Extension of sample size in Latin Hypercube Sampling with correlated variables

M. Vořechovský

Institute of Structural Mechanics, Faculty of Civil Engineering, Brno University of Technology, Brno 602 00, Czech Republic, vorechovsky.m@fce.vutbr.cz

Keywords: *Latin Hypercube Sampling; covariances; multivariate random vectors; dependence; copula; combinatorial optimization; Simulated Annealing.*

Abstract

In many computer experiments the adequacy of a given sample to give acceptable estimates of desired statistical quantities cannot be determined a priori, and thus the ability to extend or refine an experimental design may be important. This can be done very easily in crude Monte Carlo sampling. Very often, though, running each realization (physical or virtual experiment) is very expensive. Therefore the variance reduction techniques such as Latin Hypercube Sampling (LHS) is a suitable option, because it yields lower variance of estimates of statistical moments compared to crude Monte Carlo sampling at the same sample size. In conventional LHS, however, it is the necessary to specify the number of simulations (or physical realizations in the design of experiments) in advance. However, in real life problems the sample size yielding stable and statistically significant estimations of output statistics is not known beforehand. For example, when an analyst is planning to run complex nonlinear finite element computations using LHS, it is unclear how many simulations are needed. If too small sample set is used (i.e. a set that does not give acceptable statistical results), the analyst normally has to abandon the results and run new analyses with a larger sample set. It is thus desirable to start with a small sample and then extend (or refine) the design if deemed necessary. The extension would permit the use of a larger sample set without the loss of any of the already performed, and possibly quite expensive, calculations (experiments).

This problem has been overcome by the method recently proposed by the author. An algorithm is based on a hierarchy of Latin Hypercube-like samples which are proven to yield to smaller variances of results compared to the crude Monte Carlo. The subsets sampled by the proposed method can be merged together exploiting the property of variance reduction, yet retaining the sampling flexibility. The whole procedure of a cascade of LHS-like runs can be fully automated and the (automatic) stopping criterion might be e.g. the significance of output statistics or desired computational time.

References

Vořechovský, M. Stochastic computational mechanics of quasibrittle structures. *Habilitation thesis presented at Brno University of Technology*, Brno, Czech Republic, 2007.