Computers and Structures 194 (2018) 86–96

Contents lists available at ScienceDirect

Computers and Structures

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

Adaptive sparse polynomial chaos expansions for global sensitivity analysis based on support vector regression

Kai Cheng, Zhenzhou Lu *

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

ARTICLE INFO

Article history:
Received 4 April 2017
Accepted 14 September 2017

Keywords:
Support vector regression
Sparse polynomial chaos expansion
Global sensitivity analysis
Adaptive kernel function

ABSTRACT

In the context of uncertainty analysis, Polynomial chaos expansion (PCE) has been proven to be a powerful tool for developing meta-models in a wide range of applications, especially for sensitivity analysis. But the computational cost of classic PCE grows exponentially with the size of the input variables. An efficient approach to address this problem is to build a sparse PCE. In this paper, a full PCE meta-model is first developed based on support vector regression (SVR) technique using an orthogonal polynomials kernel function. Then an adaptive algorithm is proposed to select the significant basis functions from the kernel function. The selection criterion is based on the variance contribution of each term to the model output. In the adaptive algorithm, an elimination procedure is used to delete the non-significant bases, and a selection procedure is used to select the important bases. Due to the structural risk minimization principle employing by SVR model, the proposed method provides better generalization ability compared to the common least square regression algorithm. The proposed method is examined by several examples and the global sensitivity analysis is performed. The results show that the proposed method establishes accurate meta-model for global sensitivity analysis of complex models.

1. Introduction

For decades, along with the rapid development of computer science and technique, a variety of complex computational models have been developed for simulating and predicting the behavior of systems in nearly all fields of engineering and science. In the meanwhile, the uncertainty of model structures and inputs are increasing rapidly. Uncertainty propagation through such complex models may become intractable in cases when a single simulation is computationally demanding. A remedy is to substitute a complex model with a meta-model that possesses similar statistical properties, but has a simple functional form [1].

Non-intrusive polynomial chaos expansion (PCE) method is popular for uncertainty propagation and global sensitivity analysis (GSA) to determine the effect of input uncertainties on complex computational models [2–5]. The key concept in PCE is to expand the model response onto basis made of multivariate polynomials that are orthogonal with respect to the joint distribution of the input variables. In this setting, characterizing the response probability density function (PDF) is equivalent to evaluating the PC coefficients, i.e. the coordinates of the random response in this basis [6]. The coefficients of the expansion are evaluated in terms of the response of the original model at a set of points in the input space, called the experimental design [1]. Although PCE has been proven to be powerful in a wide range of applications, they are unsuitable for high dimensional problems. It is believed that the number of the unknown expansion coefficients increases exponentially with the dimension in high-dimensional approximation [6–10]. To handle this issue, some sparse representations of the PCE have been studied in [6,8–10], where only a small number of significant basis functions are retained in the response PC approximation. These methods have been proved to be able to provide a significant computational gain compared to the classic full PCE for uncertainty propagation and GSA.

Non-intrusive PCE offers a number of benefits compared to other meta-model in uncertainty analysis [11].

- The full randomness of the response is contained within the set of the expansion coefficients;
- The mean and variance of the response are available in closed-form;
- PCE can be used with any second-order random process;
- PCE can handle many probability distribution types (normal, gamma, beta, poisson, etc.);
- PCE is transparent, simple to implement and have a strong mathematical basis;

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

https://doi.org/10.1016/j.compstruc.2017.09.002
0045-7949/

- PCE reaches fast convergence when the solutions are sufficiently smooth in the random space.

Thus, PCE has been studied in a variety of fields, such as GSA [2,9,10], structural dynamics [12], heat conduction [13] and fluid dynamics problems [14].

This paper aims at building sparse PCE based on support vector machine (SVM) for uncertainty propagation and GSA. SVM was first developed by Vapnik [15,16] and colleagues at AT&T Bell laboratories based on statistical-learning theory, it is a state-of-the-art supervised learning algorithm which has been used in many cases for regression (SVR) and classification (SVC) purposes [17–20]. Compared to other meta-model, SVR is based on the structural risk minimization principle [21], which leads to better generalization capacity. In the proposed method, a full PCE meta-model is first established. Then, similar as the idea of sparse PCE, an adaptive algorithm is introduced to detect the important bases and delete the non-significant basis functions. The adaptive algorithm selects the significant basis functions based on the variance contribution of each term to the model output. The major strength of the proposed method is that it could detect a group of basis functions simultaneously, thus it is efficient for high-dimensional problems. In the meanwhile, the proposed method provides better generalization ability due to the structural risk minimization principle. For validation purpose, two benchmark examples and two engineering applications are investigated for GSA, and the results of the sensitivity analysis are compared to those computed by sparse PCE in [9].

The rest of this paper is organized as follows: Section 2 reviews the methodology of GSA. After which, SVR model is described in detail in Section 3, including the theory foundation, kernel function and the parameters selection techniques. Section 4 is the core of this paper, the adaptive algorithm for building sparse PCE is presented and the corresponding GSA indices are derived based on this meta-model. Section 5 compares the proposed method with the sparse PCE for global GSA, and the conclusion comes in the end.

2. Global sensitivity analysis and Sobol' decomposition

GSA plays an important role in exploring the respective effects of input variables on an assigned output response. It can provide complete information by accounting for variations of the input variables in their entire domain, and then the priority level of the input variables can be obtained in experiments or research. The ranking of the input variables resulting from the GSA can help designers to decide how to reduce the uncertain scope of response. Amongst the wide range of literature on sensitivity analysis, the Sobol' indices have received more attention since they possess lot of perfect properties [22–24]. In this section, the methodology of the Sobol' decomposition and the corresponding sensitivity indices are reviewed.

Considering a square-integrable function defined in the unit hypercube $[0,1]^n$, n input parameters are gathered in vector $\boldsymbol{x}$, and a scalar output y is expressed as following:

$$y = g(\boldsymbol{x}) \tag{1}$$

The Sobol' decomposition of $g(\boldsymbol{x})$ has a representation of the following form [22–24],

$$g(x_1,\ldots,x_n) = g_0 + \sum_{i=1}^{n} g_i(x_i) + \sum_{1\leqslant i\leqslant j\leqslant n} g_{ij}(x_i,x_j) + \cdots + g_{1,2,\ldots,n}(x_1,\ldots x_n) \tag{2}$$

where g_0 is a constant and the integral of each summand $g_{i_1,\ldots i_s}(x_{i_1},\ldots x_{i_s})$ over any of its independent variables is zeros. $g(\boldsymbol{x})$ is orthogonally decomposed into summands of different dimensions by the representation in Eq. (2), namely,

$$E[g_{i_r,\ldots,i_s}(x_{i_r},\ldots x_{i_s}) \times g_{i_q,\ldots i_t}(x_{i_q},\ldots x_{i_t})] = 0 \quad \forall (i_r,\ldots i_s) \neq (i_q,\ldots i_t) \tag{3}$$

When the square of Eq. (2) is integrated on $[0,1]^n$, the following equation can be obtained.

$$\int g^2(\boldsymbol{x})d\boldsymbol{x} = g_0^2 + \sum_{i=1}^{n}\int g_i^2(x_i)dx_i + \sum_{1\leqslant i\leqslant j\leqslant n}\int g_{ij}^2(x_i,x_j)dx_i dx_j + \cdots + \int g_{1,2,\ldots,n}^2(x_1,\ldots x_n)dx_1,\ldots dx_n \tag{4}$$

It can be easily deduced that,

$$V = \sum_{i=1}^{n} V_i + \sum_{1\leqslant i\leqslant j\leqslant n} V_{ij} + \cdots + V_{1,2,\ldots,n} \tag{5}$$

where V is the total variance of $g(\boldsymbol{x})$, V_i is the partial variance contribution of x_i, and $V_{i_1,\ldots,i_s}$ is the cooperative variance contribution of $\{x_{i_1},\ldots,x_{i_s}\}$ which quantifies the interactions among $\{x_{i_1},\ldots,x_{i_s}\}$.

The Sobol' indices are defined as follows:

$$S_{\boldsymbol{I}} = \frac{V_{\boldsymbol{I}}}{V} = \frac{V[E(y|\boldsymbol{x}_{\boldsymbol{I}})]}{V(y)} \tag{6}$$

where $\boldsymbol{x}_{\boldsymbol{I}}$ represents a single random input variable x_i or a set of random input variables $\{x_{i_1},\ldots,x_{i_s}\}$. $S_{\boldsymbol{I}}$ in Eq. (6) expresses the portion of $V(y)$ explained by $\boldsymbol{x}_{\boldsymbol{I}}$. Practically, two variance-based sensitivity indices are very popular,

$$S_i = \frac{V_i}{V} = \frac{V[E(y|x_i)]}{V(y)} \tag{7}$$

S_i is the main effect of x_i, also called the first order Sobol' indices, which only measures the marginal variance contribution of x_i to the variance of the output, and

$$S_i^T = \frac{V_i^T}{V} = \frac{V[E(y|\boldsymbol{x}_{-i})]}{V(y)} \tag{8}$$

S_i^T is the total effect of x_i, where $\boldsymbol{x}_{-i}$ indicates all the input variables except x_i. Thus S_i^T measures the interaction contributions of x_i with all the other inputs.

3. Review on support vector regression

3.1. SVR network model

SVR is a powerful machine learning techniques developed by the statistical learning theory. It uses the adaptive margin-based loss functions and projects the learning data into a linear feature space, and finds the best decision function in the feature space. SVR is known for their good generalization performances and their ability to handle nonlinear models using the kernels technique. The optimization problem needing to be solved in SVR is convex, and thus a unique and global optimal solution can be guaranteed.

Given a set of training samples $\{z_i = (\boldsymbol{x}_i, y_i), i = 1,\ldots,N\}$, where $\boldsymbol{x}_i \in R^n$ is the training data, N is the number of training samples, and $y_i \in R$ is the response for $\boldsymbol{x}_i$. By solving the Kuhn-Tucker conditions of the following quadratic optimization problem

$$\begin{aligned} \min \quad & \tfrac{1}{2}\|\boldsymbol{\omega}\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) \\ s.t. \quad & y_i - \boldsymbol{\omega}\cdot\varphi(\boldsymbol{x}_i) - b \leqslant \varepsilon + \xi_i,\ \xi_i \geqslant 0 \\ & \boldsymbol{\omega}\cdot\varphi(\boldsymbol{x}_i) + b - y_i \leqslant \varepsilon + \xi_i^*,\ \xi_i^* \geqslant 0, \quad \forall i, i \in (1,2,\ldots,N) \end{aligned} \tag{9}$$