WORKING PAPER

OPTIMAL DESIGN OF EXPERIMENTS:
NUMERICAL METHODS

V.V. Fedorov

October 1986
WP-86-055

IIASA
International Institute
for Applied Systems Analysis

Optimal Design of Experiments: Numerical Methods

V.V. Fedorov

October 1986
WP-86-55

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

Acknowledgement

I am grateful to Prof. R. Munn for encouraging discussions and constructive remarks.

PREFACE

Optimal experimental designs became a rather efficient tool in applications. There are numerous catalogues of optimal designs for some standard situations, for instance, when the response function is a multidimensional power polynomial or trigonometrical series. In cases when the response function is not of the approximation type but its structure is based on some "physical" assumptions, one can not hope to find an optimal design suitable for specific conditions, and usually one needs to apply numerical methods to find this design.

This short survey is devoted to numerical methods for the construction of optimal designs for experiments when a system "object under investigation – process of observation" is described by the model (see (1) in the paper) which linearly depends upon unknown parameters and contains an additional stochastic component (usually referred to as an error of observation, but it could also reflect the stochastic nature of the object). The objective of an experiment is to estimate the unknown parameters and an optimal design (for instance, optimal location of observations) has to provide the smallest errors of estimates. Usually, these errors are characterized by the variance–covariance matrix (or some functions of it) which are the inverse of the so–called information matrix used in the paper.

Sections 1 and 2 contain some general information on numerical methods. The subsequent sections deal with more specific situations. For instance, section 3 and 6 (containing some new results by the author) deal with spatially distributed observations and could be especially useful in optimization of monitoring systems, a very acute problem for many environmental studies. In Section 4, one can find algorithms which can be used for the design of experiments related to remote sensing of the earth's atmosphere by a satellite radiometer (findings of optimal frequency bands or "windows").

Prof. M. A. Antonovsky

CONTENTS

Optimal Design of Experiments: Numerical Methods

V.V. Fedorov

1. Introduction

In this paper numerical approaches for the construction of optimal designs will be considered for experiments described by the regression model

$$y_i = \vartheta^T f(x) + \varepsilon_i, i = \overline{1,N} \tag{1}$$

where $f(x)$ is a given set of basic functions, $x \in X$, and X is compact; at least some of the variables x can be controlled by an experimenter, $\vartheta \in R^m$ are estimated parameters, $y_i \in R^1$ is the i-th observation, and $\varepsilon_i \in R^1$ is the random error, $E[\varepsilon_i] = 0, E[\varepsilon_i \varepsilon_j] = \delta_{ij}$. In practice, technically more complicated problems could be faced (for instance, y_i could be a vector or errors could be correlated) but usually the methods are straightforward generalizations of the methods developed for problem (1).

The most elegant theoretical results and algorithms were created for a continuous (or approximate) design problem when a design is considered to be a probabilistic measure defined on X, and an information matrix is defined by an integral $M(\xi) = \int f(x)^T(x)\xi(dx)$. In this case, the optimal design of the experiment turns out to be the optimization problem in the space of probability measures:

$$\xi^* = \underset{\xi}{Arginf}\ \Phi[M(\xi)], \int_X \xi(dx) = 1, \tag{2}$$

where Φ is the objective function defined by an experimenter.

The first ideas on numerical construction of optimal designs can be found in the pioneer works by Box and Hunter (1965) and Sokolov (1963), where some sequential designs were suggested. These procedures can be considered as very particular cases of some iterative procedures for optimal design in construction, but nevertheless they implicitly contain the idea that one can get optimal design through improving intermediate designs by transferring a finite measure to some given point in X at every step of the sequential design.

This idea was developed and clarified by many authors and the majority of algorithms presented in this survey (which does not pretend to be a historical one) are based on it.

2. First–order iterative procedures.

It will be assumed that

(a) the functions $f(x)$ are continuous on compact X,

(b) $\Phi(M)$ is a convex function,

(c) there exists q such that

$$\{\xi:\Phi[M(\xi)]=q<\infty\}=\Xi(q)\neq\phi,$$

and

(d) for any $\xi\in\Xi(q)$ and any other $\bar{\xi}$

$$\Phi[(1-\alpha)M(\xi)+\alpha M(\bar{\xi})]=\Phi[M(\xi)]+\alpha\int_X\varphi(x,\xi)\bar{\xi}(dx)+o(\alpha). \tag{3}$$

If these assumptions hold, then the following iterative procedure will converge to an optimal design:

$$\xi_{s+1}=(1-\alpha_s)\,\xi_s+\alpha\;\xi\;(x_s), \tag{4}$$

where $\xi(x_s)$ is a design with the measure totally concentrated at the point x_s,

$$x_s=Arg\min[\varphi\;(x_s^+,\xi_s),-\varphi(x_s^-,\xi_s)], \tag{5}$$

$$x_s^+=Arg\min_{x\in X}\varphi(x,\xi_s),x_s^-=Arg\max_{x\in X_s}\varphi(x,\xi_s),$$

X_s is the supporting set of the design ξ_s, $\alpha_s=\gamma_s$, when $x_s=x_s^+$, and $\alpha_s=-min\,[\gamma_s\,,p_{si}/\,(1-p_{si})]$, p_{si} is a measure for point x_{si} of design ξ_x.

To provide weak convergence, the sequence $\{\gamma_s\}$ has to satisfy, for instance, the following condition: $\gamma_s\to 0$ and $\Sigma\gamma_s\to\infty, s\to\infty$. In addition, some other alternatives for the sequence $\{\gamma_s\}$ can be found in Ermakov, 1983; Fedorov, 1972; Fedorov and Uspensky, 1975; Fedorov, 1981; Silvey, 1982, Wu and Wynn, 1978. The iterative procedures (4), (5) comprise practically all the first–order methods widely discussed in the statistical literature since the late nineteen sixties. It should be pointed out that the iterative procedure can be realized in practice if the optimization problem (5) is not very difficult from a computational point of view, i.e., if the dimension of X is not too high. It is especially difficult to work with cases when the controllable variables belong to some functional space (see section IV).

There is a simple idea behind the iterative procedure (4), (5). If one wishes to move along the "best" direction

$$\xi_{s+1}=Arg\min_{\xi}\Phi[(1-\alpha_s)M(\xi_s)+\alpha_s M(\xi)] \tag{6}$$

then for sufficiently small α_s (see (3)):

$$\min_{\xi}\Phi[(1-\alpha_s)M(\xi_s)+\alpha_s M(\bar{\xi})]=\Phi[M(\xi_s)]+\alpha_s\min_{\xi}\int_X\varphi(x,\xi_s)\bar{\xi}(dx)$$

$$=\Phi[M(\xi_s)]+\alpha_s\min_{x\in X}\varphi(x,\xi_s).$$

Therefore

$$\bar{\xi}_s^* = Arg\min_{\xi} \int_X \psi(x, \xi_s)\bar{\xi}(dx)$$

can be chosen from the set of point measures $\xi(x), x \in X$ and has to be concentrated at the point:

$$x_x^+ = Arg\min_{x} \in X \psi(x, \xi_s)$$

The same idea is behind "deleting" some points from design ξ_s. These points $\{x_s^-\}$ are "worst" in the sense (6).

For fulfillment of this fact, assumption (d) is crucial. The majority of optimality criteria used in practice satisfy this assumption. But for some quite natural criteria, for instance,

$$\Phi[M(\xi)] = f^T(x_o)M^-(\xi)f(x_o)$$

where x_o is given, Φ is the variance of $\hat{\vartheta}^T f(x_o)$ and "-" stands for pseudoinversion, formula (3) is not generally valid. One can still use the iterative procedure (4), (5) applying to some regularized version of the initial problem:

$$\Phi_\rho[M(\xi)] = \Phi[(1-\rho)M(\xi) + \rho M(\xi_o)],$$

where $M(\xi_o)$ is regular matrix ($M(\xi_o) \neq 0$). Then

$$\lim_{s \to \infty} \Phi[M(\xi_s)] - \Phi[M(\xi^*)] \leq \rho\{\Phi[M(\xi_o)] - \Phi[M(\xi^*)]\}. \qquad (7)$$

To adjust the iterative procedure (4), (5) to particular optimality criteria, the following formulae can be useful:

$$\psi(x, \xi) = f^{T(x)}\dot{\Phi}(\xi)f(x) - tr\,\dot{\Phi}(\xi)\,M(\xi), \quad \dot{\Phi} = \frac{\partial \Phi}{\partial M}$$

where the existence of a corresponding derivative is assumed,

$$\frac{\partial M^b}{\partial M_{\alpha\beta}} = \sum_{a=0}^{b-1} M^a E_{\alpha\beta} M^{b-a-1}, \quad E_{\alpha\beta\gamma e} = \delta_{\alpha\gamma}\delta_{\beta e},$$

$$\frac{\partial M^{-b}}{\partial M_{\alpha\beta}} = -\sum_{a=0}^{b-1} M^{-a-1} E_{\alpha\beta} M^{-b+a},$$

$$\dot{\Phi} = -M^{-1}\frac{\partial \Phi}{\partial M^{-1}}M^{-1}, \quad \frac{\partial \ln|M|}{\partial M} = -M^{-1}, \quad \frac{\partial tr AM^{-1}}{\partial M} = -M^{-1}AM^{-1},$$

$$M^{-1}(\xi_{s+1}) = (1-\alpha_s)^{-1}\left[I - \frac{\alpha_s M^{-1}(\xi_s)f(x)f^T(x)}{1-\alpha_s+\alpha_s d(x, \xi_s)}\right]M^{-1}(\xi_s),$$

$$|M(\xi_{s+1})| = (1-\alpha_s)^{m-1}[1-\alpha_s+\alpha_s d(x, \xi_s)]|M(\xi_s)|,$$

$$d(x, \xi_s) = f^T(x)M^{-1}(\xi_s)f(x).$$

The convergency rate of the above mentioned algorithms decreases in the vicinity of the optimum. As in general optimization theory, attempts were made to develop second order methods. These methods are based on quadratic approximations of the function $\Phi[M]$ and it is necessary to assume the existence of derivatives $\partial\Phi/\partial p_i, \partial^2\Phi/\partial p_i \partial p_j$, where p_i is the weight for supporting point x_i. Second–order algorithms have at least two features which are handicaps for their use in practice: first, at every step it is necessary to invert the matrix $\{\partial^2\Phi/\partial p_i \partial p_j\}$, and second, all existing modifications can handle only the discrete operability region X (see Ermakov (ed), 1983 Ch 4, Wu, 1978).

In the late seventies, some attention was paid to algorithms which work in the space of information matrices $M(X)$; they are computationally effective if one can easily find the mapping $\Xi(X) \to M(X)$ (see, for example, Gribik and Kortanek, 1977). But usually it is very difficult to realize this mapping numerically.

3. Construction of optimal designs under constraints.

In (2) there is only one constraint $\int \xi(dx) = 1$. If one considers additional constraints, say $\int \psi(x)\, \xi\,(dx) \le c$, where $\psi(x)$ is a vector of given functions $(c, \psi \in R^k)$, then the iterative procedure becomes technically more complicated, although based on the same ideas, see (Fedorov and Gaivorovsky, 1984). In this case, (6) became equivalent to the following optimization problem

$$\bar{\xi}_s = Arg\min_{\bar{\xi}} \int_X \varphi(x, \xi_s)\bar{\xi}(dx), \tag{8}$$

subject to $\int_X \psi(x)\bar{\xi}(dx) \le c$.

Due to the classical theorem of Caratheodory, design $\bar{\xi}_s$ can be found within the set of designs containing no more than $k+1$ supporting points. The optimization problem (8) is essentially more complicated than (5). From a computational point of view, the dual problem (see Karlin and Studden, 1966, ch. XII):

$$\max_{u \in U} \min_{x \in X} [\varphi(x, \xi_s) + u^T \psi(x)] \tag{9}$$

where $U = \{u : u_i \ge 0, i = \overline{1,k}\}$, can be useful to define the location of the supporting points of $\bar{\xi}_s$. They have to coincide with the solutions $x_1^*, x_2^*, \ldots$ of (9). The corresponding measures can be found by any linear programming algorithm.

In some applications (see Wynn, 1982), designs have to be restricted in the following sense

$$\int_A \xi(dx) \le \int_A \xi_0(dx) \quad \text{for all } A \subset X \tag{10}$$

where $\int_X \xi_0(dx) = c$, $c \ge 1$, $A \subset X$ and a measure ξ_0 is atomless (see Karlin and

Studden, 1966, p. 233). In this case (6) leads to the following problem

$$\bar{\xi}_s = Arg\min_{\bar{\xi}} \int_X \varphi(x, \xi_s)\bar{\xi}(dx), \tag{11}$$

where a probability measure $\bar{\xi}_s$ has to satisfy (10).

It is evident that $\bar{\xi}_s$ has to coincide with ξ_0 on any subset of X, where $\varphi(x, \xi_s) \leq 0$ and has to be equal 0 otherwise (compare with theorem 1 from Wynn, 1982). Computationally, the search for these sets can be realized through discretization of X. The idea of the iterative procedure (4), (5) will apply once again if one will additionally delete from design ξ_s, some sets where $\varphi(x, \xi_s) > 0$. Thus

$$\xi_{s+1} = (1-\alpha_s)\xi_s + \alpha_s [\xi_0(E_s, dx) - \xi_0(D_s, dx)], \tag{12}$$

where E_s is the set of new included points and D_s is the set of deleted points.

The procedure similar to (12) (but without the operation of deleting) was considered by Gaivorovsky (1985), and its weak convergency was proven under rather mild conditions. Usually, deleting "bad" points essentially improves the quality of the iterative procedures (compare with the traditional case, Atwood, 1973, Fedorov and Uspensky, 1975).

Let $\Xi(\xi_0)$ be a class of probability measures ξ with supporting sets $A \subset X$, and $\xi(dx) = \xi_0(dx)$, when $x \in A$ and equal to 0 otherwise. For any design problem (2), (10), there exists an optimal design $\xi^* \in \Xi(\xi_0)$, see, for instance, Wynn, 1982, whose results have their origin in the classical moment space theory, particularly in the Liapunov Theorem, see Karlin and Studden, 1966, Ch. VIII). Therefore, it is reasonable to demand that $\xi_s = \xi(A_s, dx) \in \Xi(\xi_0), s = 1,2,\ldots$ Iterative procedure (12) does not satisfy the latter demand. Instead of this type of iterative procedure (which repeats the idea of (4), (5)), one can apply an iterative procedure of the exchange type:

$$\xi_{s+1} = \xi_0(A_{s+1}, dx) = \xi_0(A_s, dx) + \xi_0(E_s, dx) - \xi_0(D_s, dx), \tag{13}$$

where

$$\int_{E_s} \xi_0(dx) = \int_{D_s} \xi_0(dx) = \delta_s, \; A_s \cap E_s = 0, \tag{14}$$

$$D_s \subset A_s, \; E_s \cap D_s = 0.$$

If

(e) derivatives $\dot{\Phi}$ exist,

(f) ξ_0 has a continuous density $\mu_0(x)$,

and assumptions (a)–(c) hold, then for sufficiently small δ_s, the approximation

$$\Phi[M(\xi_{s+1})] = \Phi[M(\xi_s)] + [\gamma(x_s^+, \xi_s) - \gamma(x_s^-, \xi_s)]\delta_s + o(\delta_s), \tag{15}$$

where $x_s^+ \in E_s$, $x_s^- \in D_s$ and $\gamma(x, \xi) = f^T(x)\dot{\Phi}(\xi)f(x)$ can be used. If X is

covered by the grid X_δ with density proportional to $\mu_0(x)$, then x_s^+ and x_s^- can coincide with its nodes and E_s , D_s with some cells of this grid.

From (15), it is clear that to provide approximately the steepest descent on every step of the discretized version of procedure (13), one has to find

$$x_s^+ = Arg \min_{x \in X_\delta \setminus A_{\delta s}} \gamma(x, \xi_s) \text{ and } x_s^- = Arg \max_{x \in A_{\delta s}} \gamma(x, \xi_s), \tag{16}$$

where $A_{\delta s}$ is a discrete analogue of A_s. It is worthwhile to point out that for the discretized version of the iterative procedure, one can use a recursive formula for M_{s+1}^{-1} (see (17)) to simplify calculations. Complementing assumptions (a)–(f) by assumptions:

(g) for any design ξ with $\Phi[M(\xi)] \le Q < \infty$ and any C

$\xi_0\{A : \gamma(x, \xi) = C\} = 0$,

(h) $\delta_s \to 0$, $\sum_s \delta_s \to \infty$

the weak convergency:

$$\lim_{s \to \infty} \Phi[M(\xi_s)] = \min_{\xi \in \Xi(\xi_0)} \Phi[M(\xi)] \tag{17}$$

can be proven.

Result (17) is based on the fact that the fulfillment of the inequality:

$$\max_{x \in X^*} \gamma(x, \xi^*) \le \min_{x \in X \setminus X^*} \gamma(x, \xi^*)$$

is a necessary and sufficient condition for a design ξ^* to be optimal (X^* is a supporting set of ξ^*).

4. Optimal designs when controls belong to a functional space.

This case will be illuminated here by a rather specific example which nevertheless reflects the major difficulties.

Let $f(v) \in R^m$ and $x = \int_V f(v) h(v) dv$, where $h(v)$ can be controlled, $h(v) \in H$. If one manages to construct the mapping $X(H) \subset R^m$ of the set H, then all approaches discussed in the previous sections can be used without any alterations to find an optimal design ξ_x^* on X. The problem to be faced afterwards is the construction of an inverse mapping $X \to H$ to convert ξ_x^* to some design ξ_H^* defined on H. The latter problem is beyond the scope of this paper and its discussion can be find in Kozlov, 1981, Ermakov (ed), 1983, Ch.7. For the case discussed in section 2, the situation is slightly simpler because of the Equivalence Theorem (see, for instance, Ermakov (ed), 1983 Ch.2); thus only boundary points $\bar{X}(H)$ of $X(H)$ are needed for optimal design construction. Unfortunately, the numerical construction of $\bar{X}(H)$ happens to be sometimes a very difficult problem and it could be more efficient to work in the original space H.

To be more specific, let us assume that $v \in V \subset R^1$ and $0 \le h(v) \le 1$ and restrict ourself to the design problem in section 2. The most straightforward approach consists of discretization of V and approximation of $h(v)$ by same piecewise function. Under rather mild conditions, it is possible to prove that there exists optimal design which supporting points belong to the set $\bar{H} = \{h(v): h(v)[1-h(v)]=0, v \in V\}$, see, for instance Fedorov, 1986. If V is discretized by a grid with elements Δ_j, then the simplest version of procedure (4), (5) (without "deleting" operation) converts to the following one:

a) $\xi_{s+1} = (1-\alpha_s)\xi_s + \alpha_s \xi(h_s)$;

b) steps for finding h_s :

-collect all Δ_j which negatively contribute to the sum

$$\varphi(\xi_s) = \sum_{j,j'} F_j^T \dot{\Phi} F_{j'},$$

where $F_j = \int_{\Delta_j} f(v) dv$ (usually $F_j \approx f(v_j)\Delta_j$).

- put $h_s(v)=1$, $v \in \Delta_j$ if Δ_j was chosen in the previous stage, otherwise $h_s(v)=0$,
- the fulfillment of the inequality $\varphi(\xi_s) < tr\, \dot{\Phi}(\xi_s) M(\xi_s)$ tests that h_s can be used for $\xi(h_s)$.

This iterative procedure guarantees that

$$\lim_{s \to \infty} \Phi[M(\xi_s)] = \Phi[M(\xi_\Delta^*)],$$

where ξ_Δ^* is an optimal design for the discretized design problem and can be called a Δ - optimal design. When $a \le v \le b$ and functions $f(v)$ constitute a Tchebycheff system over the open interval (a,b), where a and b are possibly infinite, then the rather effective iterative procedure can be used for optimal design construction. The idea of this procedure is based on the following result (see, for instance, Fedorov, 1986).

Let $h(v) \in \bar{H}$ and let I be the number of separate nondegenerate intervals where $h(v)=1$ with the special convention that an interval whose closure contains point a or b, is counted as 1/2. For any point $x \in X$, I^* stands for the least possible I. Then a necessary and sufficient condition that x belongs to the boundary of X is that $I^* \le (m-1)/2$. Moreover, every boundary point corresponds to a unique $h(v)$ with $I(x) = I^*(x)$.

Let now $\bar{v} = (v_1, \ldots, v_{m-1})$, where $a \le v_1 \le \cdots \le v_{m-1} \le b$. According to the previous result, there exist optimal designs with all supporting points (in the operability region H) which have the following structures:

$$\bar{h}(v) = 1, v \in (a, v_1); 0, v \in (v_1, v_2); 1, v \in (v_2, v_3); \cdots$$

and $\underline{h}(v) = 1 - \bar{h}(v)$.

That fact allows for modification of the iterative procedure (4), (5) (without deleting "bad" points) to the procedure with maximization in space, with dimension less than or equal to $(m-1)$, where m is a number of basic

functions:

$$\xi_{s+1}=(1-\alpha_s)\xi_s+\alpha_s\xi(h_s),$$

$$h_s=Arg\min_{\gamma,\bar{v}}\Phi[x_\gamma(\bar{v}),\xi_s],$$

where $a\le v_1\le \ldots \le v_{m-1}\le b, x_1(\bar{v})=\int_a^b f(v)\,\bar{h}(v)dv$ and $x_2(\bar{v})=\int_a^b f(v)\underline{h}(v)dv$

The design problem considered in this section comprises the major difficulties which can be met when X is a functional space. Other examples can be found in Mehra, 1976, Pazman, 1986. These authors use mainly the same ideas surveyed here. In concluding this section, it would be worthwhile to notice that parametrization of controls (a rather standard method in optimal control theory), e.g. linear approximation $\delta^T q(v)$ of $h(v)$ in our example, could be a useful tool allowing one to convert the original problem to a finite dimension design problem.

5. Discrete designs

To construct optimal discrete (or exact) designs, a number of exchange type algorithms can be used (for detailed information, see Cook and Nachtsheim, 1980; Johnson and Nachtsheim, 1983; Steinberg and Hunter, 1984).

The idea of the simplest algorithm (originated by Mitchell, 1974) can be formulated in the following way:

After the s-th step there is a design $\xi_{Ns}=\{x_{1s},\ldots,x_{Ns}\}$, where some supporting points can coincide. This design is complemented by k points:

$$x_{N+j,s}^{+}=Arg\min_{x_j\in X}\Phi[M(\xi_{N+j-1,s}+x)],j=\overline{1,k} \tag{18}$$

Then the same number of points:

$$x_{N+k-1,s}=Arg\min_{x_l\in X_s}\Phi[M(\xi_{N+k-1,s}-x_l)],l=\overline{1,k}, \tag{19}$$

are deleted and one arrives at the new design $\xi_{N,s+1}$ containing N observations. The notation ξ_K+x (or $-x$) means that a point is included in (or excluded from) design ξ_K, X_s comprises all of the different supporting points of the design from the previous stage.

In practice, the excursion length k is usually rather modest (1-3) and there are no indications that an increase could be useful. Iterative procedure (18), (19) are computationally simple and often lead to very good results, especially when one faces discrete X, for example, $x_\alpha=\pm 1$.

In the iterative procedures (18), (19), the deletion and complementary steps are separated. If we unite them (Fedorov, 1972), then we arrive at the following iterative procedure (with excursion length 1):

$$\xi_{N,s+1}=Arg \min_{x^+\in X, x^-\in X_s} \Phi[M(\xi_{Ns}+x^+-x_j^-)], \tag{20}$$

where X_s is the supporting set of ξ_{Ns}. This procedure demands $N/2$ times more calculations at every step than (18), (19), but in most cases it gives better final results, see Johnson and Nachtsheim, 1983. The above minimization problem is equivalent to coordinate wise minimization of $\Phi[M]$:

$$\min_j \min_{x_j\in X} \Phi[M(x_{1s},\ldots,x_{js},\ldots x_{Ns}] \tag{21}$$

if one starts the numerical optimization in (21) with $x_j^0=x_{js}$.

A similar choice of an initial point is appropriate in many optimization problems, but not in the optimal design of experiments when the objective function usually has a large number of local minima along the variation of x_j, and (21) will lead to the local minimum closest to x_{js}. The application of (20) helps to approach the global minimum by explicit forcing of x^+ to be away of x_j^-. Procedures (18), (19) or (20) become a practical tool when one manage to find a simple formula for calculation of increments for $\Phi[M]$ at every stage. For the majority of widely used criteria (D-criterion, linear criteria and so on) these formulas can be found in the above cited publications (see also section 2).

In spite of the rather long history of the numerical procedures discussed above, their convergence properties are not well known except for numerous empirical results. It is not a problem, for instance, to prove the convergence of (20) to some design better than an initial one but it has not yet been proven that the limit design has to be optimal.

REFERENCES

Atwood, C.L. (1973) Sequences Converging to D-Optimal Designs of Experiments, Ann.Stat., Vol.1:342-352.

Box, G.E.P. and Hunter, W (1985) Sequential Design of Experiments for Nonlinear Models. Proc. IBM Scientific Computing Symp. Statistics, 113-120.

Cook, R.D. and Nachtsheim, C.J. (1980) A Comparison of Algorithms for Constructing Exact D-Optimal Designs, Technometrics, Vol.22:682-688.

Ermakov, S.M. (ed.) (1983) Mathematical Theory of the Design of Experiments (in Russian), p.386.

Fedorov, V.V. (1972) Theory of Optimal Experiments, Academic Press, N.Y., p. 292.

Fedorov, V.V. and Uspensky, A.B. (1975) Numerical Aspects of Design and Analysis of Experiments, Moscow State University, Moscow (in Russian), p. 167.

Fedorov, V.V. (1981) Active Regression Experiments in Mathematical Methods of Experimental Design, ed. Penenko, V.V., Nauka (in Russian), p.19-73.

Gaivoronsky, A. and Fedorov V. (1984) Design of Experiments under Constraints, WP-84-8, IIASA, Laxenburg, Austria, p. 11.

Gaivoronsky, A. (1985) Stochastic Optimization Techniques for Finding Optimal Sub-measures, WP-85-28, IIASA, Laxenburg, Austria, p.52.

Gribik, P.R. and Kortanek, K.O. (1977) Equivalence Theorems and Cutting Plane Algorithms for a Class of Experimental Design Problems, SIAM J. Appl. Math., Vol32:232-259.

Johnson, M.E. and Nachtsheim, C.J. (1983) Some Guidelines for Constructing Exact D-Optimal Designs on Convex Design Spaces, Technometrics, Vol.25:271-277.

Karlin, S. and Studden, W.J. (1966) Tchebycheff Systems: with applications in Analysis and Statistics, Wiley and Sons, N.Y., p. 578.

Kozlov, V.P. (1981) Design of Regression Experiments in Functional Spaces, in Mathematical Methods of Experimental Design, ed. Penenko, V.V., Nauka (in Russian), p.74-101.

Mehra, R.K. (1976) Synthesis of Optimal Inputs for Multiinput- Multioutput Systems with Process Noise, in "System Identification: Advances and Case Studies" ed. D.K. Mehra, D.G. Lainiotis, Academic Press, N.Y. pp. 211-250.

Mitchell, T.J. (1974) An Algorithm for the Construction of D-Optimal Experimental Designs, Technometrics, Vol16:203-210

Pazman, A. (1986) Foundations of Optimum Experimental Design, VEDA, D. Reidel Publishing Company, p.228.

Silvey, S.D. (1980) Optimal Design, Chapman and Hall, London, p.86.

Steinberg, D.M. and Hunter W.G. (1984) Experimental Design: Review and Comment, Technometrics, Vol.26:71-97.

Sokolov, S.N. (1963) Continuous Design of Regression Experiments. Teor. Verejatuost.; Primen., Vol.8:95-101 (a), 318-323 (b).

Wu, C.F. (1978) Some Iterative Procedures For Generating Nonsingular Optimal Designs, Commun. Statist. (Theor. Math.), Vol. A7(14):1399-1412.

Wu, C.F. and Wynn, H. (1978) The Convergence of General Step-Length Algorithms for Regular Optimum Design Criteria, Ann. Statist., Vol. 6:1273-1285.

Wynn, H. (1982) Optimum Submeasures with Applications to Finite Population Sampling, in "Statistical Decision Theory and Related Topics III", Vol.2, Academic Press, N.Y., pp. 485-495.