12th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP12
Vancouver, Canada, July 12-15, 2015

Dimension Reduction Methods for Reliability Problems

Karl Breitung
Engineering Risk Analysis Group, Technical University of Munich, Munich, Germany

ABSTRACT: In reliability problems in high dimensional spaces it is important to identify the essential structure of the problem. Which variables can be neglected and which not? Here statistical dimension reduction methods help to simplify the complexity of structures. FORM/SORM can be seen as early example of dimension reduction concepts. Since these approaches do not rely on the smoothness properties of the limit state functions, they can be applied also for problems with noise. Three such methods are described. Further an approach giving the SORM factor as a ratio of small surface areas is outlined.

1. INTRODUCTION

In structural reliability problems in the last years more and more problems in high dimensional spaces have been treated. Due to the increasing computer power it is nowadays possible to handle such questions in a reasonable time. But obtaining numerical results for failure probabilities is not enough, even more important is to understand the structure of the system under consideration. What are its essential features and what can be neglected for am approximate, but enough accurate description of it? Such problems have appeared in other fields of science already some time ago. Here some of these methods which seem to be useful for reliability applications will be outlined.

Another problem is that LSF's (Limit State Function) not always are smooth functions but have some additional noise which makes it difficult or impossible to apply the standard FORM/SORM methods which require smooth LSF's.

Both problems can be tackled using dimension reduction concepts where the relation between and the LSF is seen more as a statistical relation and no more as an exact functional relation. These methods are in some way complementary to response surface methods (see for example Bucher and Macke (2004)).

2. DIMENSION REDUCTION AND FORM/SORM

Given is a system described by many variables $\mathbf{X} = (X_1, \ldots, X_n)$ as a function $Y = f(\mathbf{X})$ of them, can a simpler structure having *much less* variables be found without *much* loss of information about the system? Here Y is a function of the rv's $X_1, \ldots, X_n$ and n is large. Dimension reduction methods now attempt to find a function $h : \mathbb{R}^n \to \mathbb{R}^p$ with $p \ll n$ and a function $f^* : \mathbb{R}^p \to \mathbb{R}$ such that

$$f(X_1, \ldots, X_n) = f^*(h_1(\mathbf{X}), \ldots, h_p(\mathbf{X})) + \varepsilon \quad (1)$$

where the term ε is small in some sense. An overview can be found in Burges (2009). In Hurtado (2004) dimension reduction methods were studied the first time for structural reliability but with another perspective, i.e. to approximate an LSF given by data points using scalar products in higher dimensional spaces.

If h is a linear function, such a function is found by a projection on a suitable plane M with $h(\mathbf{X}) = \mathbf{P}_M\mathbf{X}$ where $\mathbf{P}_M$ is a projection matrix onto the p-dimensional plane M.

If the plane is spanned by p orthonormal vectors $\mathbf{v}_1, \ldots, \mathbf{v}_p$, then

$$f(\mathbf{x}) \approx f^*(\mathbf{v}_1^T\mathbf{x}, \ldots, \mathbf{v}_p^T\mathbf{x}) + \varepsilon \quad (2)$$

FORM can be seen as a linear dimension reduction method. The original LSF $g(\mathbf{u})$ is replaced by the

linearized form

$$g^*(\mathbf{u}) = a_0 + \sum_{i=1}^{n} a_i U_i = a_0 + \mathbf{a}^T \mathbf{u}, \tag{3}$$

which is the projection on the vector $y = \mathbf{a}^T \mathbf{u}$, $g^*(y) = a_0 + y$, $g(\mathbf{u}) \approx g^*(y)$. This is a reduction from n to one dimension. In the same way in SORM the limit state function is replaced by a quadratic form of normal random variables.

Now, the reduction to one dimension is not always possible, so for more complex problems it seem to be useful to make reductions where p is still small compared with n, but greater than unity.

The classical counterexample for FORM/SORM is a LSF in the form

$$g(u_1, \ldots, u_n) = \beta - \sum_{i=1}^{k} (\mathbf{a}_i^T \mathbf{u})^2, \tag{4}$$

where the $\mathbf{a}_i$ are unit vectors. Here there is no unique design/beta point and the approximations of FORM/SORM can not be derived. But using dimension reduction methods it is easily possible to identify the relevant variables of this LSF.

3. DIMENSION REDUCTION METHODS

Here some dimension reduction methods are described. In the following it is assumed that $Y = g(\mathbf{U})$ is a function of a standard normal random vector $\mathbf{U}$. Such methods are for example SIR (Sliced Inverse Regression, Li (1991)) and SAVE (Sliced Average Variance Estimates, Cook and Weisberg (1991)). These two approaches are, respectively, based on the inverse mean of $\mathbf{U}$ given Y and the inverse conditional variance of $\mathbf{U}$ given Y. The idea of these methods is that if Y is a function of the $\mathbf{v}_i^T \mathbf{U}$ vectors only as in eq. (2), then the inverse regression, i.e. estimating the $\mathbf{U}$'s from the y's will afflict only these variables.

Let for example be

$$Y = g(U_1, \ldots, U_p, 0, \ldots, 0) \tag{5}$$

with $1 \leq k < n$. Then only changes in the first p variables will have an influence of Y. In the same way if one now considers variations of Y then the conditional mean $I\!E(U_1, \ldots, U_p, U_{p+1}, \ldots, U_n | Y)$ will vary only in the first k variables, i.e. lie in the plane spanned by the first p unit vectors.

In the same way, if $Y = g(\mathbf{v}_1^T \mathbf{U}, \ldots, \mathbf{v}_p^T \mathbf{U}, 0, \ldots, 0)$ i.e. a function of $\mathbf{v}_1^T \mathbf{U}, \ldots, \mathbf{v}_p^T \mathbf{U}$ only, then the curve of the inverse regression mean $I\!E(\mathbf{U}|y)$ lies in the subspace M spanned by the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_p$. Since the curve $I\!E(\mathbf{U}|y)$ lies in M, the vectors connecting the points $I\!E(\mathbf{U}|y_i)$ for various values of y_i must lie in the plane M. So if we can construct enough such points $I\!E(\mathbf{U}|y_i)$ from their differences $I\!E(\mathbf{U}|y_i) - I\!E(\mathbf{U}|y_j)$ we can find the vectors which span M. Certainly there are exceptions where these differences fail to span the whole subspace.

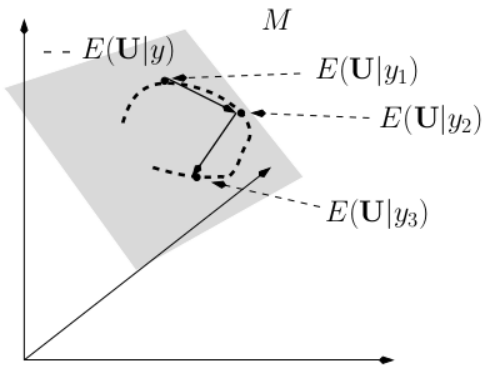

Figure 1: The variation of $I\!E(\mathbf{U}|y)$ in the plane M

We cannot find the values of $I\!E(\mathbf{U}|y))$ for a fixed y, but we can approximate them by taking slices. This means, we can find the values $I\!E(\mathbf{U}|y \in S_i)$ with $S_i = [a,b]$ an interval. These $I\!E(\mathbf{U}|y \in S_i)$ points lie in the subspace M too.

The SIR method estimates directly the matrix $\mathrm{cov}(I\!E(\mathbf{U}|y))$ using slices. Given now a set of data $(y_k, \mathbf{u}_k)$ with $k = 1, \ldots, m$ where y is a function of the $\mathbf{u}$. First the real axis is partioned into h slices $S_1, \ldots, S_h$. Let m_i the number of data in slice S_i and

$$\bar{\mathbf{u}}_i = m_i^{-1} \sum_{y_j \in S_i} \mathbf{u}_j \tag{6}$$

be the sample mean of the vectors $\mathbf{u}_j$ with $y_j \in S_i$. Then the estimator of the covariance matrix is given by:

$$\widehat{\mathbf{S}}_1 = m^{-1} \sum_{i=1}^{h} m_i (\bar{\mathbf{u}}_i \bar{\mathbf{u}}_i^T) \tag{7}$$

This is now an approximation of the matrix $\mathrm{cov}(I\!E(\mathbf{U}|y))$. By taking the p largest eigenvalues and the corresponding eigenvectors one finds an approximation for the subspace M, i.e. the space spanned by the p largest eigenvectors $\mathbf{v}_1,\ldots,\mathbf{v}_p$.

In the alternative method SAVE one uses estimates about the variation of the conditional mean to find the vectors spanning M. Since due to the law of total variance

$$\mathrm{cov}(\mathbf{U}) = I\!E(\mathrm{cov}(\mathbf{U}|y)) + \mathrm{cov}(I\!E(\mathbf{U}|y)), \qquad (8)$$

the matrix $I\!E(\mathrm{cov}(\mathbf{U}|y))$ can be estimated also from $\mathbf{I}_n - \mathrm{cov}(I\!E(\mathbf{U}|y))$.

Let $\mathbf{I}_n$ be the n-dimensional unity matrix and $\mathbf{P}$ the projection matrix onto the subspace M, then

$$\begin{aligned}
\mathrm{cov}(\mathbf{U}|y) &= \mathbf{I}_n - \mathbf{P} + \mathbf{P}\,\mathrm{cov}(\mathbf{U}|y)\mathbf{P} \\
\mathbf{I}_n - \mathrm{cov}(\mathbf{U}|y) &= \mathbf{P} - \mathbf{P}\,\mathrm{cov}(\mathbf{U}|y)\mathbf{P} \\
\mathbf{I}_n - \mathrm{cov}(\mathbf{U}|y) &= \mathbf{P}\mathbf{I}_n\mathbf{P} - \mathbf{P}\,\mathrm{cov}(\mathbf{U}|y)\mathbf{P} \\
\mathbf{I}_n - \mathrm{cov}(\mathbf{U}|y) &= \mathbf{P}(\mathbf{I}_n - \mathrm{cov}(\mathbf{U}|y))\mathbf{P} \qquad (9)
\end{aligned}$$

So the matrix $\mathbf{I}_n - \mathrm{cov}(\mathbf{U}|y)$ is equal to its projection onto the subspace M. If M is spanned by the first p unit vectors $\mathbf{e}_1,\ldots,\mathbf{e}_p$ then this means

$$\begin{aligned}
&\mathbf{I}_n - \mathrm{cov}(\mathbf{U}|y) \\
&= \begin{pmatrix} \mathbf{I}_p - \mathrm{cov}(U_1,\ldots,U_p|y) & \mathbf{0}_{p,\,n-p} \\ \mathbf{0}_{n-p,\,p} & \mathbf{0}_{n-p,\,n-p} \end{pmatrix} \qquad (10)
\end{aligned}$$

with $\mathbf{0}_{k,\,l}$ the zero matrix with k rows and l columns. So the p non-zero eigenvectors of $\mathbf{I}_n - \mathrm{cov}(\mathbf{U}|y)$ span M.

We make an estimate $\hat{\mathbf{S}}_2$ of $\mathbf{I}_n - I\!E(\mathrm{cov}(\mathbf{U}|y))$. Then the eigenvalues and eigenvectors of $\hat{\mathbf{S}}$ will approximate those of $\mathbf{I}_n - I\!E(\mathrm{cov}(\mathbf{U}|y))$. Here m is the number of the data, the real line is partitioned into h slices S_i, and m_i is the number of data in slice S_i. Then an estimate of $\mathbf{S}$ is given by:

$$\widehat{\mathbf{S}}_2 = m^{-1}\sum_{i=1}^{h} m_i(\mathbf{I}_n - \mathrm{cov}(\mathbf{U}|y \in S_i))^2 \qquad (11)$$

Here we take the square of $\mathbf{I}_n - I\!E(\mathrm{cov}(\mathbf{U}|y))$ to get non-negative eigenvalues. If Y is a function of only p variables, this matrix has $n-p$ eigenvalues approximately equal to zero and the eigenvectors for the remaining p positive eigenvalues span the subspace M.

Both methods SIR and SAVE have some weaknesses. So SIR detects linear relations quite well, but has difficulties to find quadratic dependencies, with SAVE it is exactly vice versa. Various improvements and new methods have been proposed. Here only one of these trying to combine the estimators is described. The mixed estimator derived in Zhu et al. (2007) is then

$$\widehat{\mathbf{S}}_3 = (1-a)\widehat{\mathbf{S}}_1 + a\widehat{\mathbf{S}}_2 \qquad (12)$$

where $0 < a < 1$. The combination of these estimators gives a method which is able to detect linear and quadratic functional relations quite well.

4. EXAMPLE

A random sample of 10000 points is taken with the LSF with noise term ε

$$g(\mathbf{u}) = 3 - u_1 - 0.3\cdot u_2^2 + \varepsilon \qquad (13)$$

The noise term ε has a normal distribution with mean zero and variance 0.5. Using the mixed estimator in eq. (12) with $a = 0.5$ one obtains as matrix of eigenvectors (columns)

$$\begin{pmatrix} 0.01 & 0.04 & 0.02 & 1.00 & -0.08 \\ -0.02 & -0.01 & -0.01 & -0.08 & -1.00 \\ 0.20 & -0.88 & 0.42 & 0.02 & -0.00 \\ 0.98 & 0.20 & -0.05 & -0.01 & -0.01 \\ 0.04 & -0.42 & -0.90 & 0.04 & 0.01 \end{pmatrix} \qquad (14)$$

The corresponding eigenvalues are:

$$\begin{pmatrix} 0.00 & 0.00 & 0.01 & 0.52 & 0.86 \end{pmatrix} \qquad (15)$$

The eigenvectors with the two largest eigenvalues are approximately equal to the first and second unit vector, they correctly identify the relevant components.

5. ESTIMATING THE SORM-FACTOR AS SURFACE AREA RATIO

Estimating the SORM factor cannot be done by these methods. Here an approach is given which replaces the calculation of the Hessian of the LSF by estimating surface areas.

The SORM factor is given as (Breitung (1984), Breitung (1994)) $\prod_{i=1}^{n-1}(1-\beta\kappa_i)^{-1/2}$, where the κ_i's are the main curvatures of the limit state surface at the beta point. This form uses the orientation of the limit state surface by the normal vector field $\mathbf{n}(\mathbf{u}) = |\nabla g(\mathbf{u})|^{-1}\nabla g(\mathbf{u})$. Ths means that curvatures of the surface bending towards the surface normal are defined as positive, bending away as negative. Therefore the negative sign. This ensures also that a sphere around the origin where the safe domain is inside the sphere has positive curvature which is the usual convention.

The following differential geometry facts are taken from Thorpe (1979). The Gauss-Kronecker curvature of a surface at a point is the product of its main curvatures. The Gauss map for an surface G oriented by the normal vector field $\mathbf{n}(\mathbf{u})$ is defined by $G \to S^n, \mathbf{u} \mapsto \mathbf{n}(\mathbf{u})$, where $S^n = \{\mathbf{x}; |\mathbf{x}| = 1\}$ is the unit sphere in $\mathbb{R}^n$.

Consider now the Lagrangian of the problem given by $L(\mathbf{u},\lambda) = |\mathbf{u}|^2/2 + \lambda g(\mathbf{u})$. The surface $G^* = \{\mathbf{u}; L(\mathbf{u},\beta) - \beta^2/2 = 0\}$ is a hypersurface in the n-dimensional space containing the beta point $\mathbf{u}^*$. From its definition one can deduce that all main curvatures at the beta point are positive. At the beta point the Gauss-Kronecker curvature $K(\mathbf{u}^*)$ of this hypersurface is given by

$$K(\mathbf{u}^*) = \prod_{i=1}^{n-1}(1-\beta\kappa_i) \tag{16}$$

So the Gauss-Kronecker curvature of this hypersurface is equal to the square of the inverse of the SORM-factor. This shows a way to compute it from the geometrical properties of the surface G^*.

Let $W \subset \mathbb{R}^{n-1}$ a neighborhood of the origin and $\Psi : W \to G^*$ be a local parametrization of the surface G^* near the beta point with $\Psi(\mathbf{0})) = \mathbf{u}^*$ and let $N : W \to S^n$, $\mathbf{z} \mapsto \mathbf{n}(\Psi(\mathbf{z}))$ the Gauss map of the surface G^* for all points $\mathbf{z} \in W$.

From the corollary on p. 144 in Thorpe (1979) one obtains (with $V(.)$ denoting the $(n-1)$-dimensional area in $\mathbb{R}^n$):

$$K(\mathbf{u}^*) = \lim_{\varepsilon\to 0} \frac{V(N|B_\varepsilon)}{V(G^*|B_\varepsilon)} \tag{17}$$

with $B_\varepsilon = \{\mathbf{z}; |\mathbf{z}| < \varepsilon\} \subset W$. So by computing the ratio of the area of the surface and of the area of its Gauss map on the n-dimensional unit sphere one a can approximate the SORM factor (for details about calculating the surface area for a given local parametrization see Courant and John (1974)).

6. CONCLUSIONS

Dimension reduction methods can detect essential structures in high dimensional reliability problems. Combining them with response surface approaches might allow to find more efficient failure probability estimators.

7. REFERENCES

Breitung, K. (1984). "Asymptotic approximations for multinormal integrals." *Journal of the Engineering Mechanics Division ASCE*, 110(3), 357–366.

Breitung, K. (1994). *Asymptotic Approximations for Probability Integrals*. Springer, Berlin Lecture Notes in Mathematics, Nr. 1592.

Bucher, C. and Macke, M. (2004). "Response Surfaces for Reliability Assessment." *Engineering Design Reliability Handbook*, E. Nikolaidis et al., eds., CRC Press LLC, Chapter 19, 19:1–23.

Burges, C. (2009). "Dimension reduction: A guided tour." *Machine Learning*, 2(4), 275–365.

Cook, D. and Weisberg, S. (1991). "Comment to "Sliced Inverse Regression for Dimension Reduction" by K.-Ch. Li." *Journal of the American Statistical Association*, 86(414), 328–332.

Courant, R. and John, F. (1974). *Introduction to Calculus and Analysis*, Vol. 2. John Wiley, New York.

Hurtado, J. (2004). *Structural Reliability: Statistical Learning Perspectives*. Springer.

Li, K.-C. (1991). "Sliced inverse regression for dimension reduction." *Journal of the American Statistical Association*, 86(414), 316–327.

Thorpe, J. (1979). *Elementary Topics in Differential Geometry*. Springer, New York.

Zhu, L.-X., Ohtaki, M., and Li, Y. (2007). "On hybrid methods of inverse regression-based algorithms." *Computational Statistics & Data Analysis*, 51, 2621 – 2635.