20th International Conference
ENGINEERING MECHANICS 2014
Svratka, Czech Republic, 12 – 15 May 2014

COMPARISON OF METHODS FOR CONSTRAINED DESIGN OF COMPUTER EXPERIMENTS

E. Myšáková*, M. Lepš**

Abstract: *A design of experiments creates an essential part in any experimentation, development of meta-models, sensitivity analysis or probability calculations. The creation of the design of experiments is determined by the shape of a design domain. In case of dependent input variables the design domain is constrained. Therefore classical designs developed for regular hypercubes cannot be applied here. This contribution presents comparison of several methods for designs in constrained design domains in terms of space-filling properties of the resulting designs and time required for their generation. Examples are focused on a special type of experiments called a mixture experiment, where individual parameters (ingredients) form a unit volume or unit weight. This condition leads to a simplex or a convex polytope shape of the design domain.*

Keywords: Design of Experiments, Space-filling, Constraints, Maximin, MiniMax.

1. Introduction

Design of experiments (DoE) creates an essential part in experimentations, surrogate modeling and many other fields such as sensitivity analysis or probability computations. The design of experiments consists of a set of design points whose coordinates correspond to combinations of input parameters values. The aim of the design of experiments is to gain maximal information about the solved system with a minimal number of executions. Therefore the basic requirements placed on the experimental design are orthogonality and space-filling of the design domain.

The shape of the design domain is firstly affected by presence of limiting conditions (constraints) and secondly by their form. In case of no constraints, the design domain has a regular shape, so-called hypercube. One special case of constrained design space is a mixture experiment. Here a sum of input variables (relative proportions of components) forms a unit volume or unit weight:

$$x_1 + x_2 + \ldots + x_n = 1, \quad 0 \le x_i \le 1, \quad i = 1, \ldots, n \tag{1}$$

Then the design domain is a simplex and the design points can be described by barycentric coordinates as shown in Fig. 1. The additional limiting conditions lead to the design domain in shape of a polytope.

Except of the shape of the design domain the method for generation of experimental design is influenced by other factors. The essential is the distinction between real (physical, chemical…) experiments and computer experiments called simulations. This contribution is aimed at designs for computer experiments. The experimental designs for real experiments are described in detail for example in (Montgomery, 2000). Methodology is influenced also by the difference between experiments with continuous parameters and experiments with discrete parameters. Here we focus on experiments with continuous parameters.

The quality of the designs can be evaluated by many criterions (Janouchová & Kučerová, 2013). Those aimed at orthogonality are for example a condition number or correlation coefficients. Criterions evaluating the space-filling are for example Audze-Eglais criterion, discrepancy, Euclidean Maximin distance or miniMax criterion. The last two criterions are used in this contribution.

* Ing. Eva Myšáková: Faculty of Civil Engineering, Czech Technical University in Prague, Thákurova 7, 166 29, Prague, CZ, eva.mysakova@fsv.cvut.cz

** Assoc. Prof. Ing. Matěj Lepš, PhD.: Faculty of Civil Engineering, Czech Technical University in Prague, Thákurova 7, 166 29, Prague, CZ, leps@cml.fsv.cvut.cz

Euclidean Maximin distance (*EMM*) is the shortest among all distances between the design points:

$$EMM = \min\{\ldots, L_{ij}, \ldots\}, \quad i = 1, \ldots, np, \quad j = (i+1), \ldots, np, \tag{2}$$

where *np* is a number of design points and *Lij* is the distance between points *i* and *j*. This criterion indicates undesirable proximity of points in the design. The higher the *EMM* value the better.

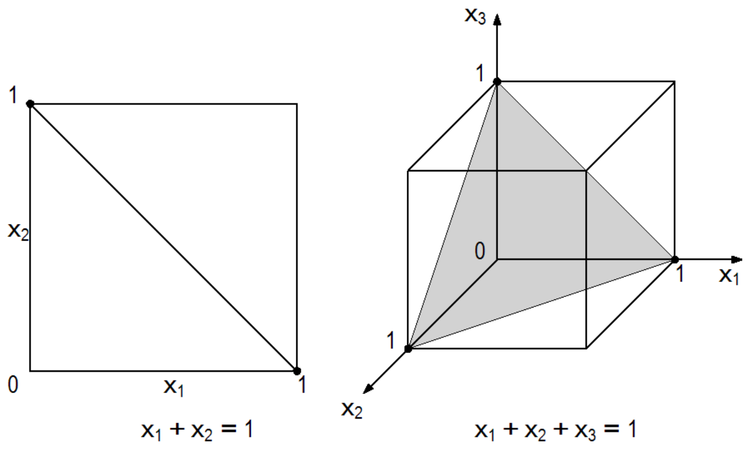

Fig. 1: The design domain for mixture experiment with two and three components.

Criterion miniMax (*mM*) corresponds to the *Largest Empty Sphere problem* (LES). The objective is to find the largest (hyper) sphere that includes no design point and whose center lies in the solved design domain. The value of the miniMax criterion is then equal to the radius of this largest sphere. MiniMax criterion serves for detection of unexplored areas inside the domain. The smaller the value is the better.

The paper follows (Myšáková, 2013) and is organized as follows: Chapter 2 presents several methods for generation of experimental designs in constrained domains. Two illustrative examples of mixture experiments used for comparison of the methods are described in Chapter 3 and Chapter 4 brings results and conclusions.

2. Methods for Design Generation

Methods for generation of the experimental design in generally irregular design space can be divided into two categories. The first one uses the bounding box – the regular domain (in shape of hypercube or hypercuboid) circumscribing the solved constrained domain. In this regular domain we can easier generate the design and then just extract the points lying in original irregular domain. Contrary, the second category is based on division of the constrained domain into simplices by Delaunay triangulation. In these simplices the generation of the points is also easier than in the whole irregular domain, so we can generate the design points in individual simplices and then create the design by union of these design points.

2.1. Generator of random points

The first method belongs to the second category mentioned above. The solved constrained domain is decomposed by Delaunay triangulation into simplices. Here the random points are generated with exponential distribution and then normalized (Devroy, 1986).The number of points generated in each simplex is given by its relative volume.

2.2. LHS on bounding box

Another method uses the bounding box of the domain. Here the LHS design is created and only points lying in solved irregular domain are extracted. Non-optimized and optimized LHS design can be used.

2.3. Removal of superfluous points

The third presented method involves removal of points from intentionally overcrowded initial designs. In each step one of points from the actual mutual closest pair is removed. There are two variants which differ in determination of a removed point. In the first variant the removed point is chosen from the pair randomly, in the second variant the point whose second shortest distance to other design points is removed.

2.4. Distmesh tool

The last method uses the Distmesh tool (Persson & Strang, 2004). It is a heuristic algorithm for generation of meshes for Finite Element Method (FEM). The quality meshes should have evenly distant nodes; therefore we can use this tool for creation of space-filling design of experiments. The algorithm is based on dynamical system of an expanding truss structure.

3. Applications

For comparison of presented methods two examples of mixture experiments were used. The first one is a three-component mixture (Snee, 1979). The corresponding design domain is a 2D irregular polygon with 6 vertices. The second example involves six-component mixture (Simon et al., 1997) with a 5D irregular polytop design domain with 51 vertices.

4. Results and Conclusions

The comparison of methods is shown in Figs. 2 and 3. The bottom graphs present the quality of resulting designs evaluated by criterions Maximin (*EMM*) and miniMax (*mM*). The top graphs show the corresponding time demands where the horizontal axis is *EMM* as in graphs below. The utopia (ideal) point lies in right bottom corner of both graphs.

The legend for the graphs is showed below:

- ■ the random points generator
- ● non-optimized LHS design in the bounding box
- ✡ optimized LHS design in the bounding box
- ◆ the removal – random point removed from the closest pair
- ★ the removal – “worse” point removed from the closest pair
- ▼ the Distmesh tool I
- △ the Distmesh tool II

Note: In methods with points removal several levels of overcrowding was tried.

The results clearly indicate a correlation between criterions *EMM* and *mM*. Optimization of one of them improves the design also in the terms of the other.

In 2D the dominance of the Distmesh tool is evident. On the other hand in higher dimensions the Distmesh tool cannot be practically used. Its time demands rise rapidly with the dimensions due to repeated Delaunay triangulations within the algorithm. More severe issue is that the quality of the resulting designs also deteriorates with grow of dimensions. In 5D the Distmesh tool even generates designs with duplicated points.

The random points generator is the fastest method but its designs have the worst quality in almost every case. Better results were achieved with the method using extraction of the points from LHS designs in the circumscribed domain. It was confirmed that usage of optimized LHS designs leads to better constrained designs; of course in cost of some additional time.

The method applying removal of points from overcrowded designs reached very good results. The more overcrowded the initial design is the better final design is created. But there are some limits for the level of overcrowding. In examples used in this paper the level of overcrowding larger than 30 does not worth the increasing time demands.

Finally, the results can be summarized as follows: in low dimensions the Distmesh tool is effective; in higher dimensions the usage of removal from the overcrowded designs can be recommended.

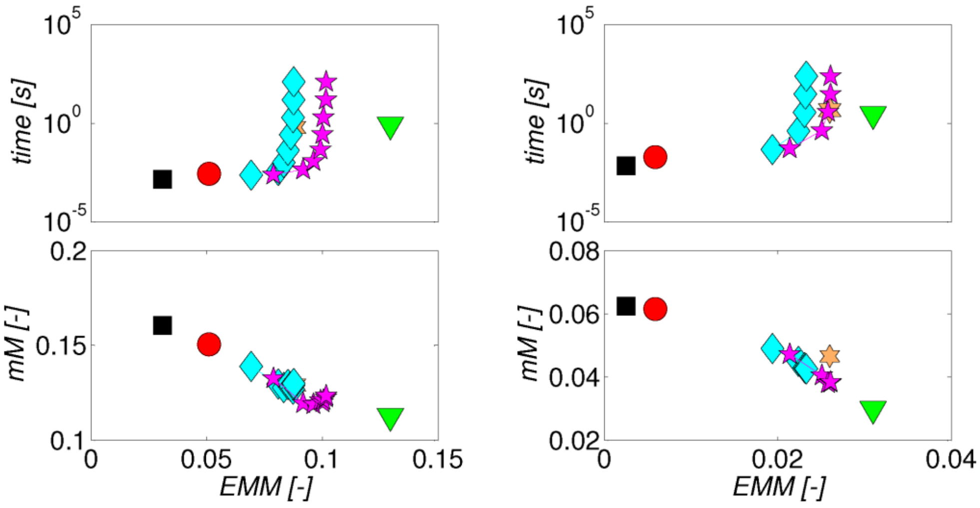

Fig. 2: Results of methods in example with three-component mixture.
Left: DoE with 10 design points; Right: DoE with 100 design points.

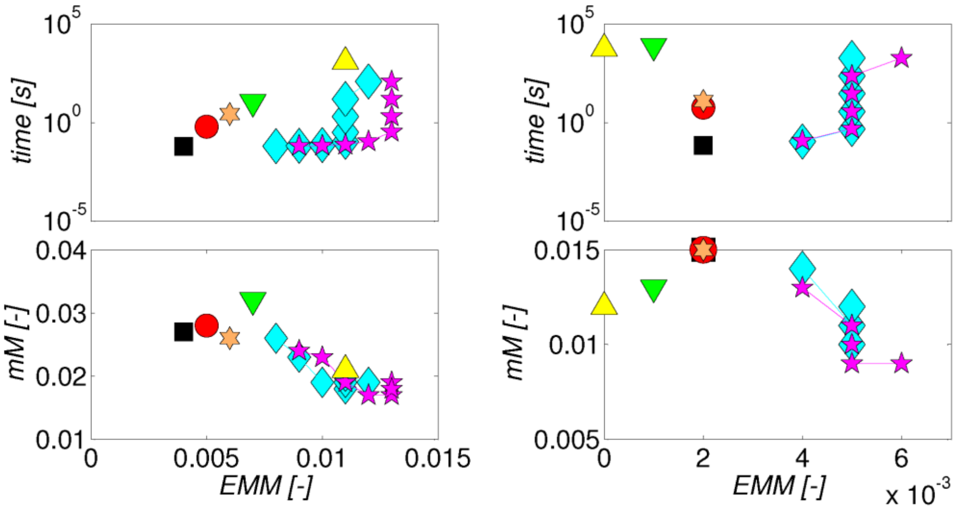

Fig. 3: Results of methods in example with six-component mixture.
Left: DoE with 10 design points; Right: DoE with 100 design points.

Acknowledgement

The authors gratefully acknowledge the financial support from the Grant Agency of the Czech Technical University in Prague, the grant SGS14/028/OHK1/1T/11.

References

Devroy, L. (1986) Non-uniform Random Variate Generation. Springer-Verlag.

Janouchová, E., Kučerová, A. (2013) Competitive Comparison of Optimal Designs of Experiments for Sampling-based Sensitivity Analysis. Computers & Structures, 124, pp. 47-60.

Montgomery, D. C. (2000) Design and Analysis of Experiments, 5th edition. Wiley.

Myšáková, E. (2013) Optimization of Uniformity of Computer Experiments for Constrained Design Spaces. Diploma Thesis. Faculty of Civil Engineering, Czech Technical University in Prague.

Persson, P. O., Strang, G. (2004) A Simple Mesh Generator in MATLAB. SIAM Review, 46, 2, pp. 329-345.

Simon, M. J., Lagergreen, E. S., Snyder, K. A. (1997) Concrete Mixture Optimization Using Statistical Mixture Design Methods. In Proceedings of the PCI/FHWA, New Orleans, Louisiana, pp. 230-244.

Snee, R. D. (1979) Experimental Designs for Mixture Systems with Multicomponent Constraints. Communications in Statistics - Theory and Methods, 8, 4, pp. 303-326.