ABSTRACT: Currently most available methods for computing structural reliability are limited to low dimensional problems, and will inevitably encounter many inherent difficulties when applied to medium to high-dimensional problems. In this paper, a new hybrid algorithm is proposed aiming at dealing with medium to high-dimensional structural reliability problems effectively. The main idea of the algorithm is to construct an appropriate auxiliary domain by means of SVM that is significantly overlapped with the target failure domain, and sequentially to convert the original problem of calculating the target failure probability of very small value to an easier one of calculating two conditional failure probabilities of relatively larger value. Markov chain Monte Carlo algorithm is adopted to construct the samples that obey the conditional distribution in the target failure domain. Two numerical examples are given to preliminarily investigate the feasibility and effectiveness of the proposed algorithm when dealing with medium to high-dimensional reliability problems.

KEYWORDS: High-dimensional reliability; Auxiliary domain; SVM; Markov chain Monte Carlo; Conditional failure probability.

1 INTRODUCTION

Civil engineering structures are probable to confront various kinds of random effects, e.g. wind load and seismic ground motion, meanwhile, parameters of the structure itself are inevitably uncertain, e.g. concrete strength and member dimension. It is of great importance to calculate the structural reliability, since inaccurate reliability result may lead to unsafe or over-conservative structural design. At present, available structural reliability methods can roughly be classified into moment method, surrogate model method, probability density evolution method and stochastic simulation method.

Moment method mainly includes First Order Reliability Method (FORM), Second Order Reliability Method (SORM) and high order moment method. FORM can yield results of acceptable accuracy for many practical problems with relatively low computational amount, and is currently the most widely used method. SORM is an improvement of FORM by taking both the first order and second order terms of the Taylor expansion of the performance function at the design point, and will often give a more accurate result at the cost of somewhat laborious calculation of the Hessian matrix. Both FORM and SORM need the sensitivity information of the problem to find the design point, which can be difficult in some cases. High order moment method makes use of not only the mean value and variance but also the skewness and kurtosis of the performance function to make the failure probability result more reliable. Meanwhile, sensitivity information is no longer needed. Despite these merits, approximation adopted in the representation of the performance function may bring in inaccuracies.

The main idea of the surrogate model method is to approximate the real implicit performance function by an explicit analytical expression with a limited number of function evaluations. So far, widely used surrogate models in structural reliability field include quadratic polynomial response surface, Artificial Neural Network (ANN) model, Kriging model and Support Vector Machine (SVM) model, among which SVM is adopted as the tool to construct the auxiliary domain in this paper due to its excellent learning and generalization ability.

Probability density evolution method is a kind of methods that aim to obtain the probability density of interested state variables during the whole response process, among which Generalized Probability Density Evolution Equation (GPDEE) proposed by Li and Chen [2][3] is proved to be suitable for both static and dynamic structural reliability problems. Although GPDEE has achieved great success in many problems, representative points selection strategy in very high-dimensional probability space is still a bottleneck for its further application.

Stochastic simulation method is the most versatile and powerful approach in stochastic structural analysis, and can be applied to all kinds of structural reliability problems theoretically. In addition, while moment method and surrogate method fail to deal with high-dimensional problems due to the so-called ‘curse of dimensionality’, some smart simulation algorithms are still qualified. Direct Monte Carlo (DMC) simulation is the simplest and the most robust, but when applied to practical engineering problems with extremely small values of failure probability, its computational cost is unacceptable. Various kinds of variance reduction techniques are employed to increase the efficiency of simulation, among which Importance Sampling (IS) is the most popular.

In this paper, a new hybrid algorithm called SVM-based Auxiliary Domain Method (SVM-ADM) is proposed which is inspired by the earlier work of Katafygiotis et al. [9] who come up with the idea of ADM and made use of the statistical equivalent linearization system as an auxiliary system.
However, the statistical equivalent linearization system is no longer a good choice for problems with medium to high degree of nonlinearity. In view of this point, SVM is employed to construct the auxiliary domain. In section 2, a brief introduction of SVM will be given followed by a very efficient Markov Chain Monte Carlo (MCMC) algorithm in section 3. Detail of the proposed algorithm is described in section 4, and two numerical examples are given to preliminarily investigate the feasibility and effectiveness of the proposed algorithm in section 5. The whole paper is ended up with the conclusion in section 6.

2 SUPPORT VECTOR MACHINE

An SVM is a kind of machine learning algorithm used for data classification in pattern recognition and is based on the principle of Structural Risk Minimization (RSM) other than the traditional Empirical Risk Minimization (ERM) [10]. It is able to construct a complex decision function that optimally separates the data into different classes even in high-dimensional spaces. In spite of its ability to deal with multi-classification problems, binary classification is enough due to the nature of reliability problems. In the following of this section, fundamental theory of both linear and nonlinear SVM is given.

2.1 Linear SVM

For a given training point set \{x_i\} ∈ R^n, i = 1, 2, ..., l, each point corresponds to a class c_i ∈ {-1, 1}. Assume that these points are linearly separable, and the function expression for the classification hyperplane is \(f(x) = \langle w, x \rangle + b = 0\). By normalization with \(1/||w||\), the classification margin boundary can be expressed by \(\langle w, x \rangle + b = \pm 1\) as shown in Fig. 1.

![Fig. 1 Linear SVM (2D)](image)

According to RSM, the hyperplane that corresponds to the maximum classification margin is the optimal classification hyperplane. Thus the problem can be equivalently expressed as follows:

\[
\begin{align*}
\text{minimize} & : & \quad \tau (w) = \frac{||w||}{2} \\
\text{subject to} & : & \quad c_i (\langle w, x_i \rangle + b) \geq \pm 1, i = 1, 2, \cdots, l
\end{align*}
\]

(1)

(2)

By introducing the Lagrangian multipliers \(\alpha \geq 0\) and conducting some mathematical derivation, we can finally get the following convex quadratic optimization problem in a dual form:

\[
\begin{align*}
\text{minimize} & : & \quad W(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \alpha_i \alpha_j c_i \langle x_i, x_j \rangle c_j \\
\text{subject to} & : & \quad \alpha_i \geq 0, \quad \sum_{i=1}^{l} \alpha_i c_i = 0
\end{align*}
\]

(3)

(4)

After solving the above optimization problem, we have the following classification hyperplane:

\[
g(x) = \text{sgn} \left[ f(x) \right] = \text{sgn} \left[ \sum_{i=1}^{l} \alpha_i c_i \langle x_i, x \rangle + b \right]
\]

(5)

According to the famous Karush-Kuhn-Tucker condition in the optimization theory, the following equation holds:

\[
\alpha_i [c_i (\langle w, x_i \rangle + b) - 1] = 0, i = 1, 2, \cdots, l
\]

(6)

From equation (6), it is easy to see that only those Lagrangian multipliers that correspond to the training points on the classification margin boundary are nonzero. These points, accordingly called support vectors, account only for a very small proportion of the total training points.

2.2 Nonlinear SVM

When the training points are not linearly separable in the input space, we can map them into a high-dimensional feature space by means of nonlinear mapping where they become linearly separable. In such a case, equation (5) becomes the following :

\[
g(x) = \text{sgn} \left[ f(x) \right] = \text{sgn} \left[ \sum_{i=1}^{l} \alpha_i c_i \langle \phi(x_i), \phi(x) \rangle + b \right]
\]

(7)

According to the Mercer theorem, the inner product in equation (7) can be calculated simply by a kernel function as follows:

\[
K(x, x') = \langle \phi(x), \phi(x') \rangle
\]

(8)

There are various forms of kernel functions in SVM, such as polynomial, Gaussian, Sigmoid, etc. Gaussian kernel function is adopted in this paper for its numerical convenience.

3 SINGLE-COMPONENT ADAPTIVE METROPOLIS ALGORITHM

MCMC is a way to produce samples that obey a specific probability distribution. Its basic idea is to construct an ergodic Markov chain, the stationary distribution of which is just the target one. Metropolis algorithm [11] is the most basic one of all the MCMC algorithms. In Metropolis algorithm, we need to select a symmetric proposal function which is used to predict the candidate point x' of the next step given the sample point x_t of the present step. Assume that π(x) is the target
distribution, so the probability to accept \( x^* \) as the sample point \( x_{t+1} \) of the next step is:

\[
\alpha(x_t, x^*) = \min\left(1, \frac{\pi(x^*)}{\pi(x_t)}\right)
\]  

(9)

Despite its feasibility, Metropolis algorithm still exposes two obvious shortcomings when applied to specific problems. One is related to the choice of the proposal function, which contributes significantly to the efficiency of the algorithm but is often selected empirically. An improper proposal function may even result in extremely slow convergence rate. The other is inapplicability of updating the sample point of the next step in a global way in high-dimensional problems since in such cases global update will lead to rather low acceptance probability and thus probably making the whole Markov chain trapped in a single sample point.

In order to overcome the first shortcoming, an adaptive sampling strategy [12] is often adopted. Gaussian density function is selected as the proposal function, the covariance matrix of which is updated step-by-step by use of the information from the past sample points. In such a way, the proposal function will be updated in time according to the information of the target distribution function, thus improving the sampling efficiency as far as possible.

As for the other issue, a single-component sampling strategy [5] proves to be an effective resolving approach, which decomposes a high-dimensional sampling problem into multiple sequential one-dimensional sampling problems. Since only a single component of the sample point is updated each time, the acceptance probability for the candidate point is largely increased, and accordingly the sampling efficiency is improved.

A single-component adaptive Metropolis algorithm is a simple combination of both the adaptive sampling strategy and the single-component sampling strategy and is used in this paper to construct the conditional sample points in the target failure domain.

4 INTRODUCTION TO SVM-ADM

In 2007, Katafygiotis et al. [9] proposed the ADM, basic idea of which is to construct an auxiliary domain of the target failure domain and then to decompose the calculation of the target failure probability into an equivalent expression as follows:

\[
P_f = \Pr(\Omega_f) = \Pr(\Omega_A) \Pr(\Omega_x | \Omega_A) / \Pr(\Omega_x | \Omega_t)
\]  

(10)

where, \( \Omega_f \), \( \Omega_A \) and \( \Omega_t \) stand for the target failure domain, the auxiliary domain and the intersection domain respectively, see Fig. 2.

![Fig. 2 Auxiliary domain (2D)](image)

From equation (10) we can see that the target failure probability is decomposed into three parts: \( \Pr(\Omega_f) \), \( \Pr(\Omega_f | \Omega_A) \) and \( \Pr(\Omega_f | \Omega_t) \). Since the auxiliary boundary always has an analytical expression, the calculation of the term \( \Pr(\Omega_f | \Omega_A) \) will not cost too much computational time. Meanwhile, as long as the auxiliary domain overlaps well with the target failure domain, values of \( \Pr(\Omega_f | \Omega_A) \) and \( \Pr(\Omega_f | \Omega_t) \) are expected at least to be of the magnitude of 0.1, and are accordingly much easier to calculate than \( \Pr(\Omega_f) \). In some sense, \( \Pr(\Omega_f | \Omega_A) / \Pr(\Omega_f | \Omega_t) \) can be seen as a correction term to the approximate failure probability \( \Pr(\Omega_f) \). In [9], the statistical equivalent linearization technique is adopted to construct an auxiliary system to the original nonlinear structural system. However, this choice does work only in a limited type of problems that are slightly nonlinear, and for many practical highly nonlinear problems there seems to be no analytical methods to construct the corresponding auxiliary system any more.

On the other hand, SVM was introduced into the field of structural reliability by Hurtado [13] and became widely used since then due to its powerful learning and generalization ability with limited samples. Despite the inspiring success it has achieved, SVM seems to be unable to construct a satisfactory surrogate model for highly nonlinear problems with a limited number of training points, especially in high-dimensional cases.

Although SVM is unable to construct an accurate surrogate model for complex problems, why not step back and use it merely to construct an auxiliary domain? In other words, combination of ADM and SVM seems to be a complementary and promising scheme. Specific steps of this idea are listed below:

1) To generate a batch of initial training points to obtain a coarse SVM model;
2) To update the SVM model sequentially by the active learning strategy until a specific number of points are reached or there is no point in the margin of the latest SVM model;
3) To construct a set of conditional samples \( S_T \) in the target failure domain by use of the single-component adaptive Metropolis algorithm and to compute the term \( \Pr(\Omega_f | \Omega_t) \);
4) To compute the term \( \Pr(\Omega_A) \) corresponding to the final auxiliary domain by DMC, and meanwhile to collect the samples that fall into the auxiliary domain as a set of conditional samples \( S_A \).
5) To compute the term \( \text{Pr}(\Omega_I|\Omega) \) with \( S_A \);
6) To compute the target failure probability \( P_I \) by equation (10).

In step 1, a direct and convenient way to construct training points is to sample in the whole space according to the standard normal distribution and then to evaluate the corresponding functions for classification. However, the efficiency of such a way is rather low since generally a large number of sampling points are required before one of them eventually falls into the target failure domain. In order to generate as many training points near the failure boundary as possible, Yuan et al. [14] proposed to train the initial SVM model with Markov chain candidate points. Therefore, step 1 and step 3 can be conducted in a unified way: a Markov chain is generated and divided into two parts, the candidate points of the former part is used to train the initial SVM model while the latter part is taken as the conditional samples in the target failure domain. An additional merit of this strategy is that stationarity of the latter part of the Markov chain can be guaranteed since the former part is believed to be longer than the burn-in period of the whole Markov chain. In order to refine the SVM model with as less subsequent training points as possible, an active learning strategy is employed. According to section 2, an SVM model is completely determined by support vectors rather than other training points, which means that not all training points contribute equally to the SVM model. Studies show that sample points in the classification margin lead to a best training result. In this paper, a large number of points that center at the central point of the preceding Markov chain are generated according to the standard normal distribution, and are used as a database of the candidate points for subsequent active learning. In each round of step 2, a fixed number of candidate points that fall into the margin domain of the latest SVM model are selected as a new set of training points, and the remaining points are used as a new database of the candidate points for the next round. The cycle is terminated when the required number of total active learning points is reached or there is no point in the margin of the latest SVM model. It is worth noting that a key problem of the algorithm is to pick up a starting point in the target failure domain that initiates the Markov chain. For practical problems, engineering experience may be helpful. But if it doesn’t work, some gradient-based search methods may have to be resorted to.

5 NUMERICAL EXAMPLES

In order to verify the feasibility and effectiveness of the proposed algorithm, two numerical examples are given below.

5.1 50-dimensional problem

In this example, results of the SVM-based surrogate model method and SVM-ADM are compared. The performance function reads:

\[
g(x) = \sum_{i=1}^{50} x_i^4 - 33 \tag{11}\]

where, \( X_i, i=1, ..., 50 \) are i.i.d. normal random variables. DMC with 1000,000 samples are conducted for 50 times, and the corresponding mean value of failure probability \( P_f=0.0014 \) is taken as the accurate result of the problem. In the SVM-based surrogate model method, 4,000 candidate points are used to train the SVM surrogate model, while in SVM-ADM, the same number of function calls are divided into four parts: 850 candidate points for initial SVM training, 150 points for active learning, 1,500 conditional sample points in the target failure domain and 1,500 conditional sample points in the auxiliary failure domain. Both methods are run for 10 times respectively, and the corresponding results are listed in Table 1 below.

<table>
<thead>
<tr>
<th>SVM</th>
<th>SVM-ADM</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_f )</td>
<td>( \text{Pr}(\Omega_I</td>
</tr>
<tr>
<td>1</td>
<td>0.0009</td>
</tr>
<tr>
<td>2</td>
<td>0.0009</td>
</tr>
<tr>
<td>3</td>
<td>0.0008</td>
</tr>
<tr>
<td>4</td>
<td>0.0010</td>
</tr>
<tr>
<td>5</td>
<td>0.0011</td>
</tr>
<tr>
<td>6</td>
<td>0.0009</td>
</tr>
<tr>
<td>7</td>
<td>0.0009</td>
</tr>
<tr>
<td>8</td>
<td>0.0010</td>
</tr>
<tr>
<td>9</td>
<td>0.0007</td>
</tr>
<tr>
<td>10</td>
<td>0.0008</td>
</tr>
<tr>
<td>Mean</td>
<td>0.0009</td>
</tr>
<tr>
<td>Emp. c.o.v.</td>
<td>0.1283</td>
</tr>
</tbody>
</table>
Two conclusions can be drawn from Table 1. Firstly, although 850 candidate points for initial SVM training and 150 points for active learning cannot give an acceptably accurate surrogate model, they provide a satisfactory auxiliary failure domain that overlaps in a large portion with the target failure domain, thus making the calculation of two conditional probabilities very easy; Secondly, with the same total number of function calls, SVM-based surrogate model method gives a rather biased result while SVM-ADM gives an almost unbiased result.

5.2 40-dimensional problem

A numerical example from Bourinet et al. [15] is used here, aiming at comparing the effectiveness of the proposed algorithm with FORM, SS and 2SMART when dealing with a medium-dimensional problem. The performance function reads:

\[ g(x) = (n + a\sigma\sqrt{n}) - \sum_{i=1}^{n} x_i \]  

(12)

where \( X_i, i=1, \ldots, n \) are i.i.d. lognormal random variables, with unit means and standard deviations equal to 0.2, the parameter \( a \) is taken as 3. Only the case \( n=40 \) is studied here. DMC with 500,000 samples are conducted for 50 times, and the corresponding mean value of failure probability \( P_f=0.0020 \) is taken as the accurate result of the problem. SVM-ADM is run for 20 times, and the comparison of the corresponding result with those from [15] is illustrated in Table 2 below.

Table 2 Comparison of the results by FORM, SS, 2SMART and SVM-ADM

<table>
<thead>
<tr>
<th>n</th>
<th>FORM</th>
<th>SS ( \times 500 )</th>
<th>2SMART ( \times 20 )</th>
<th>SVM-ADM ( \times 20 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( P_f )</td>
<td>( N_{\text{fail}} ) (( N_{\text{max}} ))</td>
<td>( N_{\text{fail}} ) (emp. c.o.v.)</td>
<td>( N_{\text{fail}} ) (emp. c.o.v.)</td>
</tr>
<tr>
<td>40</td>
<td>2.15\times10^{-4}</td>
<td>0.06</td>
<td>1.95\times10^{-3}</td>
<td>1.90\times10^{-3}</td>
</tr>
<tr>
<td></td>
<td>135(3)</td>
<td>3729(0.21)</td>
<td>3729(0.28)</td>
<td>3729(0.10)</td>
</tr>
</tbody>
</table>

It can be seen from table 2 that FORM fails to predict the failure probability well in this case. SVM-ADM yields an almost unbiased result, and performs better than SS in terms of variation of the result. Although the empirical coefficient of variation of the result by 2SMART is smaller than that by SVM-ADM, it is the author’s belief that the efficiency of the proposed algorithm will increase as long as some more delicately designed active learning strategy is adopted, such as in 2SMART.

6 CONCLUSIONS

The proposed algorithm in this paper takes advantage of both the excellent fitting and generalization ability of SVM and the characteristic of ADM to decompose the target failure probability of very small value into conditional failure probabilities of larger values, thus avoiding the dependence on the design point and being possibly suitable to deal with medium to high-dimensional reliability problems. Two numerical examples preliminarily demonstrate the unbiasedness and acceptable efficiency of the algorithm. Obviously, the work in this paper is just a beginning, and there is still a lot of work to do, e.g. to apply the algorithm to practical engineering problems. Meanwhile, more intelligent active learning strategies can be adopted in SVM instead of the coarse one in this paper to further increase the efficiency of the algorithm.

ACKNOWLEDGMENTS

The research is supported by National Science Foundation of China for key item (Grant No. 51261120374), which is greatly appreciated.

REFERENCES
