Computers and Geotechnics 89 (2017) 22–32

Contents lists available at ScienceDirect

Computers and Geotechnics

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

Research Paper

Stochastic response surface method for reliability problems involving correlated multivariates with non-Gaussian dependence structure: Analysis under incomplete probability information

Fan Wang [a,b,c,*], Heng Li [a]

[a] Department of Building and Real Estate, The Hong Kong Polytechnic University, Kowloon, Hong Kong
[b] Department of Civil Engineering and Mechanics, Huazhong University of Science and Technology, Wuhan 430074, PR China
[c] Department of Resource and Civil Engineering, Wuhan Institute of Technology, Wuhan 430073, PR China

ARTICLE INFO

Article history:
Received 23 September 2016
Received in revised form 9 January 2017
Accepted 9 February 2017

Keywords:
Reliability
Stochastic response surface method
Pair-copula
Vines
Rosenblatt transformation
Tunnels

ABSTRACT

This paper aims to provide a stochastic response surface method (SRSM) that can consider non-Gaussian dependent random variables under incomplete probability information. The Rosenblatt transformation is adopted to map the random variables from the original space into the mutually independent standard normal space for the stochastic surrogate model development. The multivariate joint distribution is reconstructed by the pair-copula decomposition approach, in which the pair-copula parameters are retrieved from the incomplete probability information. The proposed method is illustrated in a tunnel excavation example. Three different dependence structures characterized by normal copulas, Frank copulas, and hybrid copulas are respectively investigated to demonstrate the effect of dependence structure on the reliability results. The results show that the widely used Nataf transformation is actually a special case of the proposed method if all pair-copulas are normal copulas. The effect of conditioning order is also examined. This study provides a new insight into the SRSM-based reliability analysis from the copula viewpoint and extends the application of SRSM under incomplete probability information.

1. Introduction

There is an intrinsic uncertainty associated with the properties of geo-materials, and deterministic approaches to the stability analysis for geotechnical engineering cannot take these uncertainties into consideration. To address this problem, probabilistic methods, such as the first/second-order-order reliability method (FORM/SORM) and Monte Carlo simulation (MCS), have received a great deal of attentions because they are not only capable of handling uncertainty but also have the potential to support rational decision-making from a risk perspective [1].

One major challenge when applying FORM/SORM or MCS is to define the limit state function. Due to the complex underground conditions, the structure response has to be investigated using numerical approaches in many cases, which results in an implicit limit state function. To represent the limit state function in an explicit form, the response surface method (RSM) is widely adopted [2]. Various techniques, such as quadratic polynomials, artificial neural networks and support vector machines, are used to construct the response surface (e.g., [3–6]).

Unlike conventional deterministic RSM, the stochastic RSM (SRSM) uses polynomial chaos expansion (PCE) to model the input-output relationships in the standard random space [7]. The approximation by SRSM is accomplished by determining the coefficients associated with the PCE, which can be achieved by a probabilistic collocation method [8–10]. Generally, the approximation by PCE is valid across the entire *random* space [11], which is the major difference between the SRSM and the deterministic RSM. Li et al. [12] further extended the SRSM to consider correlated non-normal input variables for geotechnical reliability analysis by using the Nataf transformation [13].

It has been acknowledged that the Nataf transformation inherently assumes a Gaussian dependence structure for correlated multivariates [14]. From the copula viewpoint, it adopts a normal copula to characterize the underlying dependence structure [15,16]. However, recent investigations have demonstrated that this assumption does not always hold [17]. For non-Gaussian dependence structure cases, other copula functions should be used [18–20]. Unfortunately, in many reliability problems, the probabilistic description of the random vector is given in terms

* Corresponding author at: Department of Building and Real Estate, The Hong Kong Polytechnic University, Kowloon, Hong Kong.
E-mail address: wang_fan@hust.edu.cn (F. Wang).

http://dx.doi.org/10.1016/j.compgeo.2017.02.008
0266-352X/

of marginal distributions and correlations (referred to as the incomplete probability information [13]). Under such condition, the joint distribution cannot be uniquely determined because the dependence structure is not known [14,17,21]. In other words, the reliability evaluation based on Nataf transformation (normal copula) is only one of the various possible solutions.

To consider non-Gaussian dependent random variables under incomplete probability information, the joint distribution is constructed based on non-normal copulas with its parameter related to the correlation coefficient [14,17,19,20,22,23]. Since the dependence structure is unknown, there is no guideline on the copula selection. Several works [17,22–24] have examined the impact of different copulas on the reliability evaluations, and it is concluded that the results could be differed in a non-trivial way under different selections of copula. However, most of the extant studies are restricted to the bivariate cases [14,17,22–24]. If more than two random variables are mutually correlated, it is difficult to establish a one-to-one relationship between the pair-wise correlation coefficients and the multivariate copula parameters because their numbers are generally not equivalent [14]. This limitation can be attributed to the inflexibility of the conventional multivariate copulas in representing multivariate joint distributions with complex dependence structure.

Recently, construction of multivariate joint distribution by pair-copulas has drawn great attentions because it is highly flexible in modeling complex patterns of dependence by taking bivariate copulas as building blocks [25–28]. In this study, the pair-copula decomposition approach is adopted to represent the multivariate joint distribution. As a result, the number of pair-copula parameters is equivalent to the pair-wise correlation coefficients so that it is possible to relate them one by one. After the construction of the multivariate joint distribution, the Rosenblatt transformation [29] is used to establish a mapping relationship between the original space and the mutually independent standard normal space (i.e., U-space) for the SRSM model development. The proposed method is illustrated in a tunnel excavation example and it is compared to the SRSM with the Nataf transformation. The aim of this paper is to extend the SRSM to correlated multivariates with any dependence structure under incomplete probability information.

2. Collocation-based stochastic response surface method

2.1. Stochastic response surface method

In the SRSM, Hermite polynomials are widely adopted for functional approximation. Suppose that Y is the output random variable (i.e., system responses) and $\mathbf{X} = [x_1, x_2, \ldots, x_n]$ is the n-dimensional input random variable vector represented by a vector of independent standard normal variables $\mathbf{U} = [U_1, U_2, \ldots, U_n]$ as $\mathbf{X} = T(\mathbf{U})$. Then, the limit state function can be written as:

$$Y = G(\mathbf{X}) = G(T(\mathbf{U})) = H(\mathbf{U}) \tag{1}$$

Using the Hermite polynomials, $Y = H(\mathbf{U})$ can be written as:

$$Y = a_0 + \sum_{i_1=1}^{n} a_{i_1}\Gamma_1(U_{i_1}) + \sum_{i_1=1}^{n}\sum_{i_2=1}^{i_1} a_{i_1 i_2}\Gamma_2(U_{i_1}, U_{i_2}) + \sum_{i_1=1}^{n}\sum_{i_2=1}^{i_1}\sum_{i_3=1}^{i_2} a_{i_1 i_2 i_3}\Gamma_3(U_{i_1}, U_{i_2}, U_{i_3}) + \ldots \tag{2}$$

where $a_{i_1 i_2 \ldots i_n}$ are unknown coefficients, and $\Gamma_p(\cdot)$ is the multi-dimensional p-order Hermite polynomials given by:

$$\Gamma_p(U_{i_1}, \ldots, U_{i_p}) = (-1)^p e^{\frac{1}{2}\mathbf{U}^{\mathrm{T}}\mathbf{U}} \frac{\partial^p}{\partial U_{i_1}, \ldots, \partial U_{i_p}} e^{-\frac{1}{2}\mathbf{U}^{\mathrm{T}}\mathbf{U}} \tag{3}$$

Generally, the accuracy of approximation by Eq. (2) increases as the order p increases; however, higher order Hermite polynomials will incur a rapid increase of expansion terms. Hence, Eq. (2) is truncated at a specific order to achieve an accurate approximation, while the workload of deriving algebraic expressions is acceptable. For reference, Table 1 summarizes the closed-form of Eq. (2) from order 2 to 4.

Li et al. [12] summarizes the four major steps for establishing a stochastic surrogate model: (1) represent the input random variables by the independent standard normal random variables; (2) represent the output using the Hermite polynomials; (3) determine the coefficients associated to the Hermite polynomials using the collocation method; and (4) estimate the failure probability by available reliability techniques, e.g., MCS or FORM/SORM. The first step is critical particularly for correlated non-normal input random variables because the valid representation generally requires a nonlinear transformation from the original space to the U-space [12]. In Section 3, this transformation will be discussed in detail for correlated multivariates with non-Gaussian dependence structure.

2.2. Selection of collocation points

The unknown coefficients $a_{i_1 i_2 \ldots i_n}$ can be determined by the stochastic collocation method [30]. Similar to the deterministic collocation method, the roots of the next higher order Hermite polynomial are used as the stochastic collocation points. Then, the system responses are evaluated (e.g., by numerical approaches), and the coefficients are computed by:

$$\mathbf{a} = (\mathbf{H}^{\mathrm{T}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{F} \tag{4}$$

where $\mathbf{H}$ and $\mathbf{F}$ are the matrix of Hermite polynomials and the vector of system responses at the collocation points, respectively, and $\mathbf{a}$ is the coefficient vector to be solved.

For n-dimensional problems, the candidate collocation points for p-order polynomials are the combinations of the $(p+1)$-order Hermite polynomial roots. Note that the origin should be incorporated if it is not a collocation point because it captures the region of high probability in the standard normal space [7–9]. Thus, the number of candidate collocation points is:

$$N_{cp} = \begin{cases} (p+1)^n, & \text{if } p \text{ is even} \\ (p+1)^n + 1, & \text{if } p \text{ is odd} \end{cases} \tag{5}$$

If all the candidate collocation points are adopted to determine the polynomial coefficients, a relatively large number of real model runs have to be implemented, particularly when n and p are high. The minimum number of collocation points needed to determine the coefficients is:

$$N_a = \frac{(n+p)!}{n!p!} \tag{6}$$

Generally, $N_a \ll N_{cp}$. Thus, the number of realizations can be reduced if the collocation points are selected appropriately. Li et al. [12] noted that the collocation points should be selected to ensure that the Hermite polynomial matrix $\mathbf{H}$ has a full rank.

3. Transformation for correlated multivariates with non-Gaussian dependence structure under incomplete probability information

In engineering practice, correlations may exist among various random variables [31]. For example, the Young's modulus can be simultaneously correlated with the uniaxial compressive strength and the geological strength index of rock mass [32]. To consider