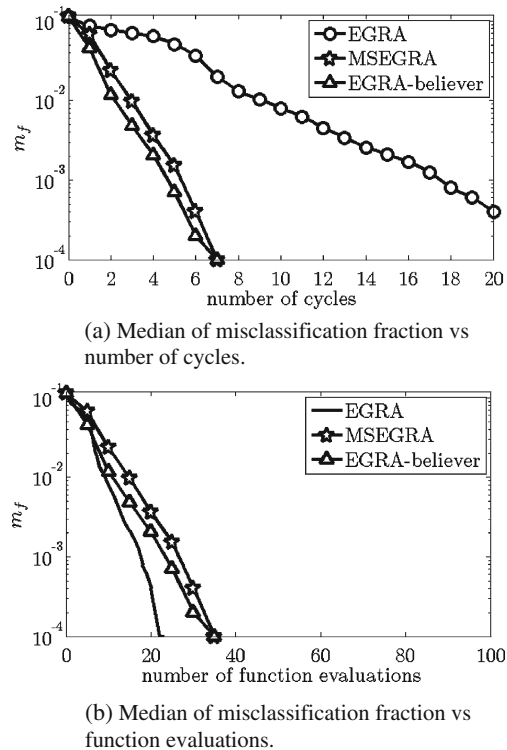

(a) Kriging model fitted to the initial points ($m_f = 0.1$).

(b) Kriging model after 20 EGRA cycles ($m_f = 0.01$).

(c) Misclassification fraction history.

Fig. 1 EGRA for a single experimental design. Points added after 20 EGRA cycles reduce misclassification (4)

(a) Median of misclassification fraction vs number of cycles.

(b) Median of misclassification fraction vs function evaluations.

Fig. 2 Median (over 100 experimental designs) of misclassification fraction (4) for EGRA variants of Table 2. Five points per cycle improve performance without a severe penalty in terms of number of function evaluations

Using multiple function evaluations per cycle is attractive when the main concern is the number of cycles, rather than number of simulations. However, one would expect a cost associated with the number function evaluations. Figure 2b shows the median of misclassification fraction versus the number of function evaluations for all three algorithms. For this problem, the penalty in increased number of evaluations is small.

6 Summary and conclusions

We proposed two approaches for contour estimation with concurrent function evaluations. The multiple surrogate efficient global reliability analysis (MSEGRA) algorithm uses a set of surrogates. EGRA-believer runs the kriging believer heuristic. We demonstrated the two approaches for an algebraic example and surrogates including kriging, radial basis function, linear Shepard and support vector regression. We found that

1. MSEGRA and EGRA-believer reduced substantially the number of cycles required for convergence, and
2. the penalty in terms of total number of function evaluations was less than a factor of two.

Results for the demonstration example need to be augmented with other problems to provide a more solid estimate of the savings associated with the procedure. In particular, the optimal number of simultaneous evaluations needs to be studied as function of the cost of function evaluations and the benefit of short wall times. The multiple surrogate version also allows one to select a single surrogate after the sampling is concluded based on accuracy measures, such as cross validation errors.

References

Basudhar A, Missoum S (2010) An improved adaptive sampling scheme for the construction of explicit boundaries. Struct Multidisc Optim 42(4):517–529

Bichon BJ (2010) Efficient surrogate modeling for reliability analysis and design. Ph.D. thesis, Vanderbilt University, Nashville

Bichon BJ, Eldred MS, Swiler LP, Mahadevan S, McFarland J (2008) Efficient global reliability analysis for nonlinear implicit performance functions. AIAA J 46(10):2459–2468

Dixon LCW, Szegö GP (1978) Towards global optimization 2. North Holland

Dubourg V, Sudret B, Bourinet JM (2011) Reliability-based design optimization using kriging surrogates and subset simulation. Struct Multidisc Optim (online first):1–18. doi:10.1007/s00158-011-0653-8

Ginsbourger D, Le Riche R, Carraro L (2010) Kriging is well-suited to parallelize optimization. In: Tenne Y, Goh C-K, Hiot LM, Ong YS (eds) Computational intelligence in expensive optimization problems. Adaptation, learning, and optimization, vol 2. Springer, Berlin, pp 131–162

Gunn SR (1997) Support vector machines for classification and regression. University of Southampton, UK. Available at http://www.isis.ecs.soton.ac.uk/resources/svminfo/. Accessed 8 April 2011

Jekabsons G (2009) RBF: radial Basis Function interpolation for MATLAB/OCTAVE. Riga Technical University, Latvia. Available at http://www.cs.rtu.lv/jekabsons/regression.html. Accessed 8 April 2011

Jones DR, Schonlau M, Welch WJ (1998) Efficient global optimization of expensive black-box functions. J Glob Optim 13(4):455–492

Kuczera R, Mourelatos Z, Nikolaidis E (2010) A re-analysis methodology for system RBDO using a trust region approach with local metamodels. SAE International Journal of Materials & Manufacturing 3(1):356–371

Lee I, Choi KK, Zhao L (2011) Sampling-based RBDO using the stochastic sensitivity analysis and dynamic kriging method. Struct Multidisc Optim 44(3):299–317

Lophaven SN, Nielsen HB, Søndergaard J (2002) Dace—a matlab kriging toolbox. Technical University of Denmark, Denmark. Available at http://www2.imm.dtu.dk/~hbn/dace/. Accessed 8 April 2011

Mathworks contributors (2004) MATLAB The language of technical computing. The MathWorks, Inc, version 7.0 release 14 edn

Picheny V, Ginsbourger D, Roustant O, Haftka RT, Kim N (2010) Adaptive design of experiments for accurate approximation of a target region. J Mech Des 132(7):9

Price KV, Storn RM, Lampinen JA (2005) Differential evolution: a practical approach to global optimization. Springer Verlag

Ranjan P, Bingham D, Michailidis G (2008) Sequential experiment design for contour estimation from complex computer codes. Technometrics 50(4):527–541

Ranjan P, Bingham D, Michailidis G (2011) Errata. Technometrics 53(1):109–110

Seok K, Hwang C, Cho D (2002) Prediction intervals for support vector machine regression. Commun Stat, Theory Methods 31(10):1887–1898

Thacker WI, Zhang J, Watson LT, Birch JB, Iyer MA, Berry MW (2010) Algorithm 905: Sheppack: modified Shepard algorithm for interpolation of scattered multivariate data. ACM Trans Math Softw 37(3):1–20

Valdebenito MA, Schuëller GI (2010) A survey on approaches for reliability-based optimization. Struct Multidisc Optim 42(5):645–663

Viana FAC (2011) SURROGATES toolbox user's guide. Available at http://sites.google.com/site/felipeacviana/surrogatestoolbox. Accessed 8 April 2011

Viana FAC, Haftka RT, Steffen Jr V (2009) Multiple surrogates: how cross-validation errors can help us to obtain the best predictor. Struct Multidisc Optim 39(4):439–457

Viana FAC, Haftka RT (2009) Importing uncertainty estimates from one surrogate to another. In: 50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, AIAA–2009–2237

Viana FAC, Haftka RT, Watson LT (2010) Why not run the efficient global optimization algorithm with multiple surrogates? In: 51st AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, AIAA–2010–3090