Reliability Engineering and System Safety 156 (2016) 64–83

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

Global sensitivity analysis using low-rank tensor approximations

Katerina Konakli*, Bruno Sudret

ETH Zürich, Chair of Risk, Safety and Uncertainty Quantification, Stefano-Franscini-Platz 5, 8093 Zürich, Switzerland

ARTICLE INFO

Article history:
Received 17 December 2015
Received in revised form
26 April 2016
Accepted 9 July 2016
Available online 16 July 2016

Keywords:
Global sensitivity analysis
Sobol' indices
Low-rank approximations
Polynomial chaos expansions

ABSTRACT

In the context of global sensitivity analysis, the Sobol' indices constitute a powerful tool for assessing the relative significance of the uncertain input parameters of a model. We herein introduce a novel approach for evaluating these indices at low computational cost, by post-processing the coefficients of polynomial meta-models belonging to the class of low-rank tensor approximations. Meta-models of this class can be particularly efficient in representing responses of high-dimensional models, because the number of unknowns in their general functional form grows only linearly with the input dimension. The proposed approach is validated in example applications, where the Sobol' indices derived from the meta-model coefficients are compared to reference indices, the latter obtained by exact analytical solutions or Monte-Carlo simulation with extremely large samples. Moreover, low-rank tensor approximations are confronted to the popular polynomial chaos expansion meta-models in case studies that involve analytical rank-one functions and finite-element models pertinent to structural mechanics and heat conduction. In the examined applications, indices based on the novel approach tend to converge faster to the reference solution with increasing size of the experimental design used to build the meta-model.

1. Introduction

Robust predictions via computer simulation necessitate accounting for the prevailing uncertainties in the parameters of the computational model. Uncertainty quantification provides the mathematically rigorous framework for propagating the uncertainties surrounding the model input to a response quantity of interest. It comprises three fundamental steps [1,2]: First, the model representing the physical system under consideration is defined; the model maps a given set of input parameters to a unique value of the response quantity of interest. The second step involves the probabilistic description of the input parameters by incorporating available data, expert judgment or a combination of both. In the third step, the uncertainty in the input is propagated upon the response quantity of interest through repeated evaluations of the model for appropriate combinations of the input parameters. In cases when the uncertainty in the response proves excessive or when it is of interest to reduce the dimensionality of the model, sensitivity analysis may be employed to rank the input parameters with respect to their significance for the response variability. Accordingly, important parameters may be described in further detail, whereas unimportant ones may be fixed to nominal values.

Methods of sensitivity analysis can be generally classified into local and global methods. Local methods are limited to examining effects of variations of the input parameters in the vicinity of nominal values. Global methods provide more complete information by accounting for variations of the input parameters in their entire domain. Under the simplifying assumption of linear or nearly linear behavior of the model, global sensitivity measures can be computed by fitting a linear-regression model to a set of input samples and the respective responses (see *e.g.* [3,4] for definitions of such measures). The same methods can be employed in cases with models that behave nonlinearly but monotonically, after applying a rank transformation on the available data. Variance-based methods represent a more powerful and versatile approach, also applicable to nonlinear and non-monotonic models. These methods, known as functional ANOVA (denoting ANalysis Of VAriance) techniques in statistics, rely upon the decomposition of the response variance as a sum of contributions from each input parameter or their combinations [5]. The Sobol' indices, originally proposed in [6], constitute the most popular tool thereof. Although these indices have proven powerful in a wide range of applications, their definition is ambiguous in cases with dependent input variables, which has led to different extensions of the original framework [7–11]. An alternative perspective is offered by the distribution-based indices, which are well-defined regardless of the dependence structure of the input [12–16]. The key idea is to use an appropriate distance measure to evaluate the effect of suppressing the uncertainty in selected variables on the probability density function (PDF) or the cumulative distribution function (CDF) of the response. These indices are especially useful

* Corresponding author.
E-mail address: konakli@ibk.baug.ethz.ch (K. Konakli).

http://dx.doi.org/10.1016/j.ress.2016.07.012
0951-8320/

when consideration of the variance only is not deemed sufficient to characterize the response uncertainty. However, contrary to the Sobol' indices, they do not sum up to unity, which may hinder interpretation. For further information on global sensitivity analysis methods, the interested reader is referred to the review papers [17,3,4,18].

It should be mentioned that different classifications of sensitivity analysis techniques can be found in the literature. In cases when one needs to perform a fast exploration of the model behavior with respect to a possibly large number of uncertain input parameters, the so-called screening methods may be employed. These methods can provide a preliminary ranking of the importance of the various input parameters at low computational cost before more precise and costly methods are applied. The Cotter method [19] and the Morris method [20] are widely used screening methods, with the latter covering the input space in a more exhaustive manner. The more recently proposed derivative-based global sensitivity measures can also be classified into this category, while they also provide upper bounds for the Sobol' indices [21–23].

The focus of the present paper is on sensitivity analysis by means of Sobol' indices. We limit our attention to cases with independent input and address the computation of these indices for high-dimensional expensive-to-evaluate models, which are increasingly used across engineering and sciences. Various methods have been investigated for computing the Sobol' indices based on Monte Carlo simulation [24–28]; because of the large number of model evaluations required, these methods are not affordable for computationally costly models. To overcome this limitation, more efficient estimators have recently been proposed [29–32]. A different approach is to substitute a complex model by a *meta-model*, which has similar statistical properties while maintaining a simple functional form (see *e.g.* [33–38] for global sensitivity analysis with various types of meta-models). Sudret [39] proposed to compute the Sobol' indices by post-processing the coefficients of polynomial chaos expansion (PCE) meta-models. The key concept in PCE is to expand the model response onto a basis made of orthogonal multivariate polynomials in the input variables. The computational cost of the associated Sobol' indices essentially reduces to the cost of estimating the PCE coefficients, which can be curtailed by using sparse PCE [40]. The PCE-based approach for computing the Sobol' indices is employed by a growing number of researchers in various fields including hydrogeology [41–43], geotechnics [44], ocean engineering [45], biomedical engineering [46], hybrid dynamic simulation [47] and electromagnetism [48,49]. Unfortunately, the PCE approach faces the so-called *curse of dimensionality*, meaning the exploding size of the candidate basis with increasing dimension.

The goal of this paper is to derive a novel approach for solving global sensitivity analysis problems in high dimensions. To this end, we make use of a recently emerged technique for building meta-models with polynomial functions based on low-rank approximations (LRA) [50–55]. LRA express the model response as a sum of a small number or rank-one tensors, which are products of univariate functions in each of the input parameters. Because the number of unknown coefficients in LRA grows only linearly with the input dimension, this technique is particularly promising for dealing with cases of high dimensionality. We herein derive analytical expressions for the Sobol' sensitivity indices based on the general functional form of LRA with polynomial bases. As in the case of PCE, the computational cost of the LRA-based Sobol' indices reduces to the cost of estimating the coefficients of the meta-model. Once a LRA meta-model of the response quantity of interest is available, the Sobol' indices can be calculated with elementary operations at nearly zero additional computational cost.

The paper is organized as follows: In Section 2, we review the basic concepts of Sobol' sensitivity analysis and define the corresponding sensitivity indices. In Section 3, we describe the mathematical setup of non-intrusive meta-modeling and define error measures that characterize the meta-model accuracy. After reviewing the computation of Sobol' indices using PCE meta-models in Section 4, we introduce the LRA-based approach in Section 5. In this, we first detail a greedy algorithm for the construction of LRA in a non-intrusive manner and then, use their general functional form to derive analytical expressions for the Sobol' indices. In Section 6, we demonstrate the accuracy of the proposed method by comparing the LRA-based indices to reference ones, with the latter representing the exact solution or Monte-Carlo estimates relying on a large sample of responses of the actual model. Furthermore, we examine the comparative performance of the LRA- and PCE-based approaches in example applications that involve analytical rank-one functions and finite-element models pertinent to structural mechanics and heat conduction. The main findings are summarized in Section 7.

2. Sobol' sensitivity analysis

We consider a computational model $\mathcal{M}$ describing the behavior of a physical or engineered system of interest. Let $\boldsymbol{X} = \{X_1, \dots, X_M\}$ denote the M-dimensional input random vector of the model with prescribed joint PDF $f_{\boldsymbol{X}}$. Due to the input uncertainties embodied in $\boldsymbol{X}$, the model response becomes random. By limiting our focus to a scalar response quantity Y, the computational model represents the map:

$$\boldsymbol{X} \in \mathcal{D}_{\boldsymbol{X}} \subset \mathbb{R}^M \longmapsto Y = \mathcal{M}(\boldsymbol{X}) \in \mathbb{R}, \tag{1}$$

where $\mathcal{D}_{\boldsymbol{X}}$ denotes the support of $\boldsymbol{X}$.

Sobol' sensitivity analysis aims at apportioning the uncertainty in Y, described by its variance, to contributions arising from the uncertainty in individual input variables and their interactions. As explained in the Introduction, the theoretical framework described in the sequel is confined to the case when the components of $\boldsymbol{X}$ are independent. Under this assumption, the joint PDF $f_{\boldsymbol{X}}$ is the product of the marginal PDF f_{X_i} of each input parameter.

2.1. Sobol' decomposition

Assuming that $\mathcal{M}$ is a square-integrable function with respect to the probability measure associated with $f_{\boldsymbol{X}}$, its Sobol' decomposition in summands of increasing dimension is given by [6]:

$$\mathcal{M}(\boldsymbol{X}) = \mathcal{M}_0 + \sum_{i=1}^{M} \mathcal{M}_i(X_i) + \sum_{1 \le i < j \le M} \mathcal{M}_{i,j}(X_i, X_j) + \dots + \mathcal{M}_{1,2,\dots,M}(\boldsymbol{X}) = \mathcal{M}_0 + \sum_{\substack{\mathbf{u} \subset \{1,\dots,M\} \\ \mathbf{u} \neq \emptyset}} \mathcal{M}_{\mathbf{u}}(\boldsymbol{X}_{\mathbf{u}}), \tag{2}$$

where $\mathbf{u} = \{i_1, \dots, i_s\}$, $1 \le s \le M$, denotes a subset of $\{1, \dots, M\}$ and $\boldsymbol{X}_{\mathbf{u}} = \{X_{i_1}, \dots, X_{i_s}\}$ is the subvector of $\boldsymbol{X}$ containing the variables of which the indices comprise $\mathbf{u}$.

The uniqueness of the decomposition is ensured by choosing summands that satisfy the conditions:

$$\mathcal{M}_0 = \mathbb{E}\left[\mathcal{M}(\boldsymbol{X})\right] \tag{3}$$

and

$$\mathbb{E}\left[\mathcal{M}_{\mathbf{u}}(\boldsymbol{X}_{\mathbf{u}})\mathcal{M}_{\mathbf{v}}(\boldsymbol{X}_{\mathbf{v}})\right] = 0 \quad \forall\, \mathbf{u}, \mathbf{v} \subset \{1, \dots, M\}, \quad \mathbf{u} \neq \mathbf{v}. \tag{4}$$

Note that the above condition implies that all summands $\{\mathcal{M}_{\mathbf{u}}, \mathbf{u} \neq \emptyset\}$ in Eq. (2) have zero mean values. A recursive construction of summands satisfying the above conditions is obtained as: