Efficient methods for optimal space filling in model reduction techniques

Vladimir Buljak[1] **and Tomasz Garbowski**[2]

[1] *Faculty of Mechanical Engineering, University of Belgrade*
80 Kraljice Marije Str., Belgrade, 11000, Serbia
email: vbuljak@mas.bg.ac.rs
[2] *Institute of Structural Engineering, Poznan University of Technology*
Piotrowo 5 street, 60-965 Poznan, Poland
e-mail: tomasz.garbowski@put.poznan.pl

Abstract

Model reduction techniques generate low-dimensional models to parametrize PDEs in order to allow for efficient evaluation of highly non-linear problems in many-query and real-time context. Computing time in these methods is divided into a time consuming 'off-line' phase needed to 'train' a surrogate model, which is further used in 'on-line' phase providing accurate results in a much faster way. The accuracy of these methods is strictly connected to the number of points in which the system response is previously computed and on their distribution in a parameter space. For a given number of analysis decided to be 'invested' for the design of surrogate model, the accuracy can be further improved by the adequate distribution of them in a parameter space. Two methods of robust samples distribution are proposed here. A first method is based on interactive nodes aimed to optimize the distribution of them. A second is based on an optimal Latin hypercube design. Both presented methods are flexible with respect to the number of nodes involved and they can provide uniform distribution for any arbitrary number of them. Furthermore, they allow also for taking into account different importance of divers parameters.

Keywords: interactive nodes, optimal Latin hypercube, model reduction

1. Introduction

Inverse analysis and parametric studies require a large number of time consuming analyses that differ between each other just by a few parameters which are changing from one simulation to another. For these purposes recently become popular the use of model reduction techniques aimed to design low-order models capable of evaluating responses of complex systems within drastically reduced computing time. Survey of different applications of model reduction techniques based on Proper Orthogonal Decomposition (POD) is given in [1]. Similar methods based on Proper Generalized Decomposition (PGD) are developed and applied in computational fluid dynamics and in computational fracture mechanics (see e.g. [2, 3]).

All of these methods have in common computational strategy which consist in an a priori time consuming computation done once-for-all, which precedes a routine repeated computation that makes the use of previously generated results. By increasing the number of simulations done as a part of a preliminary 'training' it is obvious that the error of surrogate model will be decreased, but for any given number of them, the optimization of error is strictly connected to their distribution.

2. Interactive nodes method

A method presented in this section is based on dynamic simulation of interactive nodes of evaluation points which leads to the designs that are uniformly filling the space. Method starts from a random distribution of nodes, attributing a unitary mass to each of them and introducing repulsive forces which are inversely proportional to the distances between them. Apart of these forces, also a damping force needs to be added in order to make the nodes eventually resting in a steady state, otherwise they would continue to oscillate about their equilibrium positions. As a second phase, an explicit dynamic simulation is performed and a stable solution is found that corresponds to the uniform distribution of nodes in a parameter space.

The proposed method can be easily extended to the case with different priorities of diverse parameters on which the system depends, which is a frequent case in modeling diverse phenomena using model reduction techniques. This is achieved simply by stretching the space along those directions for which there is a larger importance of the parameters, followed by a subsequent mapping into original space after the equilibrium solution is found.

3. Optimal Latin hypercube design method

Another efficient method of space filling is based on Latin hypercube design (LHD) proposed by McKay et al. [4] and Iman and Conover [5]. The LHD with n_p points is constructed in such a way that each of the n_v variables is divided into n_p equal levels and that there is only one point (or sample) at each level. Usually a random procedure is used first to determine the point locations and later the optimization is performed in order to improve points distribution in the parameter space. This, however, requires many iterations (i.e. is computationally very costly) especially when the dimensionality of parameters space is high or the number of a-priori selected samples (points) is larger than few thousands.

In the literature many efficient methods for optimization of LHD based on e.g. coordinate exchange algorithms, enhanced stochastic evolutionary algorithm or columnwise-pairwise and genetic algorithms can be found (see details in [6]). The objective of the optimization is usually the minimization of potential energy or maximization of a minimal distance between all points. Recently very efficient algorithm was proposed [6] which does not use any optimization technique but a systematic translational propagation of points arranged in particular seed design. The algorithm allows to design a high dimensional, multi-points Latin hypercube within a fraction of the second.

4. Examples

In order to illustrate a space filling properties of different algorithms a simple example using a distribution of 90 samples in 2-dimensional space is shown in Fig.1.

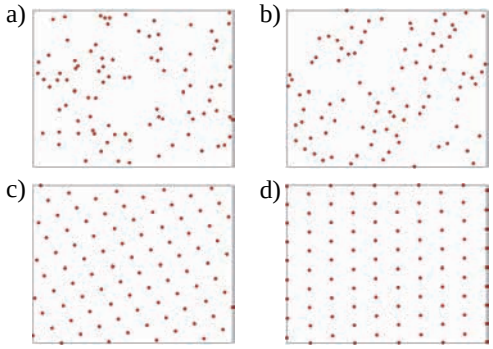

Figure 1: Distribution of 90 points in 2-dimensional space by: (a) random uniform generator, (b) Latin hypercube design, (c) interactive nodes algorithm, (d) topology propagation Latin hypercube design.

The visual inspection of space filling properties of above listed algorithms (shown in Fig. 1a–d) gives a qualitative evaluation of their ability to create an optimal design. It is visible that both Topology Propagation Latin Hypercube Design (TPLHD) and Interactive Nodes (IntNod) algorithms create well-distributed grids, while LHD and random uniform generator perform rather poor. We can also compare above algorithms in a quantitative way by measuring their performance using formula:

$$\phi_p = \left[\sum_{i=1}^{n_p-1} \sum_{j=i+1}^{n_p} d_{ij}^{-p} \right]^{1/p} \tag{1}$$

where p is a positive integer, n_p is the number of points in the design, and d_{ij} is the inter-point distance between all point pairs in the design. The general inter-point distance between any point pair $\mathbf{x}_i$ and $\mathbf{x}_j$ can be expressed as follows:

$$d_{ij} = d\left(\mathbf{x}_i, \mathbf{x}_j\right) = \left[\sum_{k=1}^{n_v} |x_{ik} - x_{jk}|^t \right]^{1/t} \tag{2}$$

Here $p = 50$ and $t = 1$ are used. Minimizing the value of performance measure ϕ_p leads to the maximization of point-to-point distance in the design. The values of ϕ_p of selected algorithms are presented in the Tab. 1. The regular grid used as reference in Tab. 1 consists of 10×9 points.

Table 1: The performance measurement of 6 methods of points distribution in 2-dimensional parameter space

RAND	LHD	IntNod	TPLHD	regular grid
48.56	31.27	9.37	10.83	9.83

In the second example a design in 6-dimensional parameters space with 7,000 samples is considered. The model reduction is performed by POD-ANN system where Proper Orthogonal Decomposition truncates the 9-dimensional input vector to just 4-dimensional and the General Regression Artificial Neural Network is used for model approximation. Surrogate is trained both on the samples generate by IntNod and TPLHD algorithms and later tested on randomly generated 6,000 points. The performance is presented in Fig. 2, where the reconstruction error of one over 6 parameters is shown.

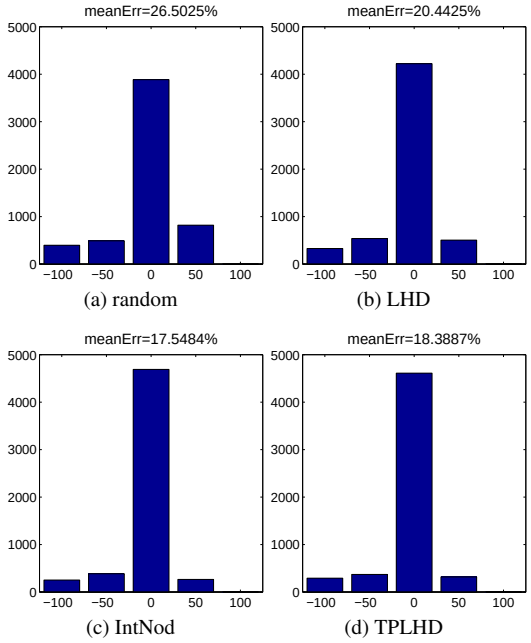

Figure 2: The approximation error of a design in 6-dimensional parameters space with 7,000 samples

5. Conclusions

The space filling property in the design of sampling points distribution plays a crucial role in model reduction performance. The better distribution in the parameters space the better approximation of the model. If the space dimensionality is relatively small and there is no limitation of number of the grid points the regular grid is the easiest and the fastest way to distribute points in the space. If, however, the space has for example 9 dimensions then a regular grid with only 3 levels for each parameter requires $3^9 = 19,683$ sampling points. In such situation it is desirable to use other sampling methods (e.g. TPLHD or IntNod) with smaller number of sampling points but with similar performance.

References

[1] Ruckelynck D., Chinesta F., Cueto E. and Ammar A., On the a priori Model Reduction: Overview and Recent Developments. *Arch Comput Methods Eng*, 13(1) pp. 91-128, 2006.

[2] Rozza G. Huynh D.B.P. and Patera A.T., Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations - application to transport and continuum mechanics. *Arch Comput Methods Eng*, 15 (3), pp. 229-275, 2008.

[3] Nouy A., A priori model reduction through Proper Generalized Decomposition, *Comput. Methods Appl. Mech. Engrg.*, 199, pp. 1603-1626, 2010.

[4] McKay M.D., Beckman R.J., Conover W.J., A comparison of three methods for selecting values of input variables from a computer code, *Technometrics*, 21, 239–245, 1979.

[5] Iman R.L., Conover W.J., Small sample sensitivity analysis techniques for computer models, with an application to risk assessment, *Commun Stat A-Theor*, 17, 1749–1842, 1980.

[6] Viana F.A.C., Venter G., Balabanov V., An algorithm for fast optimal Latin hypercube design of experiments, *Int. J. Numer. Meth. Engng*, 82, 135-156, 2010.