Reduction of Random Variables in Structural Reliability Analysis

S. Adhikari & R. S. Langley
Cambridge University Engineering Department, Trumpington Street, Cambridge, U.K

Abstract

In the reliability analysis of a complex engineering structure a very large number of the system parameters can be considered to be random variables. The difficulty in computing the failure probability increases rapidly with the number of variables, and the aim of the present paper is to propose methods whereby the number of variables can be reduced without compromising the accuracy of the reliability calculation. The reduction methods are based on the sensitivity of the failure surface. Limited numerical examples are provided to illustrate the proposed methods.

Introduction

The most common methods of reliability prediction, that is (a) FORM (First Order Reliability Method), (b) SORM (Second Order Reliability Method), and (c) asymptotic reliability analysis, are now well documented (Thoft-Christensen and Baker 1982; Ditlevsen and Madsen 1996; Melchers 1999). Suppose the random variables describing the uncertainties of the structure and loading are considered to form a vector $\mathbf{y} \in \mathbb{R}^n$. The statistics of the system are fully described by the joint probability density function $p(\mathbf{y})$. In general the random variables $\mathbf{y}$ are non-Gaussian. In principle these random variables can be transformed to a set of uncorrelated Gaussian random variables via the Rosenblatt transformation (Rosenblatt 1952). Further, they can be scaled so that each random variables has zero mean and unit variance. Suppose these transformed and scaled random variables are $\mathbf{x} \in \mathbb{R}^n$ with a joint probability density function $p(\mathbf{x}) = (2\pi)^{-n/2} e^{-|\mathbf{x}|^2/2}$ where $|\mathbf{x}| = \sqrt{\mathbf{x}^T\mathbf{x}}$ is the euclidean norm of $\mathbf{x}$. For a given set of variables $\mathbf{x}$ the structure will either fail under the applied (random) loading or will be safe. The condition of the structure for every $\mathbf{x}$ can be described by a safety margin $g(\mathbf{x})$ so the structure has failed if $g(\mathbf{x}) \leq 0$ and is safe if $g(\mathbf{x}) > 0$. Thus the probability of failure is given by $P_f = \int_{g(\mathbf{x})\leq 0} p(\mathbf{x}) d\mathbf{x}$. The function $g(\mathbf{x})$ is also known as failure surface or limit-state function. The central theme of a reliability analysis is to evaluate this multidimensional integral. For most real-life cases the dimentionality of the integral is large and consequently the exact evaluation of the integral is not possible. For this reason some kind of approximate method is required to evaluate this integral. Using the first-order reliability method (FORM) the probability of failure is given by

$$P_f = \Phi(-\beta) \quad \text{with} \quad \beta = (\mathbf{x}^{*^T}\mathbf{x}^*)^{1/2} \tag{1}$$

where $\mathbf{x}^*$, the 'design point' is the solution of following optimization problem:

$$\min \beta = (\mathbf{x}^T\mathbf{x})^{1/2} \quad \text{subject to} \quad g(\mathbf{x}) = 0. \tag{2}$$

Once the reliability index (β) is known other more accurate approximate reliability analyses, for example, second-order reliability method (SORM) or asymptotic reliability method (Breitung 1984), can also be performed. Thus the calculation of the design point and the reliability index is very crucial for all approximate reliability methods. In this paper, based on sensitivity of $g(\mathbf{x})$, several methods for the efficient calculation of $\mathbf{x}^*$ and β are proposed.

Method 1: Gradient Projection Method

For some point $\mathbf{x}^*$ in $\mathbb{R}^n$, $g(\mathbf{x})$ can be expanded in a first-order Taylor series expansion of about $\mathbf{x}^*$ as $g(\mathbf{x}) \approx g(\mathbf{x}^*) + (\mathbf{x} - \mathbf{x}^*)^T \frac{\partial g(\mathbf{x})}{\partial \mathbf{x}}|_{\mathbf{x}=\mathbf{x}^*}$ If $g(\mathbf{x})$ is linear then $\frac{\partial g(\mathbf{x})}{\partial \mathbf{x}}$ is independent of $\mathbf{x}$. In this case $\mathbf{x}^*$ will be simply the projection of the origin ($\mathbf{x} = \mathbf{0}$) on $g(\mathbf{x}) = 0$ (see Chapter 3, Melchers (1999)). Also note that the outward normal vector to the hypersurface $g(\mathbf{x}) = 0$ is $c_i = \lambda \frac{\partial g(\mathbf{x})}{\partial x_i}|_{\mathbf{x}=\mathbf{0}}$. Assume that $\boldsymbol{\nabla} g = \left\{ \frac{\partial g(\mathbf{x})}{\partial x_i} \right\} \in \mathbb{R}^n$ is normalized so that $|\boldsymbol{\nabla} g| = 1$. Then for linear $g(\mathbf{x})$ it can be shown that $\mathbf{x}^* = -\beta \boldsymbol{\nabla} g$. Motivated by this, we express $\mathbf{x} = v\boldsymbol{\nabla} g$ where $v \in \mathbb{R}$ is a new random variable and consequently the constrained optimization problem (2) becomes a simple search problem, *i.e.*, we need to solve for v such that $g(v\boldsymbol{\nabla} g) = 0$ or $g'(v) = 0$. Here $g'(\bullet) = g(\bullet \boldsymbol{\nabla} g)$ is a (non-linear) function of a *single* variable v. Note that $v = -\beta$.

This method yields accurate result when $g(\mathbf{x})$ is linear or very close to linear. This is because for linear $g(\mathbf{x})$, $\boldsymbol{\nabla} g$ is independent of the choice of $\mathbf{x}$ so that the direction of the outward normal from the failure surface does not change with position along the failure surface. For this reason $\boldsymbol{\nabla} g|_{\mathbf{x}=\mathbf{0}}$ becomes the unit vector along true $\mathbf{x}^*$ and consequently β becomes simply the 'length' of this vector from the origin to the failure surface. These simple facts do not

hold when $g(\mathbf{x})$ is a nonlinear function. In this case $\boldsymbol{\nabla} g$ depends on the choice of $\mathbf{x}$ and the direction of the outward normal from the failure surface changes with position along the failure surface. Therefore $\boldsymbol{\nabla} g|_{\mathbf{x}=\mathbf{0}}$ is in general not the unit vector along true $\mathbf{x}^*$. To solve this problem we propose an iterative method so that the $\boldsymbol{\nabla} g$ is updated at each iteration step. We first obtain an initial $\mathbf{x}^*$, say $\mathbf{x}_0^*$, by projecting $\boldsymbol{\nabla} g|_{\mathbf{x}=\mathbf{0}}$ to the failure surface. Next we use this point to obtain a new unit vector $\boldsymbol{\nabla} g|_{\mathbf{x}=\mathbf{x}_0^*}$. Projecting this vector from the origin to the failure surface we obtain a next estimate of $\mathbf{x}^*$. The method then uses this point to calculate $\boldsymbol{\nabla} g$ and continues until two successive estimates of design points are close enough. In summary, for nonlinear $g(\mathbf{x})$, the iterative procedure can be described as follows:

1. For $k = 0$, select $\mathbf{x}^{(k)} = \mathbf{0}$, a small value of ϵ, (say 0.001) and a large value of $\beta^{(k)}$ (say 10).
2. Construct the normalized vector $\boldsymbol{\nabla} g^{(k)} = \left\{ \frac{\partial g(\mathbf{x})}{\partial x_i}|_{\mathbf{x}=\mathbf{x}^{(k)}} \right\}, \forall i = 1, .., n$ so that $|\boldsymbol{\nabla} g^{(k)}| = 1$.
3. Solve $g(v \boldsymbol{\nabla} g^{(k)}) = 0$ for v.
4. Increase the index: $k = k + 1$; denote $\beta^{(k)} = -v$ and $\mathbf{x}^{(k)} = v \boldsymbol{\nabla} g^{(k)}$.
5. Denote $\delta\beta = \beta^{(k-1)} - \beta^{(k)}$.
6. (a) If $\delta\beta < 0$ then the iteration is going in the wrong direction. Terminate the iteration procedure and select $\beta = \beta^{(k)}$ and $\mathbf{x}^* = \mathbf{x}^{(k)}$ as the best values of these quantities.
   (b) If $\delta\beta < \epsilon$ then the iterative procedure has converged. Terminate the iteration procedure and select $\beta = \beta^{(k)}$ and $\mathbf{x}^* = \mathbf{x}^{(k)}$ as the final values of these quantities.
   (c) If $\delta\beta > \epsilon$ then go back to step 2.

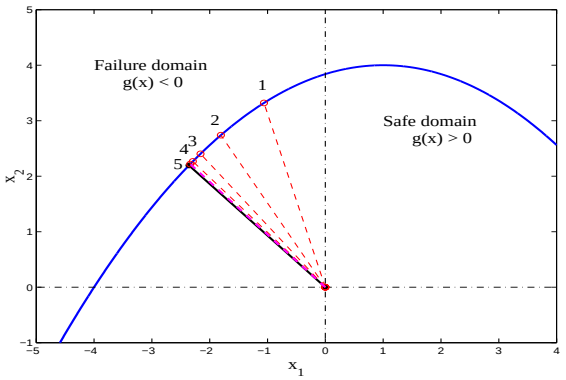

Figure 1: Nonlinear failure surface: '—' actual design vector $\mathbf{x}^*$; '−−' design vectors obtained at each iteration.

To illustrate a typical application of the proposed method we consider a system with two random variables x_1 and x_2 used by Melchers (1999, page 105). The non-linear failure surface is given by $g(\mathbf{x}) = -\frac{4}{25}(x_1-1)^2 - x_2 + 4 = 0$. Figure 1 shows the failure surface together with the design vector. The actual design vector and reliability index obtained by Melchers (1999) is given by $\mathbf{x}^* = \{-2.36, 2.19\}^T$ and $\beta = 3.22$. To apply the proposed method we need to obtain the gradients of the failure surface which are given by $\frac{\partial g(\mathbf{x})}{\partial x_1} = -\frac{8}{25}(x_1 - 1)$ and $\frac{\partial g(\mathbf{x})}{\partial x_2} = -1$. The starting point of the iterative scheme is $\mathbf{x} = \{0, 0\}^T$. Figure 1 shows $\mathbf{x}^*$ at each iteration step using the proposed method. After five iterations we obtain $\mathbf{x}^* = \{-2.34, 2.21\}^T$ and $\beta = 3.22$ which are very close to the exact values. However, It should be noted that the convergence of the proposed iterative method cannot be always guaranteed. It is expected that if the failure surface is fairly regular and linear then the method would converge.

Method 2: Dominant Gradient Method

Here we keep more than one random variable in the analysis. The criteria for selecting the random variables is based on the sensitivity of the failure surface with respect to the random variables. Suppose only n_d entries of normalized $\boldsymbol{\nabla} g$ have significantly non-zero (negative or positive) values while all other entries are close to zero. Consider the index set of these dominant variables is given by I_d. We construct the vector of dominant random variables $\mathbf{v} \in \mathbb{R}^{n_d}$ so that $v_j = x_i, \forall j = 1, .., n_d$ and $i =\in I_d$. All other random variables are assumed to be zero, that is $x_i = 0, \forall i \notin I_d$. This implies that these variables are assumed to be deterministic so that they do not play any role in the reliability analysis (see the discussion on 'omission sensitivity factors' by Madsen (1988)). Using these reduced set of variables one may obtain the failure surface in the reduced space $g'(\mathbf{v})$. The design point in the reduced space, $\mathbf{v}^*$, can be obtained from the solution of the optimization problem similar to (2) but involving $\mathbf{v}$ and $g'(\mathbf{v})$. From $\mathbf{v}^*$, the design point in the actual space can simply be obtained by substituting $x_i^* = v_j^*, \forall j = 1, .., n_d; i \in I_d$ and $x_i^* = 0, \forall\, i \notin I_d$.

It should be noted that this approach is based on sensitivity of $g(\mathbf{x})$ at $\mathbf{x} = \mathbf{0}$. For linear failure surface $\boldsymbol{\nabla} g|_{\mathbf{x}=\mathbf{0}}$ is the unit vector along true $\mathbf{x}^*$ and the dominant random variables selected by above procedure are actually the dominant random variables at the design point. For nonlinear $g(\mathbf{x})$, $\boldsymbol{\nabla} g|_{\mathbf{x}=\mathbf{0}}$ is in general not the unit vector along true $\mathbf{x}^*$. Thus the dominant random variables selected from $\boldsymbol{\nabla} g|_{\mathbf{x}=\mathbf{0}}$ are not necessarily the dominant random variables at the design point. Ideally the dominant random variables should be selected on the basis of $\boldsymbol{\nabla} g$ evaluated at the design point. Keeping this in mind, we try to update $\boldsymbol{\nabla} g$ in an iterative way so that the dominant

random variables are selected on the basis of $\boldsymbol{\nabla} g$ evaluated at points gradually closer to the true design point. The iterative procedure can be implemented by following these steps:

1. For $k = 0$, select $\mathbf{x}^{(k)} = \mathbf{0}$, a small value of ϵ, a large value of $\beta^{(k)}$ and also $n_d < n$.
2. Construct $\boldsymbol{\nabla} g^{(k)} = \left\{ \nabla g_i^{(k)} \right\} \forall i = 1, .., n$, where $\nabla g_i^{(k)} = \left\{ \frac{\partial g(\mathbf{X})}{\partial x_i} |_{\mathbf{X}=\mathbf{X}^{(k)}} \right\}$, so that $|\boldsymbol{\nabla} g^{(k)}| = 1$.
3. Sort $|\nabla g_i^{(k)}|$ to obtain the index set I_d corresponding to the biggest (positive or negative) n_d values.
4. Set $v_j = x_i^{(k)}, \forall j = 1, .., n_d$, $i \in I_d$ and $x_i^{(k)} = 0, \forall\, i \notin I_d$. Construct $\mathbf{v} = \{v_j\} \in \mathbb{R}^{n_d}$.
5. Using this transformation obtain $g'(\mathbf{v})$ from $g(\mathbf{x})$ and minimize $\beta = (\mathbf{v}^T \mathbf{v})^{1/2}$ subject to $g'(\mathbf{v}) = 0$.
6. Increase the index: $k = k + 1$. Using the solutions from step 5 denote $\beta^{(k)} = \beta$ and $\mathbf{v}^{(k)} = \mathbf{v}$.
7. Obtain $\mathbf{x}^{(k)}$ from the inverse transformation in step 4, *i.e.*, $x_i^{(k)} = v_j, \forall j = 1, .., n_d$, $i \in I_d$ and $x_i^{(k)} = 0, \forall\, i \notin I_d$.
8. Denote $\delta\beta = \beta^{(k-1)} - \beta^{(k)}$.
9. (a) If $\delta\beta < 0$ then the iteration is going in the wrong direction. Terminate the iteration procedure and select $\beta = \beta^{(k)}$ and $\mathbf{x}^* = \mathbf{x}^{(k)}$ as the best values of these quantities.
   (b) If $\delta\beta < \epsilon$ then the iterative procedure has converged. Terminate the iteration procedure and select $\beta = \beta^{(k)}$ and $\mathbf{x}^* = \mathbf{x}^{(k)}$ as the final values of these quantities.
   (c) If $\delta\beta > \epsilon$ then go back to step 2.

The initial choice of the dominant random variables, that is, n_d and I_d, can be automated by imposing a selection criteria, for example, by fixing the ratio of $\nabla g_i^{(k)}$ corresponding to the most sensitive random variable and the least sensitive random variable. Note that the index set of the dominant variables I_d may change in different iterations, however we fix n_d for all iterations. Like the previous method, the convergence of the proposed iterative method cannot be always guaranteed. In the next section another variant of this approach, which considers the contribution of the neglected variables, is proposed.

Method 3: Relative Importance Variable Method

Based on the entries of $\boldsymbol{\nabla} g$ we group the random variables into 'important' and 'unimportant' random variables. The random variables for which the failure surface is more sensitive are called important variables. Like the method in the previous section, suppose only $n_d < n$ entries of $\boldsymbol{\nabla} g$ with an index set $i \in I_d$ are important and they are casted in the vector $\mathbf{v} \in \mathbb{R}^{n_d}$. However, here we do not completely neglect all the 'unimportant' random variables, but consider that they can be represented by a single random variable, say u, such that $x_i = u \nabla g_i, \forall i \notin I_d$. This implies that the 'direction' of the unimportant random variables are fixed according to the gradient vector so that u is effectively a scaling parameter in that direction. This method can be viewed as a combination of the methods described in the previous two sections. Constructing the vector of reduced random variables $\mathbf{z} = \{\mathbf{v}^T, u\}^T \in \mathbb{R}^{n_d+1}$ one may obtain the failure surface in the reduced space $g'(\mathbf{z})$ and solve the optimization problem similar to (2) to obtain $\mathbf{z}^*$. From $\mathbf{z}^* = \{\mathbf{v}^{*^T}, u^*\}^T$, the design point in the actual space can simply be obtained by substituting $x_i^* = v_j^*, \forall j = 1, .., n_d; i \in I_d$ and $x_i^* = u^* \nabla g_i, \forall\, i \notin I_d$. Like the previous two methods, this approach can also be used in an iterative way for nonlinear $g(\mathbf{x})$ as follows:

1. For $k = 0$, select $\mathbf{x}^{(k)} = \mathbf{0}$, a small value of ϵ, a large value of $\beta^{(k)}$ and also $n_d < n$.
2. Calculate $\nabla g_i^{(k)}$ and sort $|\nabla g_i^{(k)}|$ to obtain the index set I_d corresponding to the biggest n_d values.
3. Set $v_j = x_i^{(k)}, \forall j = 1, .., n_d$, $i \in I_d$ and $x_i^{(k)} = u \nabla g_i^{(k)}, \forall\, i \notin I_d$. Construct $\mathbf{z} = \{v_j, u\} \in \mathbb{R}^{n_d+1}$.
4. Using this transformation obtain $g'(\mathbf{z})$ from $g(\mathbf{x})$ and minimize $\beta = (\mathbf{z}^T \mathbf{z})^{1/2}$ subject to $g'(\mathbf{z}) = 0$.
5. Increase the index: $k = k + 1$. Using the solutions from step 4 denote $\beta^{(k)} = \beta$ and $\mathbf{z}^{(k)} = \mathbf{z}$.
6. Obtain $\mathbf{x}^{(k)}$ from the inverse transformation in step 3.
7. Denote $\delta\beta = \beta^{(k-1)} - \beta^{(k)}$.
8. (a) If $\delta\beta < 0$ then the iteration is going in the wrong direction. Terminate the iteration procedure and select $\beta = \beta^{(k)}$ and $\mathbf{x}^* = \mathbf{x}^{(k)}$ as the best values of these quantities.
   (b) If $\delta\beta < \epsilon$ then the iterative procedure has converged. Terminate the iteration procedure and select $\beta = \beta^{(k)}$ and $\mathbf{x}^* = \mathbf{x}^{(k)}$ as the final values of these quantities.
   (c) If $\delta\beta > \epsilon$ then go back to step 2.

Numerical Example: Reliability Analysis of a Mutlistoried Portal Frame

We consider a mutlistoried portal frame with 20 members. The details of the structure and applied loading is shown in Figure 2 together with the element numbering and node numbering. It is assumed that the axial stiffness

(EA) and the bending stiffness (EI) of each member are Gaussian random variables so that there are in total 40 random variables, *i.e.*, $\mathbf{x} \in \mathbb{R}^{40}$. It is also assumed that EA and EI of different members are uncorrelated. The joint probability density function of the random variables is completely characterized by the mean and the standard deviation of EA and EI of each member. There are three types of elements and the numerical values of their properties are:

Element Type	EA (KN) Mean	EA (KN) Standard Deviation	EI (KNm^2) Mean	EI (KNm^2) Standard Deviation
1	5.0×10^9	7.0%	6.0×10^4	5.0%
2	3.0×10^9	3.0%	4.0×10^4	10.0%
3	1.0×10^9	10.0%	2.0×10^4	9.0%

The standard deviations are expressed as a percentage of the corresponding mean values. It is assumed that the column members are of type 1, the beam members are of type 2 and the diagonal members are of type 3. The failure condition is given by specifying a maximum allowable horizontal displacement at node 11, say d_{max}. We construct the failure surface $g(\mathbf{x}) = d_{max} - |\delta h_{11}(\mathbf{x})|$ where the random variable δh_{11}, the horizontal displacement at node 11, is in turn a function of all 40 random variables considered in the study. For the numerical calculations it is assumed that $d_{max} = 0.184 \times 10^{-2}$m. Numerical results obtained by using the proposed methods (arranged in the order of increasing computational cost) with only one iteration is shown below:

	Method 1	Method 2 $n_d = 5$	Method 3 $n_d = 5$	FORM $n = 40$	MCS
β	3.399	3.397	3.397	3.397	–
$P_f \times 10^3$	0.338	0.340	0.340	0.340	0.345

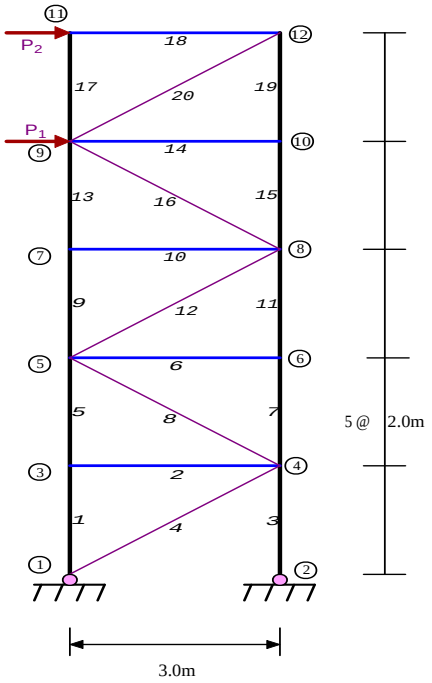

Figure 2: Mutlistoried portal frame, $P_1 = 400$MN and $P_2 = 500$MN.

It is clear that the all the proposed methods using reduced number of random variables produces very close results obtained by conventional FORM with full set of 40 variables. Further, also note that all the approximate reliability estimate methods show satisfactory agreement with Monte Carlo simulation (MCS) with 11600 samples (considered as benchmark).

Conclusions

Methods have been proposed to reduce the number of random variables in structural reliability problems involving a large number of random variables. In total three iterative methods, namely (a) gradient projection method, (b) dominant gradient method, and (c) relative importance variable method, have been proposed. All the three methods are based on the sensitivity vector of the failure surface. Initial numerical results show that there is a possibility to put these methods into real-life problems involving a large number of random variables. Further studies will involve analyzing the efficiency of the proposed methods when applied to problems with highly non-linear failure surfaces, for example, those arising in structural dynamic problems.

References

Breitung, K. (1984, March). Asymptotic approximations for multinormal integrals. *Journal of Engineering Mechanics, ASCE 110*(3), 357–367.

Ditlevsen, O. and H. O. Madsen (1996, January). *Structural Reliability Methods*. Chichester, West Sussex, England: John Wiley and Sons Ltd.

Madsen, H. O. (1988). Omission sensitivity factors. *Structural Safety 5*, 35–45.

Melchers, R. E. (1999, February). *Structural Reliability Analysis and Prediction* (2nd ed.). Chichester, West Sussex, England: John Wiley and Sons Ltd.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. *The Annals of Mathematical Statistics 23*, 470–472.

Thoft-Christensen, P. and M. J. Baker (1982). *Structural Reliability Theory and its Applications*. Berlin, Germany: Springer-Verlag.