Structural Safety 73 (2018) 42–53

Contents lists available at ScienceDirect

Structural Safety

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

A sequential surrogate method for reliability analysis based on radial basis function

Xu Li [a], Chunlin Gong [a], Liangxian Gu [a], Wenkun Gao [a], Zhao Jing [b,*], Hua Su [a]

[a] School of Astronautics, Northwestern Polytechnical University, Xi'an, PR China
[b] Department of Mechanics, Huazhong University of Science and Technology, Wuhan, PR China

ARTICLE INFO

Article history:
Received 27 December 2016
Received in revised form 18 October 2017
Accepted 19 February 2018

Keywords:
Reliability analysis
Radial basis function
Surrogate model
Limit state function

ABSTRACT

A radial basis function (RBF) based sequential surrogate reliability method (SSRM) is proposed, in which a special optimization problem is solved to update the surrogate model of the limit state function (LSF) iteratively. The objective of the optimization problem is to find a new point to maximize the probability density function (PDF), subject to the constraints that the new point is on the approximated LSF and the minimum distance to the existing points is greater than or equal to a given distance. By updating the surrogate model with the new points, the surrogate model of LSF becomes more and more accurate in the important region with a high failure probability and on the LSF boundary. Moreover, the accuracy of the unimportant region is further improved within the iteration due to the minimum distance constraint. SSRM takes advantage of the information of PDF and LSF to capture the failure features, which decrease the samples of implicit LSF defined by expensive finite element analysis. Several numerical examples show that SSRM improves the accuracy of the surrogate model in the important region around the failure boundary with a small number of samples and has a better adaptability to the nonlinear LSF, hence increases the accuracy and efficiency of the reliability analysis.

1. Introduction

In engineering design, when uncertainties are involved, the failure probability P_f of the limit state function (LSF) $g(\mathbf{x})$ should be examined. P_f is essentially a multi-dimensional integral formulated as

$$P_f = \int\int\cdots\int_D \mathrm{p}(\mathbf{x})\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_m \tag{1}$$

where D is the failure region defined as $D = \{\mathbf{x}|g(\mathbf{x}) \leqslant 0, \mathbf{x} \in \mathbb{R}^m\}$, and $\mathrm{p}(\mathbf{x})$ is the joint probability function. The numerical integral and direct Monte Carlo Simulation (MCS) are difficult for a complex system with implicit time-consuming analysis models. To calculate the integral with the original LSF faces enormous computational challenges [1–2]. Therefore, different approximations of the LSF are adopted to improve the computational efficiency of reliability analysis.

The mean value method (MVM) [3–4] performs a first-order Taylor expansion of the LSF at the mean point (as shown in Fig. 1), in which the LSF is assumed to be normal distribution. The MVM is one of the most efficient reliability methods. However, it requires the independent input variables obey normal distribution, which is difficult to be satisfied in practical problems. Moreover, the approximation error increases with the increase of the nonlinearity. Therefore, this method is appropriate for fast estimation of the structural reliability with low nonlinearity.

The first-order reliability method (FORM) [5–8] transforms random variables with different distributions into the same standard normal space **U** by Rosenblatt transformation [9] and then performs a first-order Taylor expansion at the most probable point (MPP) which has the maximum failure probability on the LSF (as shown in Fig. 1). Eventually, the normal distribution parameters (mean value and standard deviation) of LSF are figured out with the gradient of the approximation function, and then the failure probability is obtained analytically. Compared with the MVM, the FORM does not require the input variables to obey normal distribution and has a higher accuracy with the LSF expanded at the MPP. However, the optimization with an equality constraint to find the MPP increases the number of the LSF evaluations. Moreover, it increases the nonlinearity of the LSF $g(\mathbf{x})$ during the Rosenblatt transformation, thus the approximation error is large when the nonlinearity of LSF is high [10–11].

* Corresponding author.
E-mail address: jingzhao.behunters@mail.google.com (Z. Jing).

https://doi.org/10.1016/j.strusafe.2018.02.005
0167-4730/

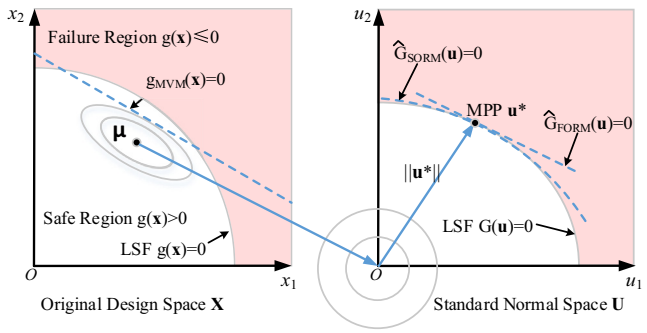

Fig. 1. Approximations of LSF in **X** and **U** space.

The second-order reliability method (SORM) [7,12–13] is similar to the FORM. First, the input random variables are transformed into the standard normal space **U** to get the LSF G(**u**). Second, G(**u**) is expanded with second-order Taylor expansion at the MPP to obtain a quadratic hypersurface. Finally, the reliability of the approximate model is analyzed with analytical methods [10]. The SORM has a higher nonlinear adaptability than the FORM, but it needs to perform the time-consuming second-order gradients. Moreover, the adaptability to nonlinear boundary is still limited [10–11].

The response surface is a different commonly used reliability analysis method [14–16]. By sampling in the neighborhood of the design point, the local approximation model of LSF is constructed. Since the response surface model evaluation time is far less than that of the original LSF, it is possible to use the approximate model for Mont Carlo Simulation (MCS) or Importance Sampling (IS) [17–19]. However, since the accuracy of the response surface method is poor with numbered samples, the complex structural behavior might not be captured. When considering factors such as the failure boundary and probability density, more samples are required to improve the local accuracy. Therefore, iterative response surface reliability analysis methods are proposed [20–30], which increases the samples in the important region near the MPP, and gradually improves the accuracy of the approximation model. This type of iterative method is also known adaptive or active learning method [31–32,45]. In general, the response surface approximation model can be replaced by other surrogate model (also known as meto-model) techniques such as radial basis function (RBF), Kriging, support vector regression (SVR), artificial neutral net (ANN), etc. [18,33–44]. The existing methods converge in a local region after increasing the sample density of the important region in some degree, but the accuracy does not increase in the less important region.

In order to make full use of the information of the added samples to improve the approximate accuracy of the important and less important region, this paper proposes a sequential surrogate reliability method (SSRM) based on RBF. By adding points sequentially to the surrogate model, the failure features in the important region and on the boundary of the LSF are captured, and the failure probability is obtained with MCS by using the surrogate model. SSRM does not need to use the original LSF to search for MPP directly, but rather becomes close to the important area near MPP in the process of adding points, thus reducing the number of sample evaluations and avoiding the failure to find the optimal solution. Meanwhile, the SSRM method makes a trade-off between precision and computational cost.

The remainder of this paper is structured as follows. Section 2 introduces the surrogate model technology used in SSRM; Section 3 describes the implementation process of SSRM, and analyzes its characteristics; Section 4 verifies the effectiveness and efficiency of SSRM with several numerical examples; conclusions are given in Section 5.

2. Surrogate model

2.1 Construction of surrogate model

The goal of the surrogate model is to construct a prediction model of a complex or unknown model with the observation samples. Instinctually, surrogate model is an interpolation or regression model, which is also a branch of machine learning [47]. The common surrogate models include polynomial response surface method (PRSM) [48–49], radial basis function (RBF) [48–49], Kriging [50–52], support vector regression (SVR) [50–52] and artificial neutral net (ANN) [53]. Since RBF has good nonlinear adaptability and is easy to implement, this paper constructs the sequential surrogate model with RBF. Assume that the observation samples are presented as

$$S = \{(\mathbf{x}_i, y_i) | i = 1, 2, \ldots, n\} \tag{2}$$

denoted by a matrix form

$$\begin{aligned} \mathbf{X} &= [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n]^{\mathrm{T}} \\ \mathbf{y} &= [y_1, y_2, \ldots, y_n]^{\mathrm{T}} \end{aligned} \tag{3}$$

where n is the sample size, $\mathbf{X}$ denotes the input sample matrix, $\mathbf{y}$ is the output sample vector. $\mathbf{x}$ is an m-dimensional design variable. RBF uses a series of linear combinations of radial basis functions to approximate the expensive limit state function, which can be formulated as

$$\hat{\mathrm{y}}(\mathbf{x}) = \sum_{i=1}^{n} \beta_i \mathrm{f}(\|\mathbf{x} - \mathbf{x}_i\|) = \mathbf{f}(\mathbf{x})^{\mathrm{T}} \boldsymbol{\beta} \tag{4}$$

where $\hat{\mathrm{y}}(\mathbf{x})$ denotes the predictive response at point $\mathbf{x}$, β_i is the ith component of the radial base coefficient vector $\boldsymbol{\beta}$, and $\mathrm{f}(\|\mathbf{x} - \mathbf{x}_i\|)$ (see Table 1) is the ith component of the radial basis function vector $\mathbf{f}(\mathbf{x})$. As shown in Table 1, $r = \|\mathbf{x} - \mathbf{x}_i\|$ is the Euclidian distance between two samples, and c is the shape parameter.

Substitute the samples of Eq. (2) into Eq. (4),

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \mathrm{f}(r_{11}) & \mathrm{f}(r_{12}) & \cdots & \mathrm{f}(r_{1n}) \\ \mathrm{f}(r_{21}) & \mathrm{f}(r_{22}) & \cdots & \mathrm{f}(r_{2n}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{f}(r_{n1}) & \mathrm{f}(r_{n2}) & \cdots & \mathrm{f}(r_{nn}) \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} \tag{5}$$

Rewritten as a matrix form

$$\mathbf{y} = \mathbf{F}\boldsymbol{\beta} \tag{6}$$

As the samples are different with each other, $\mathbf{F} \in \mathbb{R}^{n \times n}$ is a non-singular matrix, and Eq. (6) has a unique solution $\boldsymbol{\beta} = \mathbf{F}^{-1}\mathbf{y}$. Thus the prediction model is given by

$$\hat{\mathrm{y}}(\mathbf{x}) = \mathbf{f}(\mathbf{x})^{\mathrm{T}} \mathbf{F}^{-1} \mathbf{y} \tag{7}$$

where $\mathbf{f}(\mathbf{x})$ is related to the prediction point $\mathbf{x}$ and sample input matrix $\mathbf{X}$; $\mathbf{F}^{-1}\mathbf{y}$ is only related to $\mathbf{X}$ and $\mathbf{y}$. For a new prediction sample $\mathbf{x}$, $\mathbf{f}(\mathbf{x})$ is calculated one time to get its predicted value$\hat{\mathrm{y}}(\mathbf{x})$.

Table 1
Radial basis functions.

Type	Function form $\mathrm{f}(r)$
Gaussian	$\exp(-cr^2)$
Inverse Multiquadric	$(1+cr^2)^{-1/2}$
Thin plate spline	$r^2\log(1+cr^2)$