European Journal of Operational Research 000 (2017) 1–11

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Stochastics and Statistics

Sequential design strategies for mean response surface metamodeling via stochastic kriging with adaptive exploration and exploitation

Xi Chen [a,*], Qiang Zhou [b]

[a] Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, VA 24601, USA
[b] Department of Systems and Industrial Engineering, University of Arizona, Tucson, AZ 85721, USA

ARTICLE INFO

Article history:
Received 14 May 2016
Accepted 15 March 2017
Available online xxx

Keywords:
Simulation
Sequential experimental design
Simulation metamodeling
Simulation analysis and methodology

ABSTRACT

Stochastic kriging (SK) methodology has been known as an effective metamodeling tool for approximating a mean response surface implied by a stochastic simulation. In this paper we provide some theoretical results on the predictive performance of SK, in light of which novel integrated mean squared error-based sequential design strategies are proposed to apply SK for mean response surface metamodeling with a fixed simulation budget. Through numerical examples of different features, we show that SK with the proposed strategies applied holds great promise for achieving high predictive accuracy by striking a good balance between exploration and exploitation.

Published by Elsevier B.V.

1. Introduction

Complex computer simulation models of proposed or existing real systems are often used to aid system design. This usually happens when it is impractical to construct multiple prototype versions of the real system, or other constraints such as cost or time prohibit experimentation with the real system. Analysts often use the simulation model as a surrogate to do the system design. However, simulation models themselves can be quite complicated, and the decision to build and use a simulation model of a large-scale, complex system represents a nontrivial investment of time and money. Furthermore, in applications where intense simulation is necessary to evaluate even one scenario or where it is required to learn the impact of many "what if" scenarios, simulation cannot deliver the desired answer in a timely manner. As a result, a metamodel is often built on outputs from simulations run at some selected design points to "map" the performance response surface as a function of the controllable decision variables, or uncontrollable environmental variables, to approximate the behavior of the original simulation model. This metamodel can be used as an accurate, drop-in replacement for the simulation model as if the simulation can be run "on demand." Successful applications have been recorded in many cases, including urban transportation analysis (Osorio & Bierlaire, 2013), polymerization reaction process study (Ouyang, Ma, Wang, & Tu, 2017) and analysis of manufacturing process with switching output regimes (Santos & Santos, 2016).

The kriging methodology has been very popular in various engineering disciplines for approximating the output of deterministic computer experiments (i.e., the same output is produced if the simulation is run twice at the same design point); see, for instance, Santner, Williams, and Notz (2003). To build a high-quality metamodel such as kriging with a given simulation budget to expend, a carefully designed simulation experiment is required. The literature on experimental designs for deterministic computer experiments abounds and various non-sequential design schemes have been proposed, for instance, Latin hypercube designs (LHDs) (McKay, Beckman, & Conover, 1979), orthogonal array based LHDs (Tang, 1993), uniform designs (Fang, Lin, Winker, & Zhang, 2000), and maximum entropy designs (Shewry & Wynn, 1987); see more details from Section 5.5 of Kleijnen (2015). More recently, Gauthier and Pronzato (2014, 2016) investigate spectral approximations of the integrated mean squared error (IMSE) and the truncated-IMSE and obtain efficient approximations of IMSE-optimal quadrature designs. Compared to the aforementioned non-sequential designs that choose all design points up front, sequential designs can offer a huge advantage in that they improve budget allocation efficiency by learning information from previous experiment runs and allocating the remaining simulation budget more wisely.

Sequential design strategies for global metamodeling have been studied in the context of deterministic computer experiments, to perform either prediction or optimization (Kleijnen, 2017). With respect to metamodeling for prediction, the interest lies in constructing sequential designs for a metamodel to achieve high predictive

* Corresponding author.
E-mail addresses: xchen6@vt.edu (X. Chen), zhouq@email.arizona.edu (Q. Zhou).

http://dx.doi.org/10.1016/j.ejor.2017.03.042
0377-2217/Published by Elsevier B.V.

Please cite this article as: X. Chen, Q. Zhou, Sequential design strategies for mean response surface metamodeling via stochastic kriging with adaptive exploration and exploitation, European Journal of Operational Research (2017), http://dx.doi.org/10.1016/j.ejor.2017.03.042

accuracy across the entire design space. For example, Ranjan, Bingham, and Michailidis (2008) propose a sequential design method for estimating a contour (also called a level set of iso-surface) of a computer code. In reliability engineering, Bect, Ginsbourger, Li, Picheny, and Vazquez (2012) estimate the probability of failure with a sequential strategy using stepwise uncertainty reduction. Kleijnen and van Beers (2004) propose an application-driven sequential strategy using cross-validation and jackknifing to estimate the variances of computer outputs for candidate design points. Their approach tends to add more design points to those design regions with interesting input/output behavior. As to metamodeling for optimization, the focus is on establishing sequential methods to seek the global optimum. The pioneering work that introduces the efficient global optimization (EGO) approach (Jones, Schonlau, & Welch, 1998) has inspired the development of many EGO variants for solving various types of optimization problems; the interested reader is referred to details from Section 6.2 of Kleijnen (2017) and Section 6.3 of Kleijnen (2015) and references therein.

In recent years, kriging related research for stochastic simulation (or random computer experiments) has flourished, leading to a plethora of theoretical and empirical work on a wide range of topics. A majority of these studies have been dedicated to efficiently metamodeling the mean response surface implied by a stochastic simulation, e.g., van Beers and Kleijnen (2008), Ankenman, Nelson, and Staum (2010), Ng and Yin (2012) and Mehdad and Kleijnen (2015). There also exist research studies that either focus on jointly metamodeling the underlying mean and simulation variance response surfaces (e.g., Kersting, Plagemann, Pfaff, & Burgard, 2007; Boukouvalas & Cornford, 2009; Robinson, Birch, & Starnes, 2010; Marrel, Iooss, da Veiga, & Ribatet, 2012; Wang & Chen, 2016; Binois, Gramacy, & Ludkovski, 2016) or concern about quantile-based response surface approximation (e.g., Yang, Ankenman, & Nelson, 2008; Chen & Kim, 2013; Bekki, Chen, & Batur, 2014; Picheny, Ginsbourger, Richet, & Caplin, 2013; Chen & Kim, 2016), to better cope with heteroscedasticity present in the stochastic simulation outputs.

The *stochastic kriging* (SK) methodology proposed by Ankenman et al. (2010) distinguishes itself as an effective metamodeling tool for approximating a mean response surface by correctly accounting for both sampling uncertainty inherent in a stochastic simulation and the response-surface uncertainty. Recently, fruitful results have been achieved both with respect to in-depth theoretical investigations and regarding successful applications of SK (e.g., Liu & Staum, 2010; Chen, Ankenman, & Nelson, 2012; 2013; Chen & Kim, 2014; 2016).

With respect to metamodeling for mean response prediction in the stochastic simulation setting, some in-depth work has been conducted regarding asymptotic properties of kriging prediction with nugget and the implications on non-sequential experimental designs with a large simulation budget are provided (Gratiet & Garnier, 2015). However, a systematic account of sequential designs with a fixed simulation budget has yet to be established, despite the earlier efforts by van Beers and Kleijnen (2008), Ng and Yin (2012), Ajdari and Mahlooji (2014) and Mehdad and Kleijnen (2015). In particular, very little work has been done on devising efficient experimental designs for applying the SK methodology. There exist some sequential design strategies proposed by Chen and Zhou (2014), nevertheless their results are preliminary and lack a theoretical grounding.

In this paper we aim to provide a first step toward establishing efficient sequential design strategies to implement SK for mean response prediction with a fixed simulation budget. We provide some theoretical results on the predictive performance of SK, in light of which we propose novel IMSE-based sequential design strategies that show great promise in striking a good balance between exploration (namely, selecting new design points to perform simulations for reducing response-surface uncertainty) and exploitation (namely, allocating simulation budget to existing design points for reducing sampling uncertainty). The remainder of this paper is organized as follows. Section 2 provides a review on SK. Section 3 studies some properties of SK theoretically, which leads to the proposed sequential design strategies for implementing SK in Section 4. The predictive performances of SK with different design strategies used are demonstrated through numerical examples in Section 5. Section 6 concludes the paper.

2. A review on stochastic kriging

We provide a brief review on the stochastic kriging (SK) methodology in this subsection. SK assumes the following model to represent the simulation output obtained on the jth simulation replication at design point $\mathbf{x} \in \mathscr{X} \subset \mathbb{R}^d$,

$$\mathscr{Y}_j(\mathbf{x}) = \mathsf{Y}(\mathbf{x}) + \varepsilon_j(\mathbf{x}) = \mathbf{f}(\mathbf{x})^\top \boldsymbol{\beta} + \mathsf{M}(\mathbf{x}) + \varepsilon_j(\mathbf{x}) \ , \tag{1}$$

where $\mathsf{Y}(\mathbf{x})$ denotes the unknown performance-measure surface of interest, and $\varepsilon_j(\mathbf{x})$ represents the simulation noise that has mean zero and variance $\mathsf{V}(\mathbf{x}) := \mathrm{Var}(\varepsilon_j(\mathbf{x}))$. Furthermore, $\mathbf{f}(\mathbf{x})$ is a $p \times 1$ vector of known functions of $\mathbf{x}$ and $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown parameters.

As treated in the design and analysis of deterministic computer experiments literature (Santner et al., 2003), SK postulates that $\mathsf{M}(\cdot)$ represents a second-order stationary mean-zero Gaussian random field. That is, $\mathsf{M}(\mathbf{x})$ can be regarded as being sampled randomly from a space of mappings $\mathbb{R}^d \to \mathbb{R}$, in which functions are assumed to exhibit spatial correlation. Specifically, there exists a spatial correlation function $\mathcal{R}(\cdot; \boldsymbol{\theta})$ that measures the correlation of the values of $\mathsf{M}(\mathbf{x}_i)$ and $\mathsf{M}(\mathbf{x}_\ell)$. This correlation is determined by the distance between $\mathbf{x}_i$ and $\mathbf{x}_\ell$ measured along each of the d dimensions, and the $d \times 1$ parameter vector $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_d)^\top$ controls how quickly the spatial correlation diminishes as the two points become farther apart in each direction. Commonly used correlation functions include the Gaussian correlation function, Matérn correlation functions, and the exponential correlation function (see Chapter 4 of Rasmussen & Williams, 2006); we choose to use the popular Gaussian correlation function $\mathcal{R}(\mathbf{x}_i - \mathbf{x}_\ell; \boldsymbol{\theta}) = \exp(-\sum_{r=1}^d \theta_r (x_{ir} - x_{\ell r})^2)$ in this paper. Given a correlation function, the implied covariance function is

$$\mathrm{Cov}(\mathsf{M}(\mathbf{x}_i), \mathsf{M}(\mathbf{x}_\ell)) = \tau^2 \mathcal{R}(\mathbf{x}_i - \mathbf{x}_\ell; \boldsymbol{\theta}) \ , \tag{2}$$

where τ^2 denotes the variance of $\mathsf{M}(\mathbf{x})$ for all $\mathbf{x} \in \mathscr{X}$. We note that in contrast to the stochastic nature of M referred to as *extrinsic uncertainty* by Ankenman et al. (2010), the *intrinsic uncertainty* represented by ε is inherent in a stochastic simulation output; we further assume that M and ε are independent.

In the stochastic simulation setting, an experimental design consists of design-point locations and the corresponding numbers of replications to conduct simulations, e.g., $\mathscr{D} = \{(\mathbf{x}_i, n_i), i = 1, 2, \ldots, k\}$. Denote the sample average simulation response at $\mathbf{x}_i$ by $\bar{\mathscr{Y}}(\mathbf{x}_i)$ and let the vector of the sample average simulation responses be $\bar{\mathscr{Y}} = \left(\bar{\mathscr{Y}}(\mathbf{x}_1), \bar{\mathscr{Y}}(\mathbf{x}_2), \ldots, \bar{\mathscr{Y}}(\mathbf{x}_k)\right)^\top$. Following the aforementioned description, SK represents the sample average simulation response by

$$\bar{\mathscr{Y}}(\mathbf{x}_i) = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathscr{Y}_j(\mathbf{x}_i) = \mathsf{Y}(\mathbf{x}_i) + \bar{\varepsilon}(\mathbf{x}_i), \tag{3}$$

where $\bar{\varepsilon}(\mathbf{x}_i) := n_i^{-1} \sum_{j=1}^{n_i} \varepsilon_j(\mathbf{x}_i)$ denotes the simulation error associated with the performance-measure point estimate $\bar{\mathscr{Y}}(\mathbf{x}_i)$, $i = 1, 2, \ldots, k$. The simulation errors $\varepsilon_1(\mathbf{x}_i), \varepsilon_2(\mathbf{x}_i), \ldots$ are naturally independent and identically distributed (i.i.d.) across replications in the stochastic simulation setting. In this paper we do not consider using common random numbers (CRN is a widely used vari-

Please cite this article as: X. Chen, Q. Zhou, Sequential design strategies for mean response surface metamodeling via stochastic kriging with adaptive exploration and exploitation, European Journal of Operational Research (2017), http://dx.doi.org/10.1016/j.ejor.2017.03.042

Download English Version:

https://daneshyari.com/en/article/4959513

Download Persian Version:

https://daneshyari.com/article/4959513

Daneshyari.com