Reliability Engineering and System Safety 96 (2011) 825–837

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

Global sensitivity analysis by polynomial dimensional decomposition

Sharif Rahman *

College of Engineering, The University of Iowa, Iowa City, IA 52242, USA

ARTICLE INFO

Article history:
Received 1 July 2010
Received in revised form
9 February 2011
Accepted 2 March 2011
Available online 8 March 2011

Keywords:
ANOVA
Sensitivity index
Orthogonal polynomials
Dimension reduction
Curse of dimensionality

ABSTRACT

This paper presents a polynomial dimensional decomposition (PDD) method for global sensitivity analysis of stochastic systems subject to independent random input following arbitrary probability distributions. The method involves Fourier-polynomial expansions of lower-variate component functions of a stochastic response by measure-consistent orthonormal polynomial bases, analytical formulae for calculating the global sensitivity indices in terms of the expansion coefficients, and dimension-reduction integration for estimating the expansion coefficients. Due to identical dimensional structures of PDD and analysis-of-variance decomposition, the proposed method facilitates simple and direct calculation of the global sensitivity indices. Numerical results of the global sensitivity indices computed for smooth systems reveal significantly higher convergence rates of the PDD approximation than those from existing methods, including polynomial chaos expansion, random balance design, state-dependent parameter, improved Sobol's method, and sampling-based methods. However, for non-smooth functions, the convergence properties of the PDD solution deteriorate to a great extent, warranting further improvements. The computational complexity of the PDD method is polynomial, as opposed to exponential, thereby alleviating the curse of dimensionality to some extent.

1. Introduction

Mathematical modeling of complex systems often requires sensitivity analysis to determine how an output variable of interest is influenced by individual or subsets of input variables. A traditional local sensitivity analysis entails gradients or derivatives, often invoked in design optimization, describing changes in the model response due to the local variation of input. Depending on the model output, obtaining gradients or derivatives, if they exist, can be simple or difficult. In contrast, a global sensitivity analysis (GSA), increasingly becoming mainstream, characterizes how the global variation of input, due to its uncertainty, impacts the overall uncertain behavior of the model. In other words, GSA constitutes the study of how the output uncertainty from a mathematical model is divvied up, qualitatively or quantitatively, to distinct sources of input variation in the model [1].

Almost all GSA are based on the second-moment properties of random output, for which there exist a multitude of methods or techniques for calculating the global sensitivity indices. Prominent among them are a random balance design (RBD) method [2], which integrates its previous version [3] with a Fourier amplitude sensitivity test [4]; a state dependent parameter (SDP) meta-model [5] based on recursive filtering and smoothing estimation; and a variant of Sobol's method with an improved formula [6–8]. More recent developments on GSA include application of polynomial chaos expansion (PCE) [9] as a meta-model, commonly used for uncertainty quantification of complex systems [10]. Crestaux et al. [11] examined the PCE method for calculating sensitivity indices by comparing their convergence properties with those from standard sampling-based methods, including Monte Carlo with Latin hypercube sampling (MC-LHS) [12] and quasi-Monte Carlo (QMC) simulation [13]. Their findings reveal faster convergence of the PCE solution relative to sampling-based methods for smoothly varying model responses, but the convergence rate may degrade markedly when confronted with non-smooth systems. They also found the PCE method to be cost effective for low to moderate dimensional systems, even with smooth responses, imposing a heavy computational burden when there exist a mere ten variables or more. Indeed, computational research on GSA is far from complete and, therefore, development of alternative methods for improving the accuracy or efficiency of existing methods is desirable.

This paper presents an alternative method, known as the polynomial dimensional decomposition (PDD) method, for variance-based GSA of stochastic systems subject to independent random input following arbitrary probability distributions. The method is based on (1) Fourier-polynomial expansions of lower-variate component functions of a stochastic response by measure-consistent orthonormal polynomial bases; (2) analytical formulae for calculating the global sensitivity indices in terms of the

* Tel.: +1 319 335 5679.
E-mail address: rahman@engineering.uiowa.edu

0951-8320/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ress.2011.03.002

expansion coefficients; and (3) dimension-reduction integration for efficiently estimating the expansion coefficients. Section 2 reviews a generic dimensional decomposition of a multivariate function, including three distinct variants. Section 3 invokes the properties of lower-variate component functions of a dimensional decomposition, leading to a formal definition of the global sensitivity index. The Fourier-polynomial expansion, calculation of sensitivity indices, dimension-reduction integration, including the computational effort, and novelties are described in Section 4. Five numerical examples illustrate the accuracy, convergence properties, and computational efficiency of the proposed method in Section 5. Finally, conclusions are drawn in Section 6.

2. Dimensional decomposition

Let $(\Omega,\mathcal{F},P)$ be a complete probability space, where Ω is a sample space, $\mathcal{F}$ is a σ-field on Ω, and $P:\mathcal{F}\to[0,1]$ is a probability measure. With $\mathcal{B}^N$ representing the Borel σ-field on $\mathbb{R}^N$, consider an $\mathbb{R}^N$-valued independent random vector $\boldsymbol{X}=\{X_1,\ldots,X_N\}^T:(\Omega,\mathcal{F})\to(\mathbb{R}^N,\mathcal{B}^N)$, which describes statistical uncertainties in all system and input parameters of a given stochastic problem. The probability law of $\boldsymbol{X}$ is completely defined by the joint probability density function $f_{\boldsymbol{X}}(\boldsymbol{x})=\prod_{i=1}^{i=N} f_i(x_i)$, where $f_i(x_i)$ is the marginal probability density function of X_i defined on the probability triple $(\Omega_i,\mathcal{F}_i,P_i)$. Let $y(\boldsymbol{X}):=y(X_1,\ldots,X_N)$, a real-valued, square-integrable, measurable transformation on $(\Omega,\mathcal{F})$, define a relevant response of the stochastic system. A general dimensional decomposition of $y(\boldsymbol{X})$, described by [14–20]

$$y(\boldsymbol{X})=y_\emptyset+\sum_{u\subseteq\{1,\ldots,N\}} y_u(\boldsymbol{X}_u), \tag{1}$$

can be viewed as a finite, hierarchical expansion of an output function in terms of its input variables with increasing dimensions, where $\emptyset\neq u\subseteq\{1,\ldots,N\}$ is a subset with the complementary set $-u=\{1,\ldots,N\}-u$ and cardinality $1\le|u|\le N$, $y_\emptyset$ is a constant, and $y_u(\boldsymbol{X}_u)$ is a $|u|$-variate component function describing the cooperative influence of $\boldsymbol{X}_u$, a subvector of $\boldsymbol{X}$, on y. The summation in Eq. (1) comprises 2^N-1 terms, with each term depending on a group of variables indexed by a particular subset of $\{1,\ldots,N\}$. The origin of dimensional decomposition can be traced to the work of Hoeffding [14] in the 1940s and is well known in the statistics literature as analysis of variance (ANOVA) [15]. This decomposition, later referred to as high-dimensional model representation (HDMR), was subject to further refinements, including cut-HDMR [16] and random-sampling (RS)-HDMR [17]. The author's group examined this decomposition from the perspective of Taylor series expansion, calculating the statistical moments [18,19] and reliability [20] of mechanical systems.

An important feature of the decomposition in Eq. (1) is the selection of the constant $y_\emptyset$ and component functions $y_u(\boldsymbol{X}_u)$, $\emptyset\neq u\subseteq\{1,\ldots,N\}$. By defining an error functional associated with a given $y(\boldsymbol{X})$ and an appropriate kernel function, an optimization problem can be formulated and solved to obtain the desired constant and component functions. However, different kernel functions will create distinct yet formally equivalent decompositions, all exhibiting the same structure of Eq. (1). There exist three important variants of the decomposition, described as follows.

2.1. Referential dimensional decomposition

The referential dimensional decomposition (RDD) involves the Dirac measure $\prod_{i=1}^{N}\delta(x_i-c_i)$ at a reference point $\boldsymbol{c}\in\mathbb{R}^N$ as the kernel function, leading to [16,19]

$$y(\boldsymbol{X})=y(\boldsymbol{c})+\sum_{u\subseteq\{1,\ldots,N\}}\sum_{v\subseteq u}(-1)^{|u|-|v|}y(\boldsymbol{X}_v,\boldsymbol{c}_{-v}), \tag{2}$$

where $(\boldsymbol{X}_v,\boldsymbol{c}_{-v})$ denotes an N-dimensional vector whose ith component is X_i if $i\in v$ and c_i if $i\notin v$. Both the recursive form, presented as the cut-HDMR method [16], and the explicit form, in conjunction with the dimension-reduction [19] or decomposition [20] method, of Eq. (2) exist. These two forms, developed independently, have been proved to be equivalent [21]. Nonetheless, the RDD component functions lack orthogonal features, but are easy to obtain as they only involve function evaluations at a chosen reference point.

2.2. ANOVA dimensional decomposition

The ANOVA dimensional decomposition (ADD) entails the probability density function $f_{\boldsymbol{X}}(\boldsymbol{x})$ of $\boldsymbol{X}$ as the kernel function, which results in [15,22]

$$y(\boldsymbol{X})=y_0+\sum_{u\subseteq\{1,\ldots,N\}}\sum_{v\subseteq u}(-1)^{|u|-|v|}\int_{\mathbb{R}^{N-|v|}}y(\boldsymbol{X}_v,\boldsymbol{x}_{-v})f_{\boldsymbol{X}_{-v}}(\boldsymbol{x}_{-v})\,d\boldsymbol{x}_{-v}, \tag{3}$$

where y_0 is an expansion coefficient. Again, there exists a recursive form of Eq. (3) [22]. The ANOVA decomposition also has a few synonyms, notably, Sobol decomposition, which has been used by Sudret [9] and Crestaux et al. [11], among others. While ADD has desirable orthogonal properties, the ANOVA component functions are difficult to obtain, because they require calculation of high-dimensional integrals.

2.3. Polynomial dimensional decomposition

If $\{\psi_{ij}(X_i);j=0,1,\ldots\}$ is a set of orthonormal polynomial basis functions in the Hilbert space $\mathcal{L}_2(\Omega_i,\mathcal{F}_i,P_i)$ and is consistent with the probability measure P_i of X_i, then the ANOVA decomposition can be extended to generate the PDD of [23,24]

$$y(\boldsymbol{X})=y_0+\sum_{u=\{i_1,\ldots,i_{|u|}\}\subseteq\{1,\ldots,N\}}\sum_{j_{i_{|u|}}=1}^{\infty}\cdots\sum_{j_{i_1}=1}^{\infty}C_{i_1\cdots i_{|u|}j_{i_1}\cdots j_{i_{|u|}}}\psi_{i_1j_{i_1}}(X_{i_1})\cdots\psi_{i_{|u|}j_{i_{|u|}}}(X_{i_{|u|}}), \tag{4}$$

where $C_{i_1\cdots i_{|u|}j_{i_1}\cdots j_{i_{|u|}}}$, $1\le|u|\le N$, are additional expansion coefficients that also require calculating high-dimensional integrals. The PDD also has orthogonal component functions and exploits the smoothness of y, if any, for efficiently calculating its probabilistic characteristics. The author's recent work reveals that the measure-consistent PDD [24] leads to faster convergence of stochastic solutions, when compared with the traditional ANOVA decomposition employing uniform probability measure, also known as RS-HDMR [17].

3. Global sensitivity analysis

3.1. Variance decomposition

The ADD in Eq. (3) can be written more explicitly as

$$y(\boldsymbol{X})=y_0+\sum_{i=1}^{N}y_i(X_i)+\sum_{i_1=1}^{N-1}\sum_{i_2=i_1+1}^{N}y_{i_1i_2}(X_{i_1},X_{i_2})+\cdots+\sum_{i_1=1}^{N-s+1}\cdots\sum_{i_s=i_{s-1}+1}^{N}y_{i_1\cdots i_s}(X_{i_1},\ldots,X_{i_s})+\cdots+y_{12\cdots N}(X_1,\ldots,X_N), \tag{5}$$