Mechanical Systems and Signal Processing 109 (2018) 235–246

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

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

An efficient computational method of a moment-independent importance measure using quantile regression

Kaixuan Feng, Zhenzhou Lu *, Chao Pang, Wanying Yun

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

ARTICLE INFO

Article history:
Received 7 November 2017
Received in revised form 15 January 2018
Accepted 25 February 2018
Available online 8 March 2018

Keywords:
Moment-independent
Importance measure
Conditional probability density function
Non-linear quantile regression
Univariate kernel density estimation

ABSTRACT

The moment-independent uncertainty importance measure technique for exploring how uncertainty allocates from the output to the inputs has been widely used to help engineers estimate the degree of confidence of decision results and assess risks. The moment-independent importance measure (also called delta index) can better reflect the effect of the input on the whole distribution of the output instead of any specific moment. However, because the conditional probability density function (PDF) of the output is difficult to obtain, the computation process of delta index becomes quite complex. Therefore, an efficient computational algorithm by using the quantile regression is developed to estimate the delta index in this paper. Firstly, the non-linear quantile regression is employed to approximate the relationships between each input and the conditional quantiles of the output where only a set of input-output samples is needed. Secondly, at a certain value of the input, the conditional quantile points can be computed according to the obtained quantile regression models, which can be considered as the samples of the conditional output. Thirdly, the unconditional and conditional PDF of the output are evaluated by using the univariate kernel density estimation according to the original output samples and these quantile points respectively. Finally, the delta index is computed by estimating the area difference between the unconditional PDF and conditional PDF of the output. The number of model evaluations of this proposed method is dramatically decreased and is free of the dimensionality of the model inputs. Test examples show the performance of the proposed method and its usefulness in practice.

1. Introduction

Along with the constant development of computing technology and numerical methods, a variety of complicated computer models have arisen for predicting or simulating the performance of the mechanical systems at present [1–3]. These complicated models usually contain a large number of input random variables, whose uncertainties will lead the performances of the models to be random [4,5]. Nevertheless, as you can imagine, only a handful of inputs have significant effects on the model output in most cases [6]. Thereof, in order to help efficiently reduce the uncertainty of the model output, it is very necessary to distinguish the important inputs from a mass of model inputs [5–8].

* Corresponding author.
E-mail addresses: 18392592026@163.com (K. Feng), zhenzhoulu@nwpu.edu.cn (Z. Lu), pang_c16@mail.nwpu.edu.cn (C. Pang), wanying_yun@163.com (W. Yun).

https://doi.org/10.1016/j.ymssp.2018.02.044
0888-3270/

Sensitivity analysis (SA) is developed for studying the contributions of the uncertainties presented in the inputs to the outputs [4]. Specifically, SA can provide a generalized approach for identifying the important inputs and quantify their importance, fixing the unessential inputs at their nominal values and reducing model complicacy. Commonly, SA can be divided into two types: local SA and global SA [9,10]. Local SA is defined as the partial derivatives of the output with respect to the inputs at nominal values [11]. It is extensively applied to the linear models for its simplicity and efficiency. However, local SA cannot reflect the effect of the input on the output globally for the nonlinear models as it is only solved at nominal values. Compared with the local SA, global SA, also called the uncertainty importance measure (UIM) [3,7,12], focuses on quantifying the effects of the inputs on the uncertainties of outputs at the entire distribution ranges of those inputs. Recently, a variety of UIM techniques have emerged, such as the non-parametric UIM [13], the failure probability-based UIM [14–16], the variance-based UIM [4,11], and the moment-independent UIM [3,4,7,12]. The latter two kinds of techniques have gained the most attention as they are easy to interpret, model free, and stable.

The variance-based importance measure proposed by Helton and Saltelli focuses on distributing the model output variance to different sets of the model inputs according to the High Dimensional Model Representation (HDMR). It is global, quantitative and model free, thus has been studied widely by researchers in many fields [17], and many efficient solutions are investigated [5,18–20]. However, the use of variance as a measure of uncertainty implicitly supposes that the second moment is sufficient to describe output variability [4]. In fact, this assumption is not usually reasonable. As Helton and Davis mentioned, any moment of a random variable provides a summary of its distribution with the inevitable loss of resolution that occurs when the information contained in the distribution is mapped into a single number [21]. Thereof, Borgonovo and co-workers constructed the moment-independent importance measure (also called delta index) in order to overcome the shortage of the variance-based importance measure. The delta index can be interpreted as a measure of 'how much the shift on the probability density function (PDF) of model output on average while the inputs are fixed over their distribution ranges'.

At present, Wei and co-workers [22] proposed double-loop Monte Carlo simulation (MCS) method and single-loop MCS method to estimate the delta index respectively. Despite the high accuracy of the double-loop MCS method, the total computational cost involved may be not accepted by the practitioners in solving complicated engineering problems. Compared with the double-loop MCS method, the single-loop MCS method can simultaneously estimate all the delta indices by repeatedly making use of a set of samples. But it is worth noting that the joint PDF of two dependent variables needs to be estimated when solving the first order delta indices by adopting the single-loop MCS method, and only the PDF of single variable needs to be estimated by adopting the double-loop MCS method. It is well known that the estimating precision of the joint PDF of multivariate is lower and the computing complexity is higher in comparison with the univariate PDF. Therefore, to some degree, the precision of the delta index estimation is impaired by using the single-loop MCS method.

In the process of estimating the delta index, the most significant and difficult problem is how to obtain the conditional PDF of the model output rapidly and properly [12]. As has been known to many researchers, quantile regression has gradually become a quite useful tool for estimating the conditional quantile regression models [23–27]. By extending the concept of classical least-squares regression (LSR), quantile regression provides a comprehensive strategy for estimating how inputs affect the location, range and shape of the model output distribution. Therefore, the quantile regression is employed to describe the relationship between the input and the conditional quantiles of the model output in our proposed method (quantile regression based MCS method). After computing several conditional quantile points of the output when the input is fixed at a certain value, the conditional PDF of the model output can be estimated by using univariate kernel density estimation (KDE) method. Then, the rest of the computational procedure is similar to the double-loop MCS method. The quantile regression based MCS method needs only a set of samples for obtaining all the conditional quantile regression models of the model output at different quantiles, and the remainder of the computational procedure does not need to compute the model outputs, thus, it can dramatically reduce the number of model evaluations compared with the double-loop MCS method. Meanwhile, our proposed method only needs to estimate the univariate PDF, not the joint PDF of multivariate for estimating the delta indices, therefore, it is more accurate than the single-loop MCS method.

The paper is outlined as follows: Section 2 briefly reviews the definition of the delta index, and introduces the double-loop and the single-loop MCS methods for computing the delta indices. Section 3 introduces some mathematical preparations about quantile regression, and proposes the quantile regression based MCS method to estimate the delta index. Section 4 employs two numerical examples and two engineering examples to illustrate the accuracy and efficiency of the proposed method. Section 5 gives conclusions to this work.

2. Review of the moment-independent importance measure

2.1. The definition of the delta indices

Consider a computational model represented by $Y = g(\boldsymbol{X})$, where Y is the model output of interest, and $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)$ is n-dimensional input vector with random uncertainty. Denote $f_Y(y)$ and $f_{Y|X_i}(y)$ as the unconditional PDF and conditional PDF of Y respectively. $f_{Y|X_i}(y)$ can be obtained by supposing the input X_i fixed at one value.

In order to measure the entire effect of the uncertainty of an individual input X_i on the PDF $f_Y(y)$ of model output Y, the following importance measure (also called delta index) was defined by Borgonovo [3]: