Probabilistic Engineering Mechanics 39 (2015) 46–55

Contents lists available at ScienceDirect

Probabilistic Engineering Mechanics

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

Sparse grid integration based solutions for moment-independent importance measures

Changcong Zhou [a], Zhenzhou Lu [b,*], Wei Li [b]

[a] School of Mechanics, Civil Engineering and Architecture, Northwestern Polytechnical University, Xi'an, China
[b] School of Aeronautics, Northwestern Polytechnical University, Xi'an, China

ARTICLE INFO

Article history:
Received 21 July 2013
Received in revised form
16 December 2014
Accepted 16 December 2014
Available online 18 December 2014

Keywords:
Uncertainty
Sensitivity analysis
Importance measure
Moment-independent
Sparse grid

ABSTRACT

Different importance measures exist in the literature, aiming to quantify the contributions of model inputs to the output uncertainty. Among them, the moment-independent importance measures consider the entire model output without dependence on any of its particular moments (e.g., variance), and have attracted growing attention among both academics and practitioners. However, until now robust and efficient computational methods for moment-independent importance measures are unavailable in the literature. To properly address this issue, a new computational method based on sparse grid integration (SGI) is proposed to perform moment-independent importance analysis in an efficient framework. In the proposed method, SGI is combined with the moment method to obtain the conditional and unconditional distributions of the model output, which are further used to estimate the moment-independent importance measures. The proposed method makes full use of the advantages of the SGI technique, and is easy to implement. Numerical and engineering examples are studied in this work and the results have demonstrated that the proposed method is feasible and applicable for practical use.

1. Introduction

As the complexity of mathematical models in the engineering keeps increasing, it is necessary to take all the inputs into consideration thoroughly and find which of them are more important to the model output [1,2]. Global sensitivity analysis is such a tool to study "*how uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model input factors*" [3]. Global sensitivity indices are also known as importance measures, which generally include non-parametric techniques [4], variance based methods studied by Sobol [5], Rabitz and Alis [6], Saltelli et al. [7,8], Frey and Patil [9], and moment-independent approaches proposed by Borgonovo [10–13], Liu and Homma [14]. Saltelli [3] underlined that an importance measure should satisfy the requirement of being "*global, quantitative, and model free*". Borgonovo [10] extended Saltelli's three requirements by adding "*moment independent*", and proposed a new importance measure which looks at the influence of input uncertainty on the entire output distribution without reference to a specific moment of the output. In a similar manner, Liu and Homma [14] proposed another moment-independent importance measure, which describes the contribution of the input uncertainty on the cumulative distribution function (CDF) of the model output, while Borgonovo's importance measure studies the contribution on the probability density function (PDF) of the model output.

A key concern in performing importance analysis is to improve the computational efficiency [15], i.e. to obtain results with a modest number of model evaluations, especially in engineering cases where the models involved usually take a long processing time. Numerical simulation methods have been adopted by Borgonovo [10], Liu and Homma [14,16] to obtain the two moment-independent importance measures proposed by them respectively, of which the computational burden is considerably large as double loop samplings are involved. In this respect, more efficient algorithms should be developed. According to the work by Liu and Homma [16], the calculation of the two importance measures can be, as a matter of fact, transformed into calculations of CDFs. Simply speaking, once the unconditional and conditional CDFs of the model output are obtained, both the two moment-independent importance measures can be readily obtained.

In this work, we develop a new method to obtain the two moment-independent importance measures by using the sparse grid integration (SGI) technique. The sparse grid technique [17], which is based on one-dimensional formulas and then extended to higher dimensions, has been extensively utilized for high dimensional multivariate integration [18], as well as interpolation [19]. To a certain extent, this technique avoids the "curse of dimension"

* Corresponding author.
E-mail address: zhenzhoulu@nwpu.edu.cn (Z. Lu).

http://dx.doi.org/10.1016/j.probengmech.2014.12.002
0266-8920/© 2014 Elsevier Ltd. All rights reserved.

of conventional integration algorithms meaning that the computational cost grows exponentially with the dimension of the problem. In the SGI technique, multivariate quadrature formulas are constructed with combinations of tensor products of suitable one-dimensional formulas, thus the function evaluations needed and the numerical accuracy become independent of the dimension of the problem up to logarithmic factors. Existing literature has reported the high efficiency and accuracy of sparse grid applications. The SGI technique is combined with moment methods [26–28] in the proposed framework, where the first four moments of functions will be estimated by SGI and further used to obtain the unconditional and conditional CDFs of the model output. The obtained CDFs will be finally utilized to calculate the importance measures.

The remainder of the paper is organized as follows. Section 2 reviews the definitions of the two moment-independent importance measures. This is followed by a brief introduction of the mathematical formulas of the SGI technique in Section 3. In Section 4, we propose a new method based on SGI to obtain the moment-independent importance measures more efficiently. In Section 5, numerical and engineering examples are presented to demonstrate the feasibility of the proposed method. Finally, conclusions of this work are highlighted in Section 6.

2. Two moment-independent importance measures

2.1. Definitions

Suppose the model is denoted by $Y = g(\boldsymbol{X})$, where Y is the model output, and $\boldsymbol{X} = [X_1, X_2, \ldots, X_d]^T$ (d is the number of inputs) is the set of uncertain inputs. The uncertainties of the inputs are represented by probability distributions. The unconditional PDF and CDF of Y are denoted as $f_Y(y)$ and $F_Y(y)$, respectively. The conditional PDF and CDF of Y are denoted as $f_{Y|X_i}(y)$ and $F_{Y|X_i}(y)$, which are obtained by fixing the input of interest X_i at its realization x_i^* generated from its distribution. The conditional and unconditional distribution functions of the output are depicted in Fig. 1.

Borgonovo [10] proposed his importance measure by considering the contribution of inputs to the PDF of the model output. The shift between $f_Y(y)$ and $f_{Y|X_i}(y)$ is measured by the area $s(X_i)$ closed by the two PDFs, which is given as follows:

$$s(X_i) = \int |f_Y(y) - f_{Y|X_i}(y)|\, dy \quad (1)$$

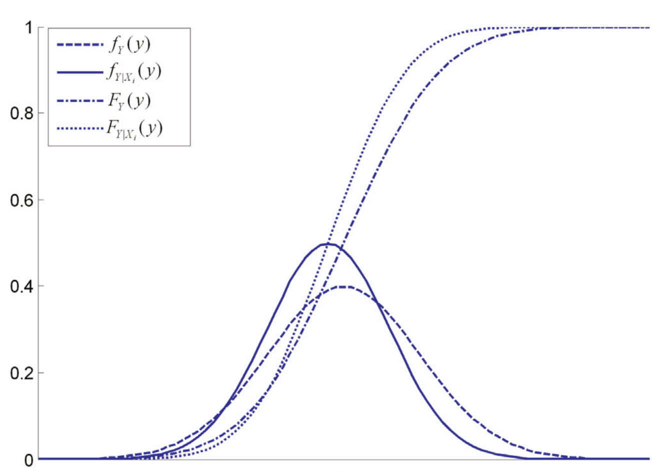

Fig. 1. Unconditional and conditional distribution functions of the model output.

The expected shift can be obtained by varying the value of X_i over its distribution range, i.e.

$$E_{X_i}[s(X_i)] = \int f_{X_i}(x_i) s(X_i)\, dx_i \quad (2)$$

where $f_{X_i}(x_i)$ is the marginal PDF of X_i.

Then Borgonovo's importance measure is defined as [10]

$$\delta_i = \frac{1}{2} E_{X_i}[s(X_i)] \quad (3)$$

In a similar manner, Liu and Homma [14] proposed a new moment-independent importance measure concerning the CDF of model output. The deviation of $F_{Y|X_i}(y)$ from $F_Y(y)$ is measured by using the area $A(X_i)$ closed by the two CDFs, and can be calculated as

$$A(X_i) = \int |F_Y(y) - F_{Y|X_i}(y)|\, dy \quad (4)$$

The expected deviation of $F_{Y|X_i}(y)$ from $F_Y(y)$ can be obtained as

$$E_{X_i}[A(X_i)] = \int f_{X_i}(x_i) A(X_i)\, dx_i \quad (5)$$

$E_{X_i}[A(X_i)]$ can be used to illustrate the influence of the input X_i on the output distribution. Furthermore, for a general model of which the unconditional expectation of the output $E(Y)$ is not equal to 0, Liu and Homma defined that the CDF based sensitivity indicator, $S_i^{(CDF)}$, can be normalized as follows:

$$S_i^{(CDF)} = \frac{E_{X_i}[A(X_i)]}{|E(Y)|} \quad (6)$$

For models of which $E(Y)$ is equal to 0, we can compare the influences of the input variables on the output CDF by directly using $E_{X_i}[A(X_i)]$. Both δ_i and $S_i^{(CDF)}$ evaluate the influence of the entire distribution of the input X_i on that of the model output Y, except that the former focuses on the PDF of output while the latter on the CDF. For both measures, the effect of X_i on the output uncertainty is estimated by varying all the other inputs over their variation range. The two moment-independent importance measures represent the expected shift in the decision maker's view on the output uncertainty provoked by X_i. The inputs with larger values of δ_i or $S_i^{(CDF)}$ will be associated with higher importance, indicating these inputs should be paid more attention to if we want to modify the performance of the model. In fact, the two measures work regardless of whether the model is linear or nonlinear, additive or nonadditive, and can be easily extended to a group of inputs, for which more detailed information can be found in the work of Borgonovo [10,11], Liu and Homma [14].

2.2. Computational issues about the two measures

Firstly, let us talk about how to obtain the value of δ_i. It can be seen from Eqs. (1) and (2) that the calculation of δ_i includes two major parts, i.e. the first would be to get the unconditional and conditional PDFs of the output to obtain $s(X_i)$, and the second is to calculate the expectation of $s(X_i)$. Now we describe the computational method utilized by Borgonovo [10] to obtain δ_i, which generally can be viewed as a double-loop simulation method. First, the unconditional PDF of output, $f_Y(y)$, is obtained by uncertainty propagation, and varying the inputs over their variation range. Second, one generates a value for X_i, namely x_i^1 sampled from $f_{X_i}(x_i)$. Given this value, the conditional PDF of output, $f_{Y|X_i}(y)$, is obtained by propagating uncertainty in the model fixing $X_i = x_i^1$. Then $s(x_i^1)$ is computed from the numerical integration of the absolute value of the difference between $f_{Y|X_i}(y)$ and $f_Y(y)$. Repeat the procedure for a certain number of times, δ_i can be finally estimated. In fact, the calculation of $S_i^{(CDF)}$ can be implemented in a