Using sobol sequences for planning computer experiments

I N Statnikov, G I Firsov
Blagonravov Mechanical Engineering Research Institute of RAS, Moscow, 101990, Russian
E-mail: firsovgi@mail.ru

Abstract. Discusses the use for research of problems of multicriteria synthesis of dynamic systems method of Planning LP-search (PLP-search), which not only allows on the basis of the simulation model experiments to revise the parameter space within specified ranges of their change, but also through special randomized nature of the planning of these experiments is to apply a quantitative statistical evaluation of influence of change of varied parameters and their pairwise combinations to analyze properties of the dynamic system.

1. Introduction
Starting from the second half of the 20th century, when solving problems of designing modern dynamic systems, the research emphasis is shifted to mathematical modeling. At the same time, the data collection process becomes a complex problem both because of the possibility of obtaining them in large quantities, and the difficulties that arise for this reason in interpreting the results of the computational experiment. In this case, according to R. Fisher, the first problem can be solved with the help of planning experiments, and the second - with the help of the logic of scientific inference. The method of the planned LP search (PLP-search) developed by the Blagonravov Mechanical Research Institute of the Russian Academy of Sciences [1] makes it possible to solve both of the above-mentioned problems at the same time.

2. Procedure of PLP-search
Developed by I.M. Sobol uniformly distributed sequences [2] of binary-rational pseudo-random numbers, originally intended for the calculations of multidimensional integrals, began to be used in search procedures, realizing the ideology of "blind" search, as is typical for the methods of the Monte Carlo family. The structure of these sequences allowed the combinatorial way [1] to construct in the J-dimensional space of variable parameters of the lattice (the matrix of planned experiments), in calculating the functions under study it was possible to obtain answers to the following questions:
- which of the varied parameters with a given (required) probability have a significant effect on the values of the function (the quality criterion of the system, in other words, to statistically evaluate the derivative of the quality criterion for each variable parameter;
- on the chosen metric between the current value of the quality criterion and its extreme value, known either as a priori or determined by the results of computational experiment, to determine the areas of concentration of the best (achievable) quality criterion values;
construct a Pareto set in the multidimensional space quality criterion or, if a compromise scheme is specified between the quality criterion, select a subdomain containing the largest concentration of compromise solutions in the J-dimensional parameter space;

- approximate the selected subdomains of quality criterion values by some types of regression dependencies.

The essence of PLP search is that the method combines ideas of discrete review of the space of analyzed parameters and the theory of planning of mathematical experiments [3]. The PLP-search method, probabilistic in nature, adequately corresponds to modern principles of simulation and heuristic modeling. Method allows you to build a matrix of planned experiments and the simulation model simulations to implement quasi-uniform in probability view space varied parameters within specified ranges of their changes, and also as a result of the special nature of these experiments planning randomized apply quantitative statistical) evaluation of the impact of variable parameters and their paired combinations to analyze properties of investigated (projected) object. If this turns out to be possible when installed (assigned) probability estimates received as result analysis of the properties of an object to its analysis of the same properties, but according to the "medium" from each variable parameter, i.e., to analyze the "one-dimensional" dependencies that psychologically and practically becomes very useful to the researcher (designer), so as to significantly overcome the curse of dimensionality "on Bellman. At the same time, the presence of the dependencies mean values of quality criteria, as well as the ability to determine the influence of parameters on quality criteria with the required confidence probability, largely solve the problem of integrating a huge numerical information obtained in the experiment, in clear and quantifiable logical analysis (up to "common sense") features.

At the same time, by constructing approximation models of criteria depending on the variable parameters, it is possible to estimate the sensitivity of the criteria on average for these parameters. The effectiveness of the plans for experiments in the PLP search is determined not only by the possibility of using them in dispersion analysis. These plans prove to be effective both in constructing regression dependencies, and generally in regression analysis, both in the computational aspect and in terms of a number of optimality criteria for these plans.

In particular, for the case of linear, quadratic and cubic regression obtained the value of the determinant of the Fisher information matrix and the limits of change of dispersion of the predicted values. Formula analysis showed that the increasing number of experiments in the series, a number of series of experiments and number of varied parameters values of the Fisher information matrix determinant grow, thereby making the plans similar properties to orthogonal; all correlation coefficients regression models for each of the regressions have good convergence to zero. For example, in the case of ten series of experiments, eight experiments in series and three varied parameters correspond to linear, quadratic and cubic regression constitute 22500, 27000 and 18750. Any of a series built plan experiments will be D-optimal.

3. Examples of PLP-search
Below are some of the examples of PLP-search in carrying out the tasks of designing various dynamic systems [4]. Here MM - mathematical model, NODE is a nonlinear ordinary differential equation, LODE is a linear differential equation, PDE - partial differential equation.

1) Rotary indexing table with hydro-mechanical drive. MM: 3 NODE of the second order. J = 9. K = 3. Result: found scope for compromise solutions, the volume of which amounted to ~ 0.2% of the originally specified.

2) High accuracy pressure pneumatic regulator. MM: 4 NODE of the second order. J = 4. K = 1. Result: found better solutions area with up to 0.5% of the originally specified.

3) Pneumatic shaking machine [5-7]. MM: 4 NODE of the second order. J = 8. K = 1. Result: 4 influential parameter; the selection amounted to 5% of the originally specified.

4) Multi-circuit planetary gear train. MM: 23 LODE second order nonhomogeneous J = 25. K = 6. Result: 8 parameters, simultaneously affected all criteria.
5) Sewing machine. MM: 5 LODE inhomogeneous second order. J = 6. K = 5. Result: in areas built regression dependences of natural frequencies of parameters mm.

6) Resonant converter for ship shafting. MM: 2 NODE of the second order. J = 6. K = 1. The result: two influential parameter; the criterion value improved 5.2 times compared with the same in the source area.

7) Transmission main drive working stand rolling mill. MM: 5 NODE of the second order. J = 5. K = 5. Result: found an area of compromise, representing ~ 3.5% of the originally specified.

8) Heat exchanger. MM: 1 PDE. J from 8 to 18. K = 4. The result: defined for each J significant parameters and built area of compromise.

4. Conclusion
Thus, the PLP-search method not only makes it possible, on the basis of simulation simulations, to perform a quasi-uniform scanning of the parameter space in given ranges of their changes, but also as a result of a special randomized nature of the planning of these experiments, to use quantitative statistical estimates of the effect of variation of variable parameters and their pairwise combinations on analyzed properties of the dynamical system under consideration.

Proving the effectiveness of PLP search in comparison with the "blind" way of searching for extremes, it is not suggested to discard the latter. However, it is obvious that PLP search not only can help speedy search for extremes if it is required, but also provides information on the influence of variable parameters and, what is also important, contributes to issues related to the study of the space of variable parameters (in particular, the results of the PLP search can help to choose an effective compromise scheme). So, it can be argued that if the decision is made to invest in gate the formulated problem first in a discrete way (which is useful even in the case when it is possible to obtain analytical dependencies, but very complex ones), PLP search seems to be a very effective method of computer technologies in the sense of the previously mentioned second component.

References
[1] Statnikov I N, Andreenkov E V 2006 PLP-search-heuristic method of solving tasks of mathematical programming. (Moscow, IIP MGUDT) (In Russian).
[2] Sobol I M 1998 Math. Comput. Simulation 47 103.
[3] Montgomery D C 2001 Design and Analysis of Experiments (Hoboken, NJ, John Wiley and Sons).
[4] Statnikov I N, Firsov G I 2016 Vestnik of Scientific and Technical Development 3(103) 38. (In Russian).
[5] S.Yu. Misyurin, V.I. Ivlev, V.M. Bozrov, and A.P. Nelyubin Parameterization of an Air Motor Based on Multiobjective Optimization and Decision Support // ISSN 105-6188, Journal of Machinery Manufacture and Reliability, 2013, Vol. 42, No. 5, pp. 353–358.
[6] G.V.Kreinin and S.Yu.Misyurin. Phased Synthesis of a Mechatronic System ISSN 1028_3358, Doklady Physics, 2014, Vol. 59, No. 11, pp. 539–543.
[7] S. Yu. Misyurin and G. V. Kreinin. On Choosing the Drive Type for the Power Unit of a Mechatronics System. Journal of Machinery Manufacture and Reliability, 2015, Vol. 44, No. 4, pp. 305–311.