A hybrid dimensional reduction method for uncertainty analysis of structures

Zhou Changcong
Associate Professor, School of Mechanics, Civil Engineering and Architecture, Northwestern Polytechnical University, Xi'an, China

Li Haihe
Graduate Student, School of Mechanics, Civil Engineering and Architecture, Northwestern Polytechnical University, Xi'an, China

Chang Qi
Graduate Student, School of Mechanics, Civil Engineering and Architecture, Northwestern Polytechnical University, Xi'an, China

Yuan Xiukai
Associate Professor, School of Aerospace Engineering, Xiamen University, Xiamen, China

ABSTRACT: In order to address the issue of uncertainty analysis of large scale structures with high-dimensional input variables in the engineering, a hybrid dimensional reduction method (H-DRM) is proposed in this work. The method can be seen as an improvement on the multiplicative dimensional reduction method (M-DRM). First, the original response function is approximated by the univariate M-DRM, based on which the variance contribution to the output variance can be estimated by three point estimate. The significant variables are thereby identified and further extended to the bivariate M-DRM. The final approximation of the original response function can be seen as a hybrid combination of univariate M-DRM of all variables and bivariate M-DRM of significant variables. The proposed H-DRM is then used for the calculation of moments and structural reliability, with the help of three point estimate. The proposed method can adaptively achieve the balance of computational efficiency and accuracy, and is able to avoid the "curse of dimension" when using three point estimate. Several examples are studied to demonstrate the feasibility of the proposed method.

1. INTRODUCTION

In the past few decades, scholars have done a lot of research on the representation and decomposition of high-dimensional problems. Rabitz et al. have done a lot of research on the input-output relationship of physical systems with multiple input variables, and proposed a high-dimensional model representation (HDMR), and deduced a cut high-dimensional model representation (Cut-HDMR) based on HDMR[1]. The Cut-HDMR assumes that the univariate or the low-order correlative terms of the input variable have major influence on the system output, while the effect of the high-order correlative terms are negligible. Sobol's research suggests that inappropriate reference points can produce poor approximation accuracy and gives recommendations for reference point selection[2]. Rahman and Xu considered the requirements of structural probability analysis, and obtained a structural response function approximation model with only univariate function (univariate decomposition approximation)[3]. In fact, Zhao did the same work at first, but he did not continue to pay attention to high-order related items[4]. Xu and Rahman extended the results of univariate research to bivariate functions (bivariate decomposition approximation) and s-dimensional function (s<n, s-variate decomposition approximation), and obtained similar results with Rabitz et al[5]. Li et al. extended the definitions of HDMR component functions to systems of which input variables may not be independent, and proposed new orthonormal polynomial approximation formulas for the random sampling-high dimensional model representation (RS-HDMR) component functions that preserve the orthogonality property[6]. Subsequently, Rahman and his collaborators extended the application of dimensional decomposition methods to structural reliability analysis[7], reliability-based design optimization[8], stochastic sensitivity analysis[9], and probabilistic fracture mechanics[10]. Recently, Zhang et al. proposed a multiplicative dimensional reduction method (M-DRM) to approximate the original high-dimensional functions[11]. That method has a significant advantage in calculating the statistical moment of the response function. However, for high-dimensional complex engineering problems, the univariate M-DRM may be not accurate enough to obtain acceptable results, and the bivariate decomposition may lead to too much computational cost. A new method has therefore been proposed in this paper to overcome the disadvantages of the univariate M-DRM and bivariate M-DRM.

This paper summarizes the previous development of the dimensional reduction method and proposes an efficient method for uncertainty analysis based on multiplicative dimensional reduction method. The multiplicative dimensional reduction method approximates the original high-dimensional performance function as the product of a series of functions increasing dimensions, and it is an efficient method for calculating fractional moments of performance function.

2. REVIEW OF THE MULTIPLICATIVE DI MENSIONAL REDUCTION METHOD

According to the high-dimensional model representation, a high-dimensional function can be expressed as a sum of functions of lower order in an increasing hierarchy as,

$$g(\mathbf{x}) = g_0 + \sum g_i(x_i) + \sum_{i<j} g_{ij}(x_i, x_j) + \cdots + \sum_{i<j<\cdots<k} g_{ij\cdots k}(x_i, x_j, \ldots, x_k) + \cdots + g_{1\cdots n}(x_1, \ldots, x_n) \tag{1}$$

where $\mathbf{x}=\{x_1, x_2, \cdots, x_n\}^{\mathrm{T}}$ is the vector of input variables and g_0 denotes the constant term of the multidimensional function , which is calculated at the cut-point $\boldsymbol{c}=[c_1, c_2, \cdots, c_n]^{\mathrm{T}}$. The function $g_i(x_i)$ represents the output of the system when only the i-th variable x_i acts independently on the system, the function $g_{ij}(x_i, x_j)$ describes the effect of the correlated terms of the i-th variable x_i and the j-th variable x_j on the system output. The meaning of the rest terms can be deduced by analogy.

We should note that:

$$\begin{cases} g_0 = g(\boldsymbol{c}) \\ g_i(x_i) = g(c_1, \cdots, c_{i-1}, x_i, c_{i+1}, \cdots, c_n) - g_0 \\ g_{ij}(x_i, x_j) \\ = g(c_1, \cdots, c_{i-1}, x_i, c_{i+1}, \cdots, c_{j-1}, x_j, c_{j+1}, \cdots, c_n) \\ -g_i(x_i) - g_j(x_j) - g_0 \\ \cdots \end{cases} \tag{2}$$

Univariate Cut-HDMR is obtained by retaining only the first two terms in Eq.(1), and the performance function is given as:

$$\hat{g}_1(\mathbf{x}) = \sum_{i=1}^{n} g(c_1, \ldots, c_{i-1}, x_i, c_{i+1}, \ldots, c_n) - (n-1) g(\boldsymbol{c}) \tag{3}$$

If the high-order terms have significant impact on the response function, the bivariate Cut-HDMR is obtained by retaining the first three terms in Eq.(1):

$$\hat{g}_2(\mathbf{x}) = \sum_{1\le i<j\le n} g\left(x_i, x_j, \boldsymbol{c}_{-ij}\right) - (n-2)\sum_{i=1}^{n} g\left(x_i, \boldsymbol{c}_{-i}\right) + \frac{(n-1)(n-2)}{2} g(\boldsymbol{c}) \tag{4}$$

In general, the s-variate Cut-HDMR of original performance function is derived as:

$$\hat{g}_s(\mathbf{x}) = \sum_{k=0}^{s} (-1)^{s-k} \binom{n-k-1}{s-k} \sum_{1\le i_1<\cdots<i_k\le n} g\left(x_{i_1}, x_{i_2}, \cdots, x_{i_k}, c_{-i_1,i_2\cdots i_k}\right) \tag{5}$$

where sth and low-order correlated contribution of all input variables are considered.

By applying logarithmic transformation of the response function, we can obtain:

$$\varphi(\mathbf{x}) = \ln(y) = \ln\left[g(\mathbf{x})\right] \tag{6}$$

Substituting Eq.(3) into Eq.(6), the approximation of $\varphi(\mathbf{x})$ can be rewritten as

$$\varphi(\mathbf{x}) = \sum_{i=1}^{n} \varphi\left(x_i, \boldsymbol{c}_{-i}\right) - (n-1)\varphi_0 \tag{7}$$

where

$$\begin{cases} \varphi_0 = \ln\left(g_0\right) \\ \varphi\left(x_i, \boldsymbol{c}_{-i}\right) = \ln\left[g\left(x_i, \boldsymbol{c}_{-i}\right)\right] \\ \varphi\left(x_i, x_j, \boldsymbol{c}_{-ij}\right) = \ln\left[g\left(x_i, x_j, \boldsymbol{c}_{-ij}\right)\right] \\ \cdots \end{cases} \tag{8}$$

Exponential transformation is inverted into Eq.(7) to achieve that:

$$\exp\left[\varphi(\mathbf{x})\right] = \exp\left[\sum_{i=1}^{n} \varphi\left(x_i, \boldsymbol{c}_{-i}\right) - (n-1)\varphi_0\right] \tag{9}$$

Substituting Eq.(6) and Eq.(8) into Eq.(9), the response function can be rewritten as

$$\hat{g}_1(\mathbf{x}) = \left[g(\boldsymbol{c})\right]^{1-n} \prod_{i=1}^{n} g\left(c_1, \ldots, c_{i-1}, x_i, c_{i+1}, \ldots, c_n\right) \tag{10}$$

Substituting Eq.(4) into Eq.(6), and following the procedures in Eq.(7) to Eq.(9), the bivariate M-DRM can be easily derived as:

$$\hat{g}_2(\mathbf{x}) = \frac{\left[g(\boldsymbol{c})\right]^{\frac{(n-1)(n-2)}{2}} \times \prod_{i=1}^{n-1} \prod_{j=i+1}^{n} g\left(x_i, x_j, \boldsymbol{c}_{-ij}\right)}{\left[\prod_{i=1}^{n} g\left(x_i, \boldsymbol{c}_{-i}\right)\right]^{n-2}} \tag{11}$$

In general, following the same procedure, the s-variate M-DRM is achieved as:

$$\hat{g}_s(\mathbf{x}) = \frac{\prod_{k_1=s,s-2,s-4,\cdots} \left[\prod_{i_1=1}^{n-k_1+1} \cdots \prod_{i_{k_1}=i_{k_1-1}+1}^{n} g\left(x_{i_1}, x_{i_2}, \cdots, x_{i_k}, \boldsymbol{c}\right)\right]^{\binom{n-k_1-1}{s-k_1}}}{\prod_{k_2=s-1,s-3,s-5,\cdots} \left[\prod_{j_1=1}^{n-k_2+1} \cdots \prod_{j_{k_2}=j_{k_2-1}+1}^{n} g\left(x_{j_1}, x_{j_2}, \cdots, x_{j_k}, \boldsymbol{c}\right)\right]^{\binom{n-k_2-1}{s-k_2}}} \tag{12}$$

3. THE PROPOSED METHOD

In this section, the univariate M-DRM introduced in the previous section is utilized to approximate the original input-output relationship. In order to estimate the first four moments of response function, three-point estimation method is introduced[13]. The basic idea of the three-point estimation method is to use selected points and corresponding weights to approximate the integral of the function. The original n-dimensional response function is approximated as the product of n one-dimensional functions, so the original n-dimensional integral is approximated by employing n one-dimensional integrals. Then the first four integer moments of response function can be estimated via the following relation

$$M_g^{\alpha} = E\left[Y^{\alpha}\right] = \int g(\mathbf{x})^{\alpha} f(\mathbf{x}) \mathrm{d}\mathbf{x} \tag{13}$$

where M_g^{α} indicates the first four integer moments of response function $g(\mathbf{x})$, α=1,2,3,4, $f(\mathbf{x})$ is the joint probabilistic density function.

Then, the three-point estimation method is used to calculate the above integrals. p_{ij} and l_{ij} are corresponding weight point and nominal value of x_i and the weight points and nominal values can be calculated by the following equations,

$$p_{x\bullet 1} = \frac{1}{2}\left(\frac{1+\alpha_{3x}/\sqrt{4\alpha_{4x}-3\alpha_{3x}^2}}{\alpha_{4x}-\alpha_{3x}^2}\right)$$
$$p_{x\bullet 2} = 1-\frac{1}{\alpha_{4x}-\alpha_{3x}^2}$$
$$p_{x\bullet 3} = \frac{1}{2}\left(\frac{1-\alpha_{3x}/\sqrt{4\alpha_{4x}-3\alpha_{3x}^2}}{\alpha_{4x}-\alpha_{3x}^2}\right) \quad (14)$$
$$l_{x\bullet 1} = \alpha_{1x}-\frac{\alpha_{2x}}{2}\left(\sqrt{4\alpha_{4x}-3\alpha_{3x}^2}-\alpha_{3x}\right)$$
$$l_{x\bullet 2} = \alpha_{1x}$$
$$l_{x\bullet 3} = \alpha_{1x}+\frac{\alpha_{2x}}{2}\left(\sqrt{4\alpha_{4x}-3\alpha_{3x}^2}+\alpha_{3x}\right)$$

where α_{jx} denotes the first four moments of x_i, $j=1,2,3,4$, i.e. mean, standard deviation, skewness, kurtosis of x_i, which can be easily estimated if the distribution type of x_i is given.

The relationship between the central moment and the integer moment is given by:

$$\mu_g = M_g^1 \quad (15)$$
$$\sigma_g = \sqrt{M_g^2-\left(M_g^1\right)^2} \quad (16)$$
$$\alpha_{3,g} = \frac{1}{\sigma_g^3}\left[M_g^3-3M_g^2M_g^1+2\left(M_g^1\right)^3\right] \quad (17)$$
$$\alpha_{4,g} = \frac{1}{\sigma_g^4}\left[M_g^4-4M_g^3M_g^1+6M_g^2\left(M_g^1\right)^2-3\left(M_g^1\right)^4\right] \quad (18)$$

In the above formula, μ_g is the mean of performance function, σ_g is the standard deviation of performance function, $\alpha_{3,g}$ and $\alpha_{4,g}$ are the coefficient of the skewness and the coefficient of kurtosis of performance function, respectively.

From the discussion in the previous sections, it can be seen that the univeriate M-DRM method may be not accurate enough and the bivariate M-DRM method is too complex in the cases of some high-dimensional engineering problems. For improving computational efficiency and accuracy, we propose to perform bivariate decomposition only on important variables. In order to measure the importance of input variables, the main sensitivity index proposed by Sobol is introduced as:

$$S_i = \frac{V_i}{V(Y)} = \frac{V\left[E\left(Y|X_i\right)\right]}{V(Y)} \quad (19)$$

where $V\left[E\left(Y|X_i\right)\right]$ denotes the average reduction in the variance of the performance function when fixing X_i over its full distribution range, $V(Y)$ is the unconditional variance of the performance function. The sensitivity index S_i can estimate the influence of individual variable X_i on the model output and can be used to identify the importance of input variables.

According to the derivation of Zhang[14], the Eq.(19) can be rewritten as:

$$S_i = \frac{V\left[E\left(Y|X_i\right)\right]}{V(Y)} = \frac{\left[M_{g_i}^2-\left(M_{g_i}^1\right)^2\right]\prod_{s=1,s\neq i}^{n} M_{g_s}^2}{\prod_{i=1}^{n}\left[M_{g_i}^2-\left(M_{g_i}^1\right)^2\right]} \quad (20)$$

As can be seen from Eq.(20), the main sensitivity index is represented by a combination of the first and second moments of a series of univariate functions. After *m* important input variables are identified, we consider the influence of the second-order correlated contribution of them on the response function. In the third term of Eq.(1), only retaining the second-order correlation terms of those m variables, the univariate-bivariate hybrid Cut-HDMR can be determined as

$$\begin{aligned}\hat{g}_{mixed}(\mathbf{x},\mathbf{y}) &= \sum_{i=1}^{n} g\left(x_i,\boldsymbol{c}_{-i}\right)-(n-1)g(\boldsymbol{c})+\frac{m(m-1)}{2}g(\boldsymbol{c}) \\ &+\sum_{1\le i<j\le m} g\left(y_i,y_j,\boldsymbol{c}_{-ij}\right)-(m-1)\sum_{i=1}^{m} g\left(y_i,\boldsymbol{c}_{-i}\right) \\ &=\sum_{1\le i<j\le m} g\left(y_i,y_j,\boldsymbol{c}_{-ij}\right)+\sum_{i=1}^{n} g\left(x_i,\boldsymbol{c}_{-i}\right) \\ &-(m-1)\sum_{i=1}^{m} g\left(y_i,\boldsymbol{c}_{-i}\right)+\left[\frac{m(m-1)}{2}-(n-1)\right]g(\boldsymbol{c})\end{aligned} \quad (21)$$

where $\mathbf{y}=\{y_1, y_2, \cdots, y_m\}^{\mathrm{T}}$ indicates the important variables in the input variables.

Substituting Eq.(21) into Eq.(6), and following the procedures in Eq.(7) to Eq.(9), the response function can be rewritten as

$$\hat{g}_{mixed}(\mathbf{x},\mathbf{y})=\prod_{i=1}^{m-1}\prod_{j=i+1}^{m}\left[g\left(y_i, y_j, \boldsymbol{c}_{-ij}\right)\right]\times \frac{\prod_{i=1}^{n}\left[g\left(x_i, \boldsymbol{c}_{-i}\right)\right]\times\left[g(\boldsymbol{c})\right]^{\left[\frac{m(m-1)}{2}-(n-1)\right]}}{\prod_{i=1}^{m}\left[g\left(y_i, \boldsymbol{c}_{-i}\right)\right]^{(m-1)}} \qquad (22)$$

This approximate model of original performance function is referred to as the hybrid multiplicative dimension-reduction method (H-M-DRM) in this work. It can be seen that the original input-output relation is approximated as the product of univariate and bivariate functions. When adopting the proposed method, the number of function evaluations in the moment calculation is $1+3n+3^2 m(m-1)/2$. In engineering applications, we usually set m equal to 2 or 3. Compared with the bivariate M-DRM in high-dimensional problems, the proposed method is apparently more efficient. But this method may not show its advantage in some cases, for example, when low-dimensional problems are tested or the high-order correlation terms of the high-dimensional problems can be ignored.

4. EXAMPLES

The numerical example is a ten-dimensional computational model, which is given as follows:

$$g(\boldsymbol{X})=3+4\sin X_1+2X_2X_3^3+X_4^3-X_5X_6-\frac{X_7X_8}{X_9X_{10}}$$

where $\boldsymbol{X}=(X_1, X_2, \cdots, X_{10})$ is the vector of random input variables following the normal distribution. The distribution information of the input variables is listed in *Table 1*.

The first four central moments of the response function are calculated by the univariate M-DRM, the proposed hybrid M-DRM and the Monte Carlo simulation method, respectively. According to the criteria for identifying important variables proposed in Section 3, we choose the two most important variables for bivariate decomposition, i.e. m=2. The calculation results are listed in the *Table 2*. The moments calculated by the univariate M-DRM and the proposed hybrid M-DRM method using three-point estimation, and the Monte Carlo simulation with 10^5 samples. The results of the Monte Carlo simulation are considered to be exact solutions and used to compare with the other two methods.

Table 1 Distribution parameters for input random variables

Variables	Distribution	Mean	Standard deviation
X_1	Normal	2	0.1
X_2	Normal	2	0.2
X_3	Normal	3	0.3
X_4	Normal	2	0.2
X_5	Normal	3	0.3
X_6	Normal	4	0.1
X_7	Normal	5	0.1
X_8	Normal	6	0.2
X_9	Normal	6	0.1
X_{10}	Normal	6	0.3

Table 2 First four central moments

Method	μ_g	σ_g	$\alpha_{3,g}$	$\alpha_{4,g}$
MCS method	113.4066	35.2197	0.6803	3.7810
Univariate M-DRM	113.2698	35.2098	0.6739	3.4020
Hybrid M-DRM	113.2698	35.2165	0.6769	3.4086

As can be seen from *Table 2*, the results obtained by univariate M-DRM method and the proposed method are all in good agreement with that estimated by the MCS procedure, but the proposed method is more accurate. In addition, for this high-dimensional problem involving 10 input variables, the number of functional evaluations of univariate M-DRM method is 31 and that of proposed method is 40, while that of MCS method is

10^5. Compared with MCS method, these two methods are very efficient. But if using bivariate M-DRM method, the number of functional evaluations will be 436, which is larger than the proposed method. That is to say, on the basis of guaranteeing accuracy, the proposed method is more accurate than the univariate M-DRM method and more efficient than the bivariate M-DRM method.

5. CONCLUSION

In this work, we propose a hybrid multiplicative dimension-reduction method. Then the proposed method and three-point estimation are employed to calculate statistical moments of the performance function. In the end, three examples with multivariate are studied to verify the accuracy and efficiency of the proposed method in comparison to the Monte Carlo simulation method. The results show that the proposed method can obtain more accurate results than the univariate M-DRM method, but the calculation amount only increase a little. And especially for high-dimensional problems where high-order terms make the significant contribution to the response function, the proposed method is more efficient than the bivariate M-DRM method. In summary, the proposed method provides an alternative and efficient method to analyze the high-dimensional structural reliability problems.

6. ACKNOWLEDGEMENT

The financial supports by the National Natural Science Foundation of China (Grant No. NSFC51608446, U1530122), the Natural Science Foundation of Shaanxi province (Grant No. 2017JQ1021), and the Fundamental Research Fund for Central Universities (No. 3102018zy011) are gratefully acknowledged.

7. REFERENCE

[1] Rabitz H, Aliş Ö F, Shorter J, Shim K. Efficient input-output model representations. Computer Physics Communications, 1999, 117(1-2): 11-20.

[2] Sobol I M. Theorems and examples on high dimensional model representation. Reliability Engineering & System Safety, 2003, 79(2): 187-193.

[3] Rahman S, Xu H. A univariate dimension-reduction method for multi-dimensional integration in stochastic mechanics. Probabilistic Engineering Mechanics, 2004, 19(4): 393-408.

[4] Zhao Y G, Ono T. New point estimates for probability moments. Journal of Engineering Mechanics, 2000, 126(4): 433-436.

[5] Xu H, Rahman S. A generalized dimension-reduction method for multidimensional integration in stochastic mechanics. International Journal for Numerical Methods in Engineering, 2004, 61(12): 1992-2019.

[6] Li G, Hu J, Wang S W, Georgopoulos P G, Schoendorf J, Rabitz H. Random sampling-high dimensional model representation (RS-HDMR) and orthogonality of its different order component functions. The Journal of Physical Chemistry A, 2006, 110(7): 2474-2485.

[7] Xu H, Rahman S. Decomposition methods for structural reliability analysis. Probabilistic Engineering Mechanics, 2005, 20(3): 239-250.

[8] Rahman S, Wei D. Design sensitivity and reliability-based structural optimization by univariate decomposition. Structural and Multidisciplinary Optimization, 2008, 35(3): 245-261.

[9] Rahman S. Stochastic sensitivity analysis by dimensional decomposition and score functions. Probabilistic Engineering Mechanics, 2009, 24(3): 278-287.

[10] Rahman S. A dimensional decomposition method for stochastic fracture mechanics. Engineering fracture mechanics, 2006, 73(15): 2093-2109.

[11] Zhang X, Pandey M D. Structural reliability analysis based on the concepts of entropy, fractional moment and dimensional reduction method. Structural Safety, 2013, 43: 28-40.

[12] Sobol I M. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Mathematics and computers in simulation, 2001, 55(1-3): 271-280.
[13] Shi Y, Lu Z, Cheng K, Zhou Y. Temporal and spatial multi-parameter dynamic reliability and global reliability sensitivity analysis based on the extreme value moments. Structural and Multidisciplinary Optimization, 2017, 56(1): 117-129.
[14] Zhang X. Efficient computational methods for structural reliability and global sensitivity analyses. University of waterloo. 2013.