Computers and Structures 126 (2013) 86–99

Contents lists available at SciVerse ScienceDirect

Computers and Structures

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

Non-parametric stochastic subset optimization for optimal-reliability design problems

Gaofeng Jia, Alexandros A. Taflanidis *

University of Notre Dame, Department of Civil and Environmental Engineering and Earth Sciences, Notre Dame, IN 46556, USA

ARTICLE INFO

Article history:
Received 26 May 2012
Accepted 7 December 2012
Available online 21 January 2013

Keywords:
Stochastic subset optimization
Reliability-based optimization
Kernel density estimation
Stochastic simulation

ABSTRACT

The stochastic subset optimization (SSO) algorithm has been recently proposed for design problems that use the system reliability as objective function. It is based on simulation of samples of the design variables from an auxiliary probability density function, and uses this information to identify subsets for the optimal solution. This paper presents an extension, termed Non-Parametric SSO, that adopts kernel density estimation (KDE) to approximate the objective function through these samples. It then uses this approximation to identify candidate points for the global minimum. To reduce the computational effort an iterative approach is established whereas efficient reflection methodologies are implemented for the KDE.

1. Introduction

Reliability-based Optimization (RBO) constitutes a powerful framework for providing robust optimal solutions for engineering design problems by explicitly addressing the effects of the modeling uncertainties stemming from our incomplete knowledge about the system state or future excitation events [1–4]. This is established by incorporating in the objective function or in the design constraints reliability measures, typically expressed through the probability of failure [5]. The latter provides a measure of the plausibility of the occurrence of unacceptable behavior of the system (frequently termed as "failure"), and is ultimately expressed through a multi-dimensional integral over the space of uncertain model parameters where the integrand is the product of the failure probability conditional on the model parameters and the probability of these parameters.

Although RBO problems are commonly formulated by adopting deterministic objective functions and reliability constraints (e.g., [6–8]), they are also frequently defined with the system reliability as the objective function (e.g., [3,9,10]). The focus of this study is on applications of the latter type. For such problems, especially ones including complex system or excitation models, this reliability cannot be always evaluated, or efficiently approximated, analytically in the context of RBO. An alternative approach [2,11] is to estimate it using stochastic simulation techniques [12], for example Monte Carlo simulation. This approach offers a high accuracy estimation of the system reliability, which makes it appropriate for RBO, but involves, though, an unavoidable estimation error and significant computational burden, since a large number of evaluations of the system-model response are required for each evaluation of the objective function. These features make the associated design optimization a challenging task [2,13].

The Stochastic Subset Optimization (SSO) algorithm was proposed in [14] for such RBO problems that adopt the system reliability as objective function and include computationally expensive numerical models. The algorithm was later extended to general optimization under uncertainty problems [15] and also efficiently integrated with sensitivity analysis tools for the model parameters in [16] and with surrogate modeling approaches in [17]. SSO is based on simulation of samples of the design variables from an auxiliary probability density function (treating them as uncertain) and uses this information to efficiently identify a subset for the optimal design variables within some predefined class of admissible subsets. The latter identification requires an optimization, with respect to some parametric description of the admissible subset class, for the one that has the smallest volume density of samples. Even though SSO has been proven efficient for various challenging optimization problems (e.g., [1,18–20]), it does have two vulnerabilities: the identification of the optimal subset involves a challenging non-smooth optimization problem [16], whereas the subset identified is the one that has the smallest average value for the objective function (among the admissible subsets) which does not always guarantee that it includes the minimum of the objective function [16].

In this paper an extension of SSO is developed to improve upon these vulnerabilities. This is established by avoiding the parametric description of subsets or the search for the one that has the smallest average value. This new version of the algorithm is termed NP-SSO (non-parametric SSO). The fundamental difference in NP-SSO

* Corresponding author. Tel.: +1 574 631 5696.
E-mail address: a.taflanidis@nd.edu (A.A. Taflanidis).

0045-7949/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.compstruc.2012.12.009

Nomenclature

RBO	reliability based optimization
$\mathbf{x}$	design variable vector
x_i	ith design variable
n_x	dimension of $\mathbf{x}$
X	admissible design space
$I_F(\mathbf{x},\boldsymbol{\theta})$	indicator function for F (failure)
$P_F(\mathbf{x})$	objective function (failure probability)
I	box-bounded search space
$P(F_I)$	probability of failure of augmented reliability problem
$p(\mathbf{x}\vert F_I)$	auxiliary (failure) density function for $\mathbf{x}$
$p(\mathbf{x},\boldsymbol{\theta}\vert F_I)$	auxiliary (failure) density function for $\mathbf{x}$ and $\boldsymbol{\theta}$
$\{\mathbf{x}^j_{..},\boldsymbol{\theta}^j_{..}\}$	jth sample from $\mathbf{x}_{..}$ or $\boldsymbol{\theta}_{..}$
$\{\mathbf{x}_F,\boldsymbol{\theta}_F\}$	(failure) sample set for $\{\mathbf{x},\boldsymbol{\theta}\}$ from $p(\mathbf{x},\boldsymbol{\theta}\vert F_I)$
$\{\mathbf{x}_c,\boldsymbol{\theta}_c\}$	sample set for which the system response is evaluated during the stochastic sampling process
$K(.)$	kernel
σ_i	standard deviation of samples $\{\mathbf{x}_F\}$ for ith design variable
I^*	subset of I with $\widetilde{P}_F(\mathbf{x}) \leqslant c^F$
$\delta(I^*\vert I)$	volume ratio between sets I^* and I
$\widehat{H}(I^*\vert I)$	approximation to $H(I^*\vert I)$ using stochastic simulation
$\{\mathbf{x}_u\}$	uniform samples in I
n_u	number of samples in $\{\mathbf{x}_u\}$
$\mathbf{X}_c$	initial coordinate system
$\mathbf{T}$	transformation between $\mathbf{X}_c$ and $\underline{\mathbf{X}}_c$
$\underline{\mathbf{x}}$	coordinates of $\mathbf{x}$ in $\underline{\mathbf{X}}_c$
ρ	target volume ratio for $\delta(I^*\vert I)$ per iteration of NP-SSO
$..^m$ (superscript m)	mth cluster of $\{\mathbf{x}_o\}$ characteristics
$.._k$ (subscript k)	kth iteration of NP-SSO characteristics
KDE	kernel density estimation
$\mathbf{x}^*$	optimal design solution
$\boldsymbol{\theta}$	model parameters
n_θ	dimension of $\boldsymbol{\theta}$
$\boldsymbol{\Theta}$	space of possible values for $\boldsymbol{\theta}$
$p(.)$	probability density function
$\widehat{P}_F(\mathbf{x})$	approximation to objective function through stochastic simulation
V_I	volume of set I
$\widehat{P}(F_I)$	approximation to $P(F_I)$ through stochastic simulation
$\widetilde{P}_F(\mathbf{x})$	approximation to objective function through NP-SSO
$p(\boldsymbol{\theta}\vert F_I)$	auxiliary (failure) density function for $\boldsymbol{\theta}$
n_s	number of samples available from $p(\mathbf{x}\vert F_I)$
$\{\mathbf{x}_F\}$	(failure) sample set from $p(\mathbf{x}\vert F_I)$
N	number of samples in set $\{\mathbf{x}_c,\boldsymbol{\theta}_c\}$
$q(.)$	proposal density for $\boldsymbol{\theta}$ for obtaining samples $\{\boldsymbol{\theta}_c\}$
h_i	bandwidth of Kernel for x_i
$\widetilde{I}^*$	box-bounded superset of I^*
$H(I^*\vert I)$	ratio of average objective function values in I^* and I
$\widetilde{H}(I^*\vert I)$	approximation to $H(I^*\vert I)$ using approximation $\widetilde{P}_F(\mathbf{x})$
$\{\mathbf{x}_o\}$	uniform samples in I^*
n_o	number of samples in $\{\mathbf{x}_o\}$
$\underline{\mathbf{X}}_c$	rotated coordinate system
$\mathbf{C}$	covariance matrix of $\{\mathbf{x}_o\}$
c^F	threshold defining subset I^*
$\tilde{\rho}$	maximum acceptable volume ratio for $\delta(\widetilde{I}^*\vert I)$ per iteration of NP-SSO
M	number of clusters for $\{\mathbf{x}_o\}$
c^S	stopping threshold for $H(I^*\vert I)$

is that Kernel Density Estimation (KDE) is adopted to approximate the objective function using the information from the available samples of the design variables, and to ultimately identify candidate points (not subsets) for the global minimum. As it can be challenging to populate the failure region with samples, especially for design choices in regions close to the minima of the objective function, an iterative approach is established to reduce the computational effort. Through this approach the design variable samples gradually move from regions with higher values of the objective function to regions with lower values, to ultimately establish a higher accuracy KDE approximation in these regions of interest. Appropriate reflection approaches are introduced for the KDE and the impact of potential multiple local minima on the computational framework is also addressed. The comparison to SSO, as well as challenges for the KDE implementation in high-dimensional problem are also extensively discussed. In the next section the RBO problem of interest is reviewed. In Section 3 the general theoretical and computational framework for the NP-SSO algorithm is presented and then in Section 4 the proposed iterative approach to increase computational efficiency is discussed. Then the NP-SSO algorithm is reviewed in Section 5 and in Section 6 it is illustrated in an example considering the optimization of a base-isolation protective system for a three story structure.

2. Reliability based optimization problem

Consider a system that involves some controllable parameters that define its design, referred to as design variables and let $\mathbf{x} = [x_1 x_2 \cdots x_{n_x}] \in X \subset \Re^{n_x}$ be the design vector where X denotes the bounded admissible design space. Let $\boldsymbol{\theta} = [\theta_1 \theta_2 \cdots \theta_{n_\theta}]$ lying in $\Theta \subset \Re^{n_\theta}$ be the vector of uncertain model parameters for the system, where Θ denotes the set of their possible values. A PDF (probability density function) $p(\boldsymbol{\theta})$, which incorporates our available knowledge about the system and its excitation, is assigned to these parameters. This PDF is assumed here independent of $\mathbf{x}$, though the approach can be directly extended [2] to cases where the latter is not true [simply substitute $p(\boldsymbol{\theta}|\mathbf{x})$ for $p(\boldsymbol{\theta})$]. The failure probability for a given choice of the design variables is then [5]:

$$P_F(\mathbf{x}) = \int_\Theta I_F(\mathbf{x},\boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta} \quad (1)$$

where $I_F(\mathbf{x},\boldsymbol{\theta})$ is the indicator function for event F (frequently characterized as system failure), which equals one if event F occurs for the system that corresponds to $\{\mathbf{x},\boldsymbol{\theta}\}$ and zero if it does not. Commonly a limit state function $g_l(\mathbf{x},\boldsymbol{\theta})$ is introduced to describe F through some appropriate convention, for example, $g_l(\mathbf{x},\boldsymbol{\theta}) > 0 \Leftrightarrow F$. Ultimately calculation of $I_F(\mathbf{x},\boldsymbol{\theta})$ requires evaluation of the system response to estimate $g_l(\mathbf{x},\boldsymbol{\theta})$.

As discussed in the introduction, we are interested here in the minimization problem:

$$\begin{aligned} &\min \ P_F(\mathbf{x}) \\ &\text{given } \mathbf{f}_c(\mathbf{x}) \geqslant 0 \end{aligned} \quad (2)$$

where $\mathbf{f}_c(\mathbf{x})$ is a vector of deterministic constraints. An equivalent formulation of this problem is

$$x^* = \arg\min_{\mathbf{x}\in X} P_F(\mathbf{x}) \quad (3)$$

where the constraints are taken into account by appropriate definition of the admissible design space X.

Using stochastic simulation the objective function for (3) can be calculated as

$$\widehat{P}_F(\mathbf{x}) = \frac{1}{n_t}\sum_{j=1}^{n_t} I_F(\mathbf{x},\boldsymbol{\theta}^j)\frac{p(\boldsymbol{\theta}^j)}{q(\boldsymbol{\theta}^j)} \quad (4)$$