Applied Mathematical Modelling 37 (2013) 6097–6109

Contents lists available at SciVerse ScienceDirect

Applied Mathematical Modelling

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

Moving least squares based sensitivity analysis for models with dependent variables

Longfei Tian, Zhenzhou Lu *, Wenrui Hao

School of Aeronautics, Northwestern Polytechnical University, Xi'an, Shaanxi 710072, PR China

ARTICLE INFO

Article history:
Received 21 March 2012
Received in revised form 9 November 2012
Accepted 18 December 2012
Available online 29 December 2012

Keywords:
Sensitivity analysis
Dependent inputs
Variance decomposition
Moving least squares
Marginal contribution
Total contribution

ABSTRACT

For models with dependent input variables, sensitivity analysis is often a troublesome work and only a few methods are available. Mara and Tarantola in their paper ("*Variance-based sensitivity indices for models with dependent inputs*") defined a set of variance-based sensitivity indices for models with dependent inputs. We in this paper propose a method based on moving least squares approximation to calculate these sensitivity indices. The new proposed method is adaptable to both linear and nonlinear models since the moving least squares approximation can capture severe change in scattered data. Both linear and nonlinear numerical examples are employed in this paper to demonstrate the ability of the proposed method. Then the new sensitivity analysis method is applied to a cantilever beam structure and from the results the most efficient method that can decrease the variance of model output can be determined, and the efficiency is demonstrated by exploring the dependence of output variance on the variation coefficients of input variables. At last, we apply the new method to a headless rivet model and the sensitivity indices of all inputs are calculated, and some significant conclusions are obtained from the results.

1. Introduction

In many risk assessment models, the uncertainty in input variables is one of the biggest challenges for researchers to deal with [1]. In order to look at the influence of input uncertainty on the model output, sensitivity analysis (SA) is a powerful technique and is now widely used in practical engineering problems. The most frequently used definition of sensitivity analysis is proposed by Saltelli [2] in 2002 as following: SA is *"The study of how uncertainty in the output of a model* (*numerical or otherwise*) *can be apportioned to different sources of uncertainty in the model input"*. Also, according to McRae et al. [3], SA can be classified into two groups: the local sensitivity analysis (LSA) and the global sensitivity analysis (GSA). LSA researches the local response variation of the model output with the variation of input variables one at a time by fixing other variables at their central values. While for GSA, its objective is to rank the importance of the system inputs considering their uncertainty and the influence they have upon the uncertainty of the system output, typically over a large region of input space [4]. Apparently, GSA has a more extensive use in engineering problems.

There is considerable amount of literature about GSA available now. In 1990, Iman and Hora [5] advocated that ideal sensitivity indices, also called importance measures (IM), should be "*unconditional, easy to interpret, easy to compute, and stable*", and based on which they proposed the Iman–Hora IM. Similar to Iman and Hora, Saltelli [6] later proposed his new requirements for SA techniques, i.e., "*global, quantitative and model free*", also he proposed his own sensitivity indicators that satisfy the three new requirements. Many other people such as Pörn [7], Rabitz et al. [8,9], Chun et al. [10], Sobol [11,12], Helton

* Corresponding author. Tel./fax: +86 29 88460480.
E-mail addresses: tianlf44@163.com (L. Tian), zhenzhoulu@nwpu.edu.cn (Z. Lu), haowr22@163.com (W. Hao).

0307-904X/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.apm.2012.12.010

et al. [13], Borgonovo [14] and De Rocquigny et al. [15] have also proposed their own sensitivity indices and the corresponding GSA techniques. Among all these GSA techniques, variance-based methods are the most familiar ones and they are still under active development because of their demonstrated merits [16]: firstly, their capacity to capture the influence of the full range of variation of each input factor and secondly, their ability to identify interaction effects among input factors. With this respect, we also adopt the variance-based sensitivity indices in this paper.

In general, the estimation of the variance-based sensitivity indices can be classified into two main methodological approaches [17]: the first one is the classical approach such as the crude Monte Carlo method [16], Fourier Amplitude Sensitivity Test (FAST) method [18] and the extended FAST [19] method. These 'classical' approaches estimate directly the conditional variances characterizing the sensitivity indices. Hence, these 'classical' methods have a significant computational cost, which renders impractical their application to expensive computational models [17]. The second one is the metamodelling approach, and its application in GSA has been demonstrated in [20]. By using metamodelling approach, a 'metamodel' of the original real model (mathematical or computational) is estimated first, then on the basis of this 'metamodel', we can obtain all sensitivity indices of interest, as well as the factor mapping between model inputs and model output. Intuitively, the metamodelling approach has a higher efficiency in estimating the conditional variances characterizing the sensitivity indices because it can make a better use of the information contained in the sample [17].

For many variance-based GSA methods based on metamodelling, they use high dimensional model representation (HDMR, Ref. [8]) to decompose the original model function into terms of increasing dimensions (including main effects and interactions). For models with independent input variables, Sobol' [21] proved that if each term in the HDMR expansion has zero mean, then all the terms of the decomposition are orthogonal in pairs and they can be univocally determined. According to the orthogonal HDMR expansion proposed by Sobol', we can also decompose the variance of model output into terms of increasing dimensionality, and this is the so-called ANOVA-HDMR (ANOVA means "*analysis of variance*") decomposition. At present, many GSA methods using HDMR expansion are available in literature, and most of them rely on the assumption that all input variables are independent. For models with dependent input variables, although variable dependence is a common situation in practical engineering problems, only a few studies have been conducted to research models with dependent inputs. Saltelli et al. [22] proposed a correlation ratio method based on McKay's method [23]. Fang et al. [24] proposed sequential sampling to approximate the differential sensitivity index. However, these methods can only derive an overall sensitivity index of one input variable, this is not convenient for engineering decision making. Xu and Gertner [25] later advocated that the uncertainty contribution of an individual input variable on the model output should be decomposed into two parts: the uncorrelated contribution and the correlated contribution. The uncorrelated one means that this part of contribution has nothing to do with other input variables, which can only be explained by the input variable itself. The correlated one means that the contribution must take into consideration the correlation between the input variable and all other input variables. One drawback of Xu and Gertner's method is that their method is based on linear regression; hence a limitation of the technique nature appears when one conducts the method on models with appreciable nonlinearity. Another drawback of their method is that their decomposition is not thorough enough. Similar to Xu and Gertner, Mara and Tarantola [26] also defined a set of variance-based sensitivity indices for models with dependent variables by performing the ANOVA-HDMR after *decorrelating* the inputs. Compared with the uncertainty decomposition proposed by Xu and Gertner in Ref. [25], Mara and Tarantola's decomposition is univocal and easier to interpret. They show that the variance-based sensitivity indices can be interpreted as the fully, partially correlated and independent contributions of the inputs to the output variance.

In this paper, we propose a new GSA method using moving least-squares (MLS, Refs. [27–29]) metamodelling to calculate the variance-based sensitivity indices defined by Mara and Tarantola. Due to the strong flexibility of the moving least-squares approximation in nonlinear models, this new method has the ability to deal with both linear and nonlinear problems. In addition, this method using MLS normally calculates all the sensitivity indices (including all the main effects and all the interaction effects up to the second or third order) with only a few thousand samples ($\leqslant 5000$), almost independent of the dimensionality of the model function.

The rest of the paper is organized as follows: in Section 2 we briefly introduce the Sobol's ANOVA-HDMR decomposition and the general variance-based sensitivity indices for models with independent input variables. In Section 3, we present the Mara and Tarantola's variance-based sensitivity indices for models with dependent input variables. In Section 4, we describe the proposed GSA method using moving least-squares metamodelling in detail and show how this method can be adopted in calculating the sensitivity indices defined by Mara and Tarantola for models with dependent inputs. In Section 5 both numerical and engineering examples are employed to demonstrate the accuracy and efficiency of the proposed GSA method. Conclusions are provided in Section 6.

2. The ANOVA–HDMR and sensitivity indices for models with independent inputs

We consider a general model function described as follows,

$$Y = f(X_1, X_2, \cdots X_n) = f(\mathbf{X}) \tag{1}$$

where $\mathbf{X} = (X_1, X_2, \ldots X_n)$ in this section is the vector of independent random input variables and Y is the output, and we assume that the original model function $f(\mathbf{X})$ is square integrable over the input space. Then the HDMR expansion of $f(\mathbf{X})$ which is closely related to variance-base GSA methods can be described as follows [16]: