Available online at www.sciencedirect.com
ScienceDirect
Procedia Social and Behavioral Sciences 2 (2010) 7712–7713
Procedia Social and Behavioral Sciences

Sixth International Conference on Sensitivity Analysis of Model Output

An original sensitivity statistic within a new adaptive accelerated Monte-Carlo method

Munoz Zuniga Miguel[a,b,*], Garnier Josselin[a], Remy Emmanuel[b]

[a] *Laboratoire de probabilités et modèles aléatoires, Université Paris VII, Case courier 7012, 2 place Jussieu, 75251 Paris cedex 05, France*
[b] *EDF R&D, 6 quai Watier, 78400 Chatou, France*

Abstract

The aim of this paper is to present a sensitivity statistic developed in the context of the design of a new accelerated Monte-Carlo method. In the field of structural reliability, we elaborated the "Adaptive Directional Stratification" method (ADS), in order to estimate small failure probabilities in a robust manner with a limited number of simulations. In order to break the curse of dimensionality, we propose an efficient statistic, evaluated at the end of the learning stage of the ADS method, to detect the input variables which are the most influential on the failure event. Thereby, we can focus the computational effort by stratifying only the most influential variables, which allows to better deal with high-dimensional spaces in the estimation step of the ADS method.

Keywords: accelerated Monte-Carlo ; sensitivity ; failure ; probability ; stratification ; directional

1. ADS with a sensitivity statistic

To deal with the problem of robust estimation of small failure probabilities with a limited number of simulations, we developed a new accelerated Monte-Carlo method, called: Adaptive Directional Stratification (ADS). It consists in an adaptive method with 2 stages in the adaptation stage, coupling the stratification and the directional simulation methods, which are well-known separately Kroese (2007) and Ditlevsen (1996).
To estimate the failure probability:

$$P_f = P(G(X) < 0) \tag{1}$$

with G the failure function, supposed greedy in computational resources, and X the p-dimensional random input vector, we first carry out a preliminary stage. First, we transform X into a random vector U with all its components as standard Gaussian variables. Then, in the Gaussian space, we factorize U into the product of a directional angle random vector, A, uniformly distributed over the unit sphere, and a radius random variable, R, such that R^2 follows a chi-square distribution with p degrees of freedom - see Nelsen (1999) and Ditlevsen (1996). Next, we decompose P_f

* Corresponding author. Tel.: +33-630639428
E-mail address: mmunozzu@hotmail.com

doi:10.1016/j.sbspro.2010.05.192

conditionally to A and stratify the unit sphere into "quadrants" - see Ditlevsen (1996) and Munoz Zuniga (2009) - which completes the preliminary stage. Then, the ADS method combines a learning and an estimation stage, which requires to split into two parts the total number of directional simulations we are limited to (around 500, which corresponds to a few thousands of calls to the failure function). With the first part of the directional draws allocated to the learning stage, we estimate, for each quadrant, the optimal allocation of the remaining directional simulations to be achieved in the estimation stage to reduce the variance of the failure probability estimator. Lastly, in the second stage, we estimate the failure probability with the estimated optimal allocations. For more details on the ADS method, we refer to Munoz Zuniga (2009). This method gives controlled and efficient results (i.e. an estimator of the failure probability and an estimator of the error) for a small number of simulations when the dimension of the input vector is up to 4. Unfortunately, when the input dimension grows, the number of quadrants of the ADS method increases exponentially: indeed, in dimension p, the number of quadrants is 2^p. As a minimum of draws is required to explore each quadrant, the number of directional simulations needed is too large for the restricted number of simulations we have.

That is why we developed a sensitivity statistic which relies on the idea to get during the learning stage , a sort of the random variables in function of their influence on the failure event, enabling, in the estimation stage, to only stratify the most significant variables. To determine if a random variable will be stratified, we propose the following method. First, we index the quadrants. The input index $k \in \{1,...,p\}$ is given the tag: i_k which takes its values in $\{-1,1\}$ and corresponds to the input sign. Thus, each quadrant is characterized by a p-uple $(i_1,...,i_p)$. Then, we define the sequence $T := (T_k)_{k \in \{1,...,p\}}$ by:

$$T_k = \sum_{i_l \in \{-1,1\}, l \neq k} \left| f\left(\widetilde{P}(i_1,...,i_{k-1},1,i_{k+1},...,i_p)\right) - f\left(\widetilde{P}(i_1,...,i_{k-1},-1,i_{k+1},...,i_p)\right) \right| \quad (2)$$

with $\widetilde{P}(i_1,...,i_p)$ the estimation of the failure probability in the quadrant $(i_1,...,i_p)$ obtained during the learning stage and f an increasing function. We propose to take f equal to the identity function. In this case, the sensitivity statistic T_k aggregates the differences of the failure probabilities between the quadrants along the dimension k. The larger T_k is, the more influential the k-th input is. Then, we sort the sequence T by decreasing order and decide to stratify only over the $p'<p$ first dimensions, the other inputs being simulated without stratification. We have numerically showed that, for the targeted number of simulations, the ADS method is efficient when the number of random variables is close to 3. So a reasonable advice will be to take $p'=3$. Then, we estimate the optimal allocation to be achieved in the new $2^{p'}$ hyper-quadrants with the learning stage simulations and we apply the second stage of the ADS method to estimate the failure probability. The statistic T offers accurate selection of the influential variables and enables to consider a dimension up to 6. Beyond six, the space is definitively too large to be surveyed with a limited number of simulations as we have. We can also indicate that we could have been more drastic and, after having determined the p' most important inputs, have set the other inputs to conservative values (if available) and carried out the ADS method in the p'-dimensional space. This kind of strategy is a common practice in risk analysis, but if we make a wrong classification of the inputs with the statistic T, then we will set one (or several) influential input variable to a conservative value and the estimation of the failure probability might be too much conservative to be of practical interest. Another possibility is to keep simulating the p-p' variables by a classical Monte-Carlo simulation: with this solution, the problem is that we loose the directional strategy in the "eliminated" dimensions and as we are looking for small probabilities, this strategy dramatically reduces the estimation accuracy when there are more than p' influential inputs. In conclusion, the ADS method coupled with the sensitivity statistic T constitutes an efficient method in our context.

2. References

Munoz Zuniga M., Garnier J., Remy E. and de Rocquigny E., 2009: Adaptive Directionnal Stratification – An adaptive directional sampling method on a stratified space, *Communication at ICOSSAR congress*.

Nelsen R.B., 1999: An introduction to copulas. Springer.

Gluver H., Ditlevsen O., Melchers R.E., 1990: Probability integration by directional simulation, *Journal of Eng. Mech.*, 36, 355–68.

Kroese D.P., Rubinstein R.Y., 2007: Simulation and the Monte-Carlo Method (second edition), Wiley.

Ditlevsen O., Madsen H.O., 1996: Structural reliability methods, Wiley.