E3S Web of Conferences **360**, 01078 (2022) https://doi.org/10.1051/e3sconf/202236001078
VESEP2022

Nataf transformation based univariate decomposable polynomial RSM for engineering reliability analysis

Gaixia Liu[1,*] and *Wentong* Zhang[2]

[1]Education Center of Experiments and Innovations, Harbin Institute of Technology, Shenzhen, Guangdong, China
[2]School of Civil and Environmental Engineering, Harbin Institute of Technology, Shenzhen, Guangdong, China

Abstract. For engineering problems, the correlation exists among the common random variables. For those with highly correlated variables accompanying the high nonlinearity, large error would be induced if ignoring the influence of their correlation matrix. Considering correlated variables, an executing mode of polynomial response surface method based on Nataf transformation and univariate decomposition is introduced in this paper, called as N-UDPRSM. The correlated variables can be converted into the independent standard normal space and all of the univariate component polynomials can be determined separately. Besides, high order terms can be adopted into N-UDPRSM to balance the accuracy and efficiency for high nonlinear engineering problems. The corresponding practical implementation for engineering reliability analysis is designed in detail. A typical engineering structure is studied. The results indicate that it performs well in balancing accuracy and efficiency, and it also preserve some superiority contrast with other state-of-art methods.

1 Introduction

For the engineering reliability analysis, the basic problem is to compute the structural failure probability P_f, namely

$$P_f = \mathrm{Prob}\left\{F(\boldsymbol{\Theta}) \le 0\right\} = \int_{F(\boldsymbol{\Theta})\le 0} p_{\Theta}\left(\boldsymbol{\Theta}\right)\mathrm{d}\Theta \tag{1}$$

in which $\boldsymbol{\Theta} = (\Theta_1, \Theta_2, \ldots, \Theta_m)$ represents the structural random variables; $F(\boldsymbol{\Theta})$ represents the performance function which express the relationship between $\boldsymbol{\Theta}$ and the structural limit state; $p_{\Theta}(\boldsymbol{\Theta})$ represents the combined probability density function of $\boldsymbol{\Theta}$.

Generally, it is difficult to obtain the closed solution of Eq. (1) through direct numerical integration methods. Surrogate models have been proposed to compute this probability [1, 2]. The response surface method (RSM) is frequent-used and reliable for its own merits of satisfactory accuracy, efficiency and implementation. The basic idea of RSM is to

* Corresponding author: liugaixia@hit.edu.cn

approximate the implicit function $F(\boldsymbol{\Theta})$ by an implementable function [3-6]. The polynomial RSM may be the most easy-to-implement. One of its tasks is to calculate the undefined coefficients [7]. However, the dimensions of equation increase much faster than that of random variables, and it is the so-called "curse of dimension" question. For the high dimensional or high nonlinear systems, the difficulties of evaluating the undetermined coefficients and balancing the accuracy and efficiency still exist.

The univariate decomposition method (UDM) was used to reduce the computational cost for Eq. (1), which can decompose a multi-dimensional integration into separate integrations [8]. Combining the polynomial RSM and UDM, a new executing mode of adaptive polynomial RSM [9-10] has been developed, in which all of the univariate component polynomials can be determined separately. On the other hand, the correlation generally exists among common variables. For the highly correlated variables accompanying high nonlinearity, large error would be induced if ignoring the correlation for the engineering reliability analysis. The common way is to transform the correlated variables into independent standard normal space. Nataf transformation [11] is a widely and feasible used method for normal-to-non-normal transformation, when only the marginal PDFs and the considered correlation factors of the random variables are known.

In order to settle the problems with correlated variables for engineering reliability analysis with balancing the accuracy and efficiency, an execution for polynomial RSM based on Nataf transformation and UDM is introduced in this paper.

2 Univariate decomposable polynomial response surface model

2.1 The univariate decomposition method

Supposing $F(\boldsymbol{\Theta})$ is continuous, real-valued and differentiable, $F(\boldsymbol{\Theta})$ can be expressed in the Taylor series expansion, namely

$$F(\boldsymbol{\Theta})=F(\boldsymbol{c})+\sum_{j=1}^{\infty}\frac{1}{j!}\sum_{i=1}^{m}\frac{\partial^j F}{\partial \Theta_i^j}(\boldsymbol{c})(\Theta_i-c_i)^j+\Omega \tag{2}$$

where $\boldsymbol{\Theta} = (\Theta_1, \Theta_2, \ldots, \Theta_m)$ is random vector of the input; $\boldsymbol{c} = (c_1, c_2, \ldots, c_m)$ is the selected reference point; $F(\boldsymbol{c})$ is the value of $F(\boldsymbol{\Theta})$ at $\boldsymbol{c}$; Ω represents the residual error of omitting cross terms.

Based on UDM, $F(\boldsymbol{\Theta})$ can be approximated in a summary of univariate functions, namely

$$F(\boldsymbol{\Theta}) \approx \sum_{i=1}^{m} F(c_1,\cdots,c_{i-1},\Theta_i,c_{i+1},\cdots,c_{\mathrm{m}})-(m-1)F(\boldsymbol{c}) \tag{3}$$

where $F(c_1, c_2, \ldots, c_{i-1}, \Theta_i, c_{i+1}, \ldots, c_m)$ represents the component function with respect to Θ_i and it can be expanded at $\boldsymbol{c}$

$$F(c_1,\cdots,c_{i-1},\Theta_i,c_{i+1},\cdots,c_m)=F(\boldsymbol{c})+\sum_{j=1}^{\infty}\frac{1}{j!}\frac{\partial^j F}{\partial \Theta_i^j}(\boldsymbol{c})(\Theta_i-c_i)^j \tag{4}$$

Substituting Eq. (3) into Eq. (2) yields

$$F(\boldsymbol{\Theta}) \approx F(\boldsymbol{c})+\sum_{j=1}^{\infty}\frac{1}{j!}\sum_{i=1}^{m}\frac{\partial^j F}{\partial \Theta_i^j}(\boldsymbol{c})(\Theta_i-c_i)^j \tag{5}$$

Comparing Eq. (1) with Eq. (4), only Θ differs, which is so small that it can be neglected [8].

2.2 Univariate decomposable polynomial response surface model - UDPRSM

Based on RSM, a univariate polynomial $g_i(\Theta_i)$ can be adopted to approximate $F(c_1, c_2, \ldots, c_{i-1}, \Theta_i, c_{i+1}, \ldots, c_m)$ in Eq.(3), i.e.,

$$F(c_1 \cdots, c_{i-1}, \Theta_i, c_{i+1} \cdots, c_m) \approx g_i(\Theta_i) = \sum_{n=0}^{N} a_{i,n} \Theta_i^n \tag{6}$$

where N is the pre-set highest order and $a_{i,n}$ is the undefined coefficient.

Selecting $V = N+1$ samples around reference point, i.e., $\Theta_{i,1} = c_i+\gamma_1 r\sigma_i$, $\Theta_{i,2} = c_i+\gamma_2 r\sigma_i$, …, $\Theta_{i,V-1} = c_i+ \gamma_{V-1} r\sigma_i$, and $\Theta_{i,V} = c_i+\gamma_V r\sigma_i$, and determining the corresponding V samples of $F(\boldsymbol{\Theta})$, i.e., $g_{i,1}$, $g_{i,2}$, $g_{i,3}$, …, $g_{i,V-1}$, $g_{i,V}$, the coefficients $a_{i,0}$, $a_{i,1}$, $a_{i,1}$,…, $a_{i,N}$ can be confirmed by solving the simultaneous equations directly. Herein, r is the sample domain parameter, r = 3 empirically; γ_v (n=1, 2, …, V) is the interpolation coefficient, and the Chebyshev nodes are recommended in this work; σ_i is the standard deviation of Θ_i.

Therefore, substituting all determined $g_i(\Theta_i)$ (i=1, 2, …, m) into Eq. (3) generate the total model $G(\boldsymbol{\Theta})$, namely

$$F(\boldsymbol{\Theta}) \approx G(\boldsymbol{\Theta}) = \sum_{i=1}^{m} g_i(\Theta_i) - (m-1)F(\boldsymbol{c}) \tag{7}$$

It is noteworthy that although Eq. (7) is a polynomial as a whole, it is decomposed into m component response surface functions; for every component function $g_i(\Theta_i)$, its undefined coefficients are solved separately instead of simultaneously for the whole function.

3 Structural reliability analysis by UDPRSM considering variable correlation

For structure reliability analysis, RSM is used to approximate the structural implicit limit state function $F(\boldsymbol{\Theta})$ by an implementable function $G(\boldsymbol{\Theta})$. Therefore, the structural failure probability P_f can be rewritten as

$$P_f = \text{Prob}\left\{F(\boldsymbol{\Theta}) \le 0\right\} \;\Leftrightarrow\; P_f = \text{Prob}\left\{G(\boldsymbol{\Theta}) \le 0\right\} \tag{8}$$

Besides, in Section 2, UDPRSM provides a convenient model based on UDM. However, due to its inherent character, it seems that it is adverse for those cases with correlated variables. In order to settle this question, Nataf transformation [11] is introduced here.

3.1 Nataf transformation and its inversion

Supposing that $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_m)$ is standard normal random vector with the correlation matrix $\boldsymbol{\rho}_0=(\rho_{0ij})_{m\times m}$ and the random vector $\boldsymbol{\Theta} = (\Theta_1, \Theta_2, \ldots, \Theta_m)$ is with the correlation matrix $\boldsymbol{\rho}=(\rho_{ij})_{m\times m}$, by the isoprobabilistic transformation, the relationship between $\boldsymbol{\Theta}$ and $\boldsymbol{Y}$ can be expressed as

$$\Theta_i = F_{\Theta i}^{-1}\left[\boldsymbol{\Phi}\left(y_i\right)\right] \quad i = 1, 2, \ldots, m \tag{9}$$

where $F_{\Theta i}(\bullet)$ is the marginal cumulative distribution function of Θ_i, and $F^{-1}_{\Theta i}(\bullet)$ is the inverse function of $F_{\Theta i}(\bullet)$.

Generally, $\boldsymbol{\rho} = (\rho_{ij})_{I\times I}$ can be determined beforehand and for some common distribution types, $\boldsymbol{\rho}_0=(\rho_{0ij})_{I\times I}$ can be approximated by the following empirical formula

$$\rho_{0ij} = F \cdot \rho_{ij} \tag{10}$$

where $F \geq 1$ and it is the function of $F_{\Theta i}(\bullet)$ and ρ_{ij}.

Then, by Cholesky decomposition, there is

$$\boldsymbol{\rho}_0 = \boldsymbol{L}_0 \boldsymbol{L}_0^T \tag{11}$$

where $\boldsymbol{L}_0$ is the lower triangular decomposition matrix of $\boldsymbol{\rho}_0$. Left multiplying by $\boldsymbol{L}_0^{-1}$, $\boldsymbol{Y}$ can be converted into independent standard normal random vector $\boldsymbol{Z} = (Z_1, Z_2, \ldots, Z_m)$, i.e.,

$$\boldsymbol{Z} = \boldsymbol{L}_0^{-1}\boldsymbol{Y} \tag{12}$$

From the above, the forward process $\boldsymbol{T}$ of Nataf transformation can be briefly expressed as the following form

$$\boldsymbol{T}: \quad \boldsymbol{Z} = \boldsymbol{L}_0^{-1}\boldsymbol{Y} = \boldsymbol{L}_0^{-1}\begin{Bmatrix} \boldsymbol{\Phi}^{-1}\left[F_{\Theta i}\left(\Theta_1\right)\right] \\ \vdots \\ \boldsymbol{\Phi}^{-1}\left[F_{\Theta m}\left(\Theta_m\right)\right] \end{Bmatrix} \tag{13}$$

By simple back analysis, its inversion $\boldsymbol{T}^{-1}$ can be obtained, expressed as

$$\boldsymbol{T}^{-1}: \quad \boldsymbol{\Theta} = \boldsymbol{Y}\boldsymbol{L}_0 = \begin{Bmatrix} F_{\Theta 1}^{-1}\left[\boldsymbol{\Phi}\left(z_1\right)\right] \\ \vdots \\ F_{\Theta m}^{-1}\left[\boldsymbol{\Phi}\left(z_m\right)\right] \end{Bmatrix}\boldsymbol{L}_0 \tag{14}$$

Therefore, by the forward process $\boldsymbol{T}$, $\boldsymbol{\Theta}$ can be transformed into the independent standard normal random vector $\boldsymbol{Z}$, and by its inversion $\boldsymbol{T}^{-1}$, $\boldsymbol{Z}$ can be transformed into $\boldsymbol{\Theta}$.

3.2 Practical implementation for engineering reliability analysis

Combined with Nataf transformation and its inversion, the implementation procedure of UDPRSM for engineering reliability analysis is designed, denoted as N-UDPRSM in this work. N-UDPRSM can be divided into the following five stages.

Stage 1. The preparatory work.

(1) Select the random vector $\boldsymbol{\Theta} = (\Theta_1, \Theta_2, \ldots, \Theta_m)$, which can influence the structure or its response uncertainty; and determine their probability distribution and the correlation coefficient matrix $\boldsymbol{\rho}$.

(2) Select the pre-set degree N of polynomial, i.e., determine the response surface model.

Stage 2. The initial response surface establishment.

(1) Take the mean point of $\boldsymbol{\Theta}$ as the reference point, i.e., $\boldsymbol{c} = \boldsymbol{\mu}$, and then calculate $F(\boldsymbol{\mu})$.

(2) For each Θ_i (i=1, 2, ..., m), run the following sub-steps to determine the component response surfaces.

(I) Select V=N+1 samples $\Theta_{i,v}$ (v=1,2,..., V) and meanwhile the other random variables Θ_j ($j\neq i$) are set to their reference values c_j, i.e., $\boldsymbol{\Theta}_{i,v}$=($c_1$, c_2, …, c_{i-1}, $\Theta_{i,v}$, c_{i+1},…, c_m); and calculate their corresponding response values $F(\boldsymbol{\Theta}_{i,v})$=$g_{i,v}$, ($v$=1,2,..., V);

(II) By Nataf transformation Eq. (13), transform those V vector samples $\boldsymbol{\Theta}_{i,v}$ into $\mathbf{Z}_{i,v}$ = (z_1, z_2, …, z_m), which are obey to the random vector $\boldsymbol{Z}$ = (Z_1, Z_2, …, Z_m) in the standard Gaussian space;

(III) Taking the above $\mathbf{Z}_{i,v}$ and $F(\boldsymbol{\Theta}_{i,v})$ as the input and output factors respectively, fit the component response surface function $g_i(Z_i)$ by solving Eq.(6), in which variable Θ is replaced by Z.

(3) Substitute all $g_i(Z_i)$ (i=1, 2, ..., m) and $F(\boldsymbol{c})$ into Eq.(7), and yield the initial response surface model with respect to Z in the standard normal space, noted as $G_0(\boldsymbol{Z})$.

Stage 3. Searching the new reference centre.

(1) Based on $G_0(\boldsymbol{Z})$, search for the design point $\boldsymbol{Z}^*$ by the first order reliability method (FORM) in the standard normal space.

(2) By the inverse Nataf transformation Eq. (14), transform $\boldsymbol{Z}^*$ into $\boldsymbol{\Theta}^*$, which would be taken as the new reference centre in the physical space.

Stage 4. Establishment of the final response surface.

(1) Set $\boldsymbol{c}$= $\boldsymbol{\Theta}^*$ and calculate $F(\boldsymbol{c})$.

(2) Repeat the **Stage 2**-(2) step to approximate all of $g_i(Z_i)$ (i=1,..., m) again.

(3) Substitute all of $g_i(Z_i)$ (i=1,..., m) and $F(\boldsymbol{c})$ into Eq. (7), and yield the final response surface model in the standard norm space, noted as $G(\boldsymbol{Z})$.

Stage 5. Estimation of the reliability index β associated with $G(\boldsymbol{Z})$ by FORM.

4 Case study

As Fig.1 shown, a frame structure is studied in this section. This example includes some unique features, such as correlation matrix highly correlated, high uncertainties, high dimensionality and highly nonlinear response.

There are 35 members in this frame structure and their properties are referred to [3], [4], [5], [6], etc. The coefficients of correlation among all loads F_i (i=1, 2, 3) are ρ_F = 0.95, those of all cross-section A_i, (i=1, 2, …, 8) are ρ_{AiAj} = 0.13, those of I_i (i=1, 2, …, 8) are ρ_{IiIj} = 0.13, that between E_1 and E_2 is ρ_E = 0.9, and all other variables are mutually independent. The limit horizontal displacement Δx of the structural top is set to 0.061 m, with regarding to the serviceability failure state. The corresponding performance function is

$$g(\boldsymbol{\Theta}) = 0.061 - \Delta x \tag{15}$$

where $\boldsymbol{\Theta}$ is the random vector concluding all the above physical parameters and forces.

Herein, the generalized reliability index β is present in order to provide a visualized result, i.e., β = $-\boldsymbol{\Phi}^{-1}(P_f)$. For this case, through Importance sampling, the reference β was obtained in Ref. [5], i.e., β=3.51 with P_f=2.24×10^{-4}. Table 1 shows the results, in which N-UDPRSM-j means the pre-set order of the univariate polynomial is j in **Stage 1** of Section 3.2. The pre-set order is set to1~4, respectively.

For N-UDPRSM, as the pre-set order increases, the number of samples increases, the calculated reliability index β decreases and the corresponding error decreases. It is revealed that the surrogate model in N-UDPRSM performs well in fitting the performance function $g(\boldsymbol{\Theta})$ and the relatively higher order can improve the accuracy with some loss of efficiency. Besides, the pre-set order 2 is good enough for this case.

The sparse polynomial chaos expansions (PCE) [4], the sparse polynomial RSM [5] and the sparse Kriging-based model [6] are used to study the reliability of this structure, respectively. Contrast with these surrogate models, N-UDPRSM-3 performs the best considering both accuracy and efficiency; further, the advantages of the proposed N-UDPRSM-2 in efficiency are superior enough while keeping better precision.

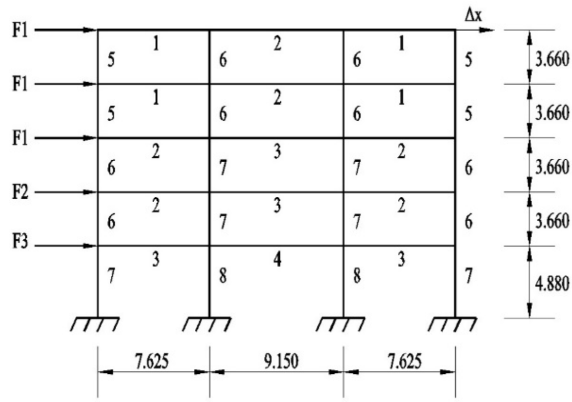

Fig. 1. A frame structure layout (length unit, m).

Table 1. Computational results of the three-bay five-story frame reliability.

Method	Number of samples	β	Error of β /%
N-UDPRSM-1	45	3.866	10.14
N-UDPRSM-2	87	3.398	-3.20
N-UDPRSM-3	129	3.565	1.56
N-UDPRSM-4	171	3.559	1.40
Full PCE [4]	3724	3.60	-2.56
Sparse PCE [4]	450	3.61	-2.85
Sparse polynomial RSM [5]	149	3.63	-3.42
Sparse Kriging-based model - 1 [6]	166	3.718	-5.93
Sparse Kriging-based model - 2 [6]	222	3.732	-6.33
Reference value [5]	500,000	3.51	--

5 Conclusions

In this paper, by combining Nataf transformation with univariate decomposable polynomial RSM, N-UDPRSM is proposed to analyse the engineering structural reliability considering correlated variables. By applying the general model of UDPRSM and Nataf transformation and its inversion, the detailed procedure of practical implementation for engineering reliability analysis is designed. Moreover, a typical example is investigated to evaluate the performance of the proposed method. It is revealed that the proposed method could be well applied to the practical engineering with satisfactory accuracy and efficiency, and it also preserve some superiority by contrast with other state-of-art methods.

References

1. R. Rackwitz, Structural Safety **23** (2001)
2. M. Aldosary, J. Wang, C. Li, Engineering Computations **35** (2018)
3. C. Bucher, U. Bourgund, Structural Safety **7** (1990)
4. G. Blatman, B. Sudret, Probabilistic Engineering Mechanics **25** (2010)

5. N. Roussouly, F. Petitjean, M. Salaun, Probabilistic Engineering Mechanics **32** (2013)
6. H. Guimarães, JC. Matos, AA. Henriques, Structural Safety **73** (2018)
7. S. Goswami, S. Ghosh, S. Chakraborty, Structural Safety **60** (2016)
8. S. Rahman S, H. Xu, Probabilistic Engineering Mechanics **19** (2004)
9. W. Zhang, and Y. Xiao, Engineering Computations **36** (2019)
10. W. Zhang, Y. Xiao, Structural Engineering and Mechanics **76** (2020)
11. PL. Liu, D. Kiureghian, Probabilistic Engineering Mechanics **1** (1986)