Improved response surface method based on triple weighted regression

Jie Wu¹,², Jianguo Zhang¹,², Lingfei You¹,² and Nan Ye¹,²

¹ School of Reliability and Systems Engineering, Beihang University, Beijing, China
² Science and Technology on Reliability and Environmental Engineering Laboratory, Beihang University, Beijing, China

E-mail: zjg@buaa.edu.cn

Abstract. The response surface method (RSM) is frequently used to reduce the computational burden of structural reliability analysis. In this paper, an improved RSM is proposed. The response surface is fitted by the triple weighted regression technique, which integrates weighting systems to assign weight factors to each sampling point including (1) the absolute value of the limit state function (LSF), (2) the value of the joint probability density function and (3) the distance from the design point. Numerical examples are presented to demonstrate that the proposed method gives more accurate results than RSM for both probability of failure (PoF) and LSF evaluations.

1. Introduction

The main purpose of structural reliability analysis is to evaluate the PoF, denoted as \( P_f \), given by

\[
P_f = P\{g(x) \leq 0\} = \int_{g(x) \leq 0} f_X(x) \, dx
\]

in which \( f_X(x) \) is the joint probability density function for the vector of basic random variables \( x = [x_1, x_2, \ldots, x_n]^T \), which represents uncertain quantities (material properties, loads, geometrical data, etc.). \( g(x) \) is the LSF of the structure and \( g(x) \leq 0 \) represents the failure domain [1].

Among the existing structural reliability analysis methods, the first order reliability method (FORM) and the second order reliability method (SORM) suffer relatively poor performance in terms of accuracy. In some cases, FORM and SORM may have issues with convergence [2, 3]. Monte Carlo simulation (MCS) requires a large amount of computational resources, especially for low PoF, which limits its practical application [4, 5]. The RSM uses a polynomial function to approximate the actual LSF. The RSM has been widely used due to its simple principle and easy operation.

Bucher and Bourgund [6] used the quadratic polynomial function without cross terms to approximate the LSF, and applied the RSM to structural reliability analysis. Rajashekhar and Ellingwood [7] further developed the RSM through adaptive iteration. Kaymaz and McMahon [8] proposed a new RSM named ADAPRES which applied a weighted regression method in place of normal regression. Nguyen et al. [9] improved the weighted regression method by resorting to two weighting systems. Chen et al. [10] designed and constructed an improved RSM based on weighted

[1]
regression for the anti-slide reliability analysis of concrete gravity dam. Xu and Cao [11] developed an improved weighted nonlinear response surface method which improved the computational efficiency by combining the fractional weighting method with the exponential weighting method.

The most important part of applying weighted regression is to determine the appropriate weighting factors. Therefore, this paper proposes an enhanced weighting strategy which integrates three weighting systems to improve the accuracy of results.

2. Establishment of the improved response surface method based on triple weighted regression

The main features of the proposed method are weighting systems. In this section, the establishment of the weight matrix is presented first, followed by basic steps of the improved response surface method based on triple weighted regression.

2.1. Establishment of the weight matrix

The first weighting system is relating to the absolute values of the LSF $g(x)$. It aims to penalize the sampling points located far from the failure surface. In the $k$th iterative step, the first weight factor of the $j$th sampling point is stated as:

$$g_{best} = \min_{j=1}^{k \times (2n+1)} |g(x_j)|, \ g(x_j) \neq 0$$

in which $n$ is the number of random variables, $k \times (2n+1)$ is the number of existing sampling points.

$$w_{g_j} = \begin{cases} 
\frac{g_{best}}{|g(x_j)|}, & g(x_j) \neq 0 \\
1, & g(x_j) = 0 
\end{cases}$$

The second weighting system takes into account the values of the joint probability density function $f(x)$. In the $k$th iterative step, the second weight factor of the $j$th sampling point is stated as:

$$f_{best} = \max_{j=1}^{k \times (2n+1)} f(x_j)$$

$$w_{f_j} = \frac{f(x_j)}{f_{best}}$$

The third weighting system is expected to penalize the sampling points located far from the design point. In the $k$th iterative step, the third weight factor of the $j$th sampling point is stated as:

$$d_j = \|x^{(k)} - x_j\|$$

in which $d_j$ is the distance from the sampling point to the design point $x^{(k)}$.

$$d_{max} = \max_{j=1}^{k \times (2n+1)} d_j$$

$$s_j = \frac{d_j}{d_{max}}$$
Then, the average of the three weighting factors is taken to obtain the mixed weight factor $w_j$ at each sampling point, as

$$w_j = \frac{w_{sj} + w_{sj} + w_{sj}}{3}$$

(10)

The weight matrix $W$ is constructed by taking the mixed weight factor of each sampling point as the diagonal element:

$$W = \begin{pmatrix}
w_1 & & & \\
& w_2 & & \\
& & \ddots & \\
& & & w_{k \times (2n+1)}
\end{pmatrix}$$

(11)

### 2.2. Basic steps of the improved response surface method based on triple weighted regression

Step 1: Adopt the traditional RSM in the first iteration ($k = 1$).

Step 2: Choose the sampling points in the $k^{th}$ iterative step:

$$\mathbf{x}^{(k-1)} = (x_1^{(k-1)}, x_2^{(k-1)}, \ldots, x_n^{(k-1)})$$

(12)

$$(x_1^{(k-1)}, x_2^{(k-1)}, \ldots, x_n^{(k-1)} \pm f^{(k)}(\sigma_1, \ldots, \sigma_n), x_n^{(k-1)}) \quad (i = 1, 2, \ldots, n)$$

(13)

where $\mathbf{x}^{(k-1)}$ is the center point calculated in the previous iteration. $f^{(k)}$ is the interpolation coefficient, generally taken $1 \sim 3$.

Step 3: Evaluate the actual LSF values $g$ for the sampling points.

Step 4: Construct the weight matrix $W$ with $k \times (2n+1)$ sampling points.

Step 5: Apply a weighted regression to get the response surface function:

$$b = (A^TWA)^{-1}A^Tg$$

(14)

in which $b$ is the undetermined coefficient in the response surface function, $A$ is the design matrix. For quadratic response surfaces without cross terms, the design matrix $A$ is consisting of $m = k \times (2n+1)$ sampling points:

$$A = \begin{pmatrix}
1 & x_{11} & \cdots & x_{1m} & x_{11}^2 & \cdots & x_{1m}^2 \\
1 & x_{21} & \cdots & x_{2n} & x_{21}^2 & \cdots & x_{2m}^2 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\
1 & x_{n1} & \cdots & x_{nm} & x_{n1}^2 & \cdots & x_{nm}^2
\end{pmatrix}$$

(15)

Step 6: Obtain the design point $\mathbf{x}^{(k)}$ and the reliability index $\beta^{(k)}$ with FORM/SORM.

Step 7: Calculate a new sampling center $\mathbf{x}^{(k)}$ as
\[ x^{(k)} = \mu_x + \frac{g(\mu_x)}{g(\mu_x) - g(x^{(k)})}(x^{(k)} - \mu_x) \]  

(16)

in which \( \mu_x \) is the mean point.

Step 8: Check the convergence if \( k \neq 1 \). If convergence is achieved, stop the iterative procedure; otherwise, \( k = k + 1 \), continue to Step 2.

The convergence in the reliability index as

\[ \frac{\beta^{(k)} - \beta^{(k-1)}}{\beta^{(k)}} < \varepsilon \]  

(17)

in which \( \varepsilon \) is the accuracy chosen by the user.

3. Examples

3.1. Example 1: a nonlinear LSF

A hypothetical LSF with two independent standard normal variables is considered:

\[ g(x) = \exp(0.2x_1 + 6.2) - \exp(0.47x_2 + 5.0) \]  

(18)

Table 1 contains numerical results obtained from various methods including MCS (N=1000000).

| Method       | Reliability index | PoF   | Error(%) |
|--------------|-------------------|-------|----------|
| MCS          | 2.3510            | 0.00937| —        |
| RSM          | 2.3706            | 0.00888| 5.229    |
| Proposed method | 2.3567       | 0.00922| 1.601    |

As seen from the numerical results in Table 1, the proposed method can obtain a more accurate estimation of PoF as well as reliability index than RSM by applying the triple weighted regression technique. For this example, Figure 1 shows the actual LSF and obtained response surface functions using the traditional RSM and proposed method. As seen from Figure 1, the proposed method approximates the LSF better than RSM over a large region around the design point.
Figure 1. Comparison of RSM and the proposed method.

3.2. Example 2: a cantilever beam
A cantilever beam subjected to a concentrated force is considered (as shown in Figure 2). \( L \) (length) is considered as deterministic with the values of 5m. \( E \) (modulus of elasticity), \( I \) (moment of inertia of the cross section) and \( P \) (concentrated force) are independent normal random variables. The properties of random variables are tabulated in Table 2.

![Figure 2. A cantilever beam subjected to a concentrated force.](image)

Table 2. Properties of random variables -- Example 2.

| Variable | Mean       | Standard deviation | Distribution |
|----------|------------|--------------------|--------------|
| \( E \) (kN/m\(^2\)) | \( 2 \times 10^7 \) | \( 0.5 \times 10^7 \) | Normal       |
| \( I \) (m\(^4\)) | \( 10^{-4} \) | \( 0.1 \times 10^{-4} \) | Normal       |
| \( P \) (kN) | 8          | 2.5                | Normal       |

Considering the maximum deformation of the cantilever beam, the LSF is established as follows:

\[
g(E, I, P) = EI - 78.125P
\]

Table 3 contains numerical results obtained from various methods including MCS (N=1000000).

Table 3. Comparison of numerical results -- Example 2.

| Method      | Reliability index | PoF     | Error(%) |
|-------------|-------------------|---------|----------|
| MCS         | 2.5157            | 0.00594 | —        |
| RSM         | 2.5308            | 0.00569 | 4.209    |
| Proposed method | 2.5133         | 0.00598 | 0.673    |

The numerical results of reliability analyses for this example are listed in Table 3. It is seen from the numerical results that the improvement of RSM leads to an increase in the accuracy of estimated PoF, just like the previous example. The advantage in terms of accuracy is related to the weighting systems which assign weight factors to each sampling point.

4. Conclusion
In this paper, an improvement of the RSM which includes weighting systems to assign weight factors to each sampling point has been proposed. For the first iteration, normal regression is chosen and weighted regression is applied for the following iterations. To get a better response surface, three weight factors are considered in the weighted regression and all known sampling points are utilized in the iterative process. From the example 1 and example 2, relating to nonlinear LSFs, it is found
that, in comparison to the traditional RSM, the triple weighted regression technique for fitting response surface guarantees accurate results in a better approximate response surface function.

References
[1] Ditlevsen O and Madsen HO 1996 Structural reliability methods (Chichester: Wiley)
[2] Wang L and Grandhi RV 1996 Comput Struct 59 1139–48
[3] Thoft Christensen P and Baker MJ 1982 Structural reliability theory and its applications (Berlin: Springer-Verlag)
[4] Bjerager P 1988 J Eng Mech ASCE 114 1285–302
[5] Rubinstein RY and Kroese DP 2016 Simulations and Monte-Carlo method. Third Edition (Wiley)
[6] Bucher C and Bourgund U 1990 Struct Saf 7 57–66
[7] Rajashekar M R and Ellingwood B R 1993 Struct Saf 12 205–20
[8] Kaymaz I and McMahon C A 2005 Probabil Eng Mech 20 11–17
[9] Nguyen XS, Sellier A, Duprâte F and Pons G 2009 Probab Eng Mech 24 135–43
[10] Chen J Y, Xu Q, Li J and Fan S L 2010 J Zhejiang Univ Sci A 11 432–9
[11] Xu M R and Cao Z J 2018 2nd Int. Conf. on Mechanical, Electronic, Control and Automation Engineering (Qingdao) vol 149 (Paris/Atlantis Press) p 195