DESIGN OF EXPERIMENTS WITH SPATIALLY-AVERAGED OBSERVATIONS

Valeri Fedorov

May 1986
WP-86-24

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
2361 Laxenburg, Austria

ACKNOWLEDGEMENTS

I am very grateful to Prof. A.B. Kurzhanski for encouraging discussions and to A. Tedards for her patience in editing and proofing the paper.

PREFACE

This paper deals with experiments when only some average values (over time or space intervals) can be measured. This kind of experiment can be encountered in the areas of remote sounding of atmosphere, spectrometry, sample surveys, radioactivity analysis, etc. In these cases, the optimal experimental design means that the choice of intervals of observation (sometimes referred to as "windows") and corresponding shares of totally available time (or expenses) will maximize the final experimental information.

Formalization of the problem leads to a special class of optimization problems closely related to the classical "Markov's moments" problem. This paper contains new analytical and numerical results, together with a short but illuminative survey of previous researches.

Prof. M.A. Antonovsky
Environmental Program

CONTENTS

DESIGN OF EXPERIMENTS WITH SPATIALLY-AVERAGED OBSERVATIONS

Valeri Fedorov

1. INTRODUCTION

A number of publications concerning optimal design of experiments when controls belong to some functional space were published in the late 1970's. Now it is evident that the basic ideas behind these theoretical approaches are the same as in traditional experimental design theory (e.g., Fedorov, Uspensky 1977; Mehra 1974; Kozlov 1981; Pazman 1986). The differences become tangible in the application of general theoretical results to specific experimental problems.

In this paper these differences will be traced for experiments with spatially-averaged observations.

The simplest statistical model describing at least a part of the above mentioned experiments is the following one:

$$y_i = \vartheta^\tau x_i + \varepsilon_i \ , \quad i = \overline{1,N} \tag{1}$$

where $\vartheta \in R^m$ is a vector of unknown parameters; ε_i are independent random values with zero means and finite variances σ_i^2 (a more detailed assumption will be formulated later); τ stands for transposing. Variables x_i are defined by the integral

$$x_i = \int_V f(v) h_i(v) dv \ , \tag{2}$$

$f(v)$ is a vector of given basic functions; $h(v)$ are some functions which can be chosen (controlled) by an experimenter, $0 \le h \le 1$. In some cases integral (2) must be a Lebesque one. Function $h(v)$ describes the physical nature of an experiment and most typical examples will be given below in Section 2. If the least squares estimators

$$\hat{\vartheta} = Arg \min_{\vartheta} \sum_{i=1}^{N} \sigma_i^{-2} \{y_i - \vartheta^\tau x_i\}^2$$

are used to analyze experiments described by (1), then the quality of these estimators is defined by their dispersion (variance-covariance) matrices $D = E\{(\hat{\vartheta} - \vartheta_t)\ (\hat{theta} - \vartheta_t)^\tau\}$, where the subscript τ stands for true values. It is well known (see Fedorov 1972) that in regular cases

$$D^{-1} = \bar{M} = \sum_{i=1}^{N} \sigma_i^{-2} x(h_i) x^\tau(h_i) \tag{3}$$

where σ^2 could depend upon $h(t)$ also. Matrix M is usually called "information matrix."

The objective of optimal experimental design is the search for controls $h_i^*(v)$ providing better dispersion matrices or (more accurately) some functions of them (see Sections 3, 4).

2. EXPERIMENTAL PROBLEMS

Example 1. Spectrometric Experiments

In these experiments the measurement tools (e.g., a spectrometer, a radiometer, etc.) have a finite spectral resolution and can measure only a spectrum (of absorbtion or radiation) averaged over some frequency interval which is defined by the so-called "slit" function. Formally the model of a measurement process can be described in the simplest case (when Beer's law is still working and the spectrum is linearly dependent upon concentrations of components) by the following model:

- The total spectrum intensity linearly depends upon the spectrum intensities of any component

$$\eta(v,\vartheta) = \vartheta^{\mathsf{T}} f(v)$$

where v is frequency; $\vartheta^{\mathsf{T}} = (\vartheta_1, \ldots, \vartheta_m)$ is the vector of concentrations.

- The observed signal y_i is a linear functional of the spectrum under analysis [compare with (2)]

$$y_i = \int_V \eta(v,\vartheta) h_i(v)\,dv + \varepsilon_i \quad , \tag{4}$$

where $h_i(v)$ is a slit function (or resolution function) of the i-th observation; V is the frequency interval available for observations; ε_i is the error of observation.

Usually a slit function can be satisfactorily approximated by

$$h_i = \begin{cases} 1, & v_{ij} \le v \le v_{ij} + \Delta_j \\ 0, & v < v_{ij}\ , v \ge v_{ij} + \Delta_j \end{cases} \quad , \quad j = \overline{1,k} \quad . \tag{5}$$

- Errors ε_i are assumed to be random, independent for different i and their variance σ_i^2 can depend upon $h_i(v)$ and upon the time t_i spent on the i-th observation. In practice one can face several possibilities. For instance,

$$(a)\ \ \sigma_i^2 = \sigma^2 \qquad (b)\ \ \sigma_i^2 = \sigma^2 \left[\int_V h_i(v)\,dv\right]$$

$$(c)\ \ \sigma_i^2 = \sigma^2 t_i \qquad (d)\ \ \sigma_i^2 = \sigma^2 t_i \left[\int_V h_i(v)\,dv\right]. \tag{6}$$

The latter two cases correspond usually to the situation when an experimenter can choose the duration of an observation. If instead of y_i (these values can be called "total radiation") one considers average values ("average radiation"):

$$\bar{y}_i = y_i \left\{\int_V h_i(v)\,dv\right\}^{-1}$$

or

$$\bar{y}_i = y_i \left\{t_i \int_V h_i(v)\,dv\right\}^{-1} \tag{7}$$

then their variances will be correspondingly equal:

$$(a)\ \ \sigma_i^2 = \sigma^2 \qquad (b)\ \ \sigma_i^2 = \sigma^2 \left\{\int_V h_i(v)\,dv\right\}^{-1}$$

or

$$(c)\ \ \sigma_i^2 = \sigma^2 t_i^{-1} \qquad (d)\ \ \sigma_i^2 = \sigma^2 t_i^{-1} \left\{\int_V h_i(v)\,dv\right\}^{-1} \tag{8}$$

and variables x_i will be defined by the following formulae:

$$x_i = \int_V f(v)h_i(v)dv \{\int_V h_i(v)dv\}^{-1}$$

or

$$x_i = \int_V f(v)h_i(v)dv \{t_i \int_V h_i(v)dv\}^{-1}$$

In what follows below, the notation y_i will be used everywhere and the structure of variances σ_i^2 and values x_i will define the case.

Example 2. Remote Sounding of Atmosphere

The satellite measurements of the outgoing radiance in the infrared spectrum band become routine for distant retrieval of different physical parameters of the atmosphere (for instance, the vertical profile of temperature, humidity, ozone concentration, etc.; see Condratjev and Timofeev 1970, Strand and Westwater 1968, Twomey 1966). In a simplified form, the measured radiation $u(v)$ depends upon the vertical temperature profile $T(z)$, where z is an altitude, and the transmittance function of the atmosphere $\rho(v,z)$:

$$u(v) = B\{v,T(0)\}\rho(v,0) + \int_0^Z B\{v,T(z)\}\frac{\delta\rho(v,z)}{\delta z}dz$$

where Z corresponds to the altitude of a measurement tool; $B(v,T)$ is the Planck's function. Both $\rho(v,z)$ and $B(v,T)$ are assumed to be known.

The most crucial assumption is the possibility to approximate the function $T(z)$ by some parametric function $T(z,\vartheta)$ with subsequent linearization of the integral equation in the vicinity of an initial estimate of the temperature profile $\bar{T}(z)$ so that

$$u(v) \approx \vartheta^T f(v) = \vartheta^T\{\psi(v,0)\rho(v,0) + \int_0^Z \psi(v,z)\frac{\delta\rho(v,z)}{\delta z}\delta z\}$$

and integration with a weight function $h(v)$ over feasible frequency band V leads to the model described by (1), (2), (6) or (7) and (8).

3. OPTIMIZATION PROBLEMS IN EXPERIMENTAL DESIGN PROBLEM

Following Section 1, let us try to formalize a design problem for experiments from Section 2. As with traditional design theory (see above cited publications), the set of values

$$\xi_x = \{p_i,x_i\}_1^n \ , \quad \Sigma p_i = 1 \ , \tag{9}$$

where the weights p_i are the shares n_i/N of total number of measurements (or the shares t_i/T of total time available), which have to be done under the conditions x_i (or at the supporting points x_i), will be called a design (of an experiment). In the traditional case, the set $X \in R^m$ of feasible controls (operability region) is explicitly given. In the considered case, X is the mapping of a feasible set H (in some functional space) of controls $h_i(v)$, and usually the construction of X (say, its boundaries) is a problem of great difficulty. Therefore, it could be useful to consider designs in the original space also:

$$\xi_h = \{p_i,h_i(v)\} \tag{10}$$

From (3) it is evident that for model (1), (2) the information matrix (and subsequently dispersion matrix) depends upon the location of the supporting points x_i, and if at some points several measurements n_i are done, then on functions h_i also, but does not depend upon the results of measurements. In other words, the information matrix M depends upon a design ξ_x (or ξ_h).

Due to this fact, the design problem can be formulated as the following minimization problem

$$\xi^* = Arg \min_{\xi} \Phi\{\bar{M}(\xi)\} \tag{11}$$

where ξ can have both possible subscripts. The function Φ (optimality criteria) describes the objectives of an experimenter.

If the whole set of parameters present some interest, then it is reasonable to minimize the volume of the ellipsoid of concentration which is proportional to $|\bar{M}(\xi)|^{-1/2}$ and one can use $\Phi[M] = |\bar{M}(\xi)|^{-1/2}$ (everywhere $|A|$ means the determinant of matrix A).

If it is necessary to know a behaviour of some linear function of ϑ: $\psi^T(z)\vartheta$, then it is reasonable to minimize the average variance of this function over some region of interest. In this case

$$\Phi[M] = \int_Z \psi^T(z)\bar{M}^{-1}(\xi)\psi(z)dz$$

Rather detailed lists of the most popular optimality criteria can be found in Fedorov 1972; Silvey 1980.

In the future it is reasonable to distinguish between two types of designs: continuous and discrete ones.

In the first case, weights p_i can vary continuously between 0 and 1. This takes place when the weight is proportional to the time of measurement. We can go further and assume that any probabilistic measure $\xi_x = \xi(dx)$ or $\xi_H = \xi(dh)$ describes some design. In experimental practice, it could be impossible to realize, for instance, continuous measures. But fortunately for any design with continuous measure, it is possible to find the design with the same information matrix, but with measure concentrated in the final number of supporting points. It is easy to see that in the considered situation

$$\bar{M}(\xi_x) = N\int_X xx^T\xi(dx) = NM(\xi_x)$$

or

$$\bar{M}(\xi_H) = N\int_H x(h)x^T(h)\xi(dh) = NM(\xi_H)$$

In what follows below, the subscript x or h will be omitted without any comments if it will not lead to confusion.

Assuming that $\Phi(NM) = a(N)\Psi(M)$ (and it is true for the majority of optimality criteria used in practice) minimization problem (11) can be replaced by

$$\xi^* = Arg \min_{\xi} \Psi\{M(\xi)\} \tag{12}$$

where no values depend upon the total time or the total number of available measurements. This means that a continuous optimal design does not depend upon them also.

This useful property is not valid in the discrete case.

4. CONTINUOUS OPTIMAL DESIGNS

For the sake of simplicity in this section and all subsequent sections only the case when

$$\Psi\{M\} = |M^{-1}| \quad \text{and} \quad \xi^* = Arg \max_{\xi} |M|$$

will be considered. Other criteria can be handled in a similar way (see Fedorov 1980; Silvey 1980). It is most convenient to start with continuous version of designs ξ_x defined by (9). Then the celebrated Kiefer-Wolfovitz equivalence theorem (see Fedorov 1972) can be used and only one assumption is necessary for its fulfillment:

(a) Operability region X is compact.

Theorem 1

(1) There exists an optimal design ξ_x^* containing no more than $m(m+1)/2$ supporting points.

(2) The following problems are equivalent:

- maximization of $|M(\xi_x)|$,
- minimization of $\max_{x \in X} \lambda(x) d(x, \xi_x)$,
- $\max_{x \in X} \lambda(x) d(x, \xi_x) = m$,

  where $\lambda(x_i) = \sigma_i^{-2}$ and $d(x, \xi_x) = x^T M^{-1}(\xi_x) x$.

(3) At the supporting points of an optimal design ξ_x^* the function $d(x, \xi_x)$ approaches its maximum.

(4) The set of optimal designs is convex.

In a number of comparatively simple situations Theorem 1 gives a chance to construct optimal design analytically. For more compiicated models it helps to develop numerical procedures and to understand some general features of optimal designs.

For instance, if one manages to prove that operability region X is compact, then he can be sure that (see point (3) of the theorem) all supporting points of an optimal design ξ_x^* are some boundary points of X.

Example 1

Let

$$f^T(v) = (1, v), \quad |v| \le 1, \quad \Psi(M) = |M|^{-1/2},$$

$$\gamma(x) = const, \quad x_1 = \int_{-1}^{1} h(v) dv, \quad x_2 = \int_{-1}^{1} v h(v) dv$$

The set X can be easily constructed because of the simplicity of integral (2) and its boundary is described by the following curves:

$$x_2 = \pm\{\frac{1}{2} - (x_1 - 1)^2/2\}, \quad 0 \le x_1 \le 2$$

From point 2 of Theorem 1 it follows that the supporting points of an optimal design must coincide with the points where the ellipse $x^T M^{-1}(\xi_x^*)x$ is tangent to X. This ellipse must have the least area (or $|M|$) between all ellipses containing X. The simple calculations show that the points with coordinates (2;0), $(\sqrt{5}-1,\ 2\sqrt{5}-4)$, $(\sqrt{5}-1,\ 4-2\sqrt{5})$ could be supporting ones.

From symmetry of X and point 4 of Theorem 1 it follows that the weights of the two last points must be equal. Straightforward maximization of the determinant $|M|$ gives

$$p_2 = p_3 = 0.5(\sqrt{5}-1)^{-1},\ p_1 = 1-2p_2$$

Finally, from the simple integral equations defined at the beginning of the example it is easy to find that

$$\xi_h^* = \begin{Bmatrix} h_1^*(v)\equiv 1 & h_2^*(v)=\begin{cases}1, & v\le\sqrt{5}-2\\ 0, & v>\sqrt{5}-2\end{cases} & h_3^*(v)=h_2(-v) \\ 0.19 & 0.405 & 0.405 \end{Bmatrix}, \quad |v|\le 1$$

with the information matrix

$$M = \begin{Bmatrix} 2 & 0 \\ 0 & 0.18 \end{Bmatrix}$$

For comparison of the traditional design with two Δ-windows at the points $v=\pm 1$, the same matrix equals

$$M = \begin{Bmatrix} \Delta & 0 \\ 0 & \Delta^2 \end{Bmatrix}.$$

This means the ratio of standard errors will be 0.5Δ (for $\hat{\vartheta}_1$) and $5.5\Delta^2$ (for $\hat{\vartheta}_2$). So the optimal design is essentially more effective than the traditional approach, especially for small Δ.

Example 2.

The characteristics of optimal designs (for instance, the location of supporting points) essentially depend upon the chosen basic function. To illuminate this, let us consider the regression problem formulated in the previous example with a new basic function $f^T(v) = (\sin\pi v,\ \cos\pi v)$. It can be proved that for any slit function $h(v)$ the vector (Krein, Nudelman 1973, VII:3)

$$x^T = \{\int_{-1}^{1} h(v)\sin\pi v dv\ ,\ \int_{-1}^{1} h(v)\cos\pi v dv\}$$

must belong to the circle $\{X: x_1^2 + x_2^2 \le 2\}$.

The optimal designs for this operability region and response function $v^T x = \vartheta_1 x_1 + v_2 x_2$ can be easily constructed. For instance, optimal design can consist of the supporting points coinciding with all vertexes of any regular polygon refined to the circle X and their weights must be equal. One of the simplest optimal designs is

$$\xi_x^* = \begin{Bmatrix} (0;2) & (2;0) \\ 0.5 & 0.5 \end{Bmatrix}$$

To find out the corresponding optimal design in slit function space the integral equations

$$x_{1l}^{*} = \int_{-1}^{1} h_l(v)\sin\pi v dv \quad \text{and} \quad x_{2l}^{*} = \int_{-1}^{1} h_l(v)\cos\pi v dv$$

must be solved. One of the solutions is

$$h_1^{*}(v) = \begin{cases} 1, & -0.5 \le v \le 0.5, \\ 0, & v < -0.5, \ v > 0.5, \end{cases}$$

and

$$h_2^{*} = \begin{cases} 0, & -0.5 \le v \le 0.5, \\ 1, & v < -0.5, \ v > 0.5. \end{cases}$$

It is worthwhile to note that the widths of slit function "windows" have the same order as intervals of typical variations of basic functions.

5. NUMERICAL METHODS

If assumption (a) holds and there is a way to find X then the following numerical can be used for optimal design construction (Fedorov 1972):

$$\xi_{s+1} = (1-\alpha_s)\xi_s + \alpha_s \xi(x_s) \tag{13}$$

where $\xi(x_s)$ is the design concentrated at the single point

$$x_s = Arg \max_{x \in X} \lambda(x) d(x, \xi_s) \tag{14}$$

where

$$d(x, \xi_s) = x^{\tau} M^{-1}(\xi_s) x$$

$$M(\xi_{s+1}) = (1-\alpha_s)^{-1} \left[1 - \frac{\lambda(x_s)\alpha_s M^{-1}(\xi_s) x_s x_s^{\tau} M^{-1}(\xi_s)}{1-\alpha_s + \lambda(x_s)\alpha_s d(x_s, \xi_s)} \right] M^{-1}(\xi_s)$$

The sequence $\{\alpha_s\}$ can be, for instance:

$$(a) \quad \sum_{s=0}^{s=\infty} \alpha_s = \infty \ , \quad \lim_{s \to \infty} \alpha_s = 0 \qquad (b) \quad \alpha_s = \frac{\lambda(x_s) d(x_s, \xi_s) - m}{\{\lambda(x_s) d(x_s, \xi_s) - 1\} m}$$

Both of them guarantee that

$$\lim_{s \to \infty} |M(\xi)| = \min_{\xi} |M(\xi)| \tag{15}$$

More sophisticated versions of this method are discussed in Ermakov 1983; Fedorov and Uspensky 1975.

In spite of the formal simplicity of iterative procedure (13), (14), its practical usefullness is rather restricted: one must find out the way to construct X before using this procedure.

Procedure (13), (14) can be replaced by equivalent procedures in the space of functions $h(v)$. For that one has to replace the vector $\lambda^{1/2}(x)x$ with the vector $\int_V h(v) f(v) dv$ (for cases (a),(c) from (6) and (8)), and with the vector $\int_V h(v) f(v) dv \Big/ \sqrt{\int_V h(v) dv}$ (for cases (b),(d) from (6) and (8)).

For the sake of simplicity, the iterative procedure will be considered for the first cases when it can be presented in the following form:

$$\xi_{s+1} = (1-\alpha_s)\xi_s + \alpha_s \xi(h_s), \tag{16}$$

$$h_s = Arg \max_{h \in H} d(h, \xi_s), \tag{17}$$

where

$$d(h, \xi_s) = \int_V \int_V h(v)h(v')f^T(v)M^{-1}(\xi_s)f(v')dvdv'.$$

Unfortunately, maximization problem (17) is more complicated than (14). One of the simplest could be the following one:

- discretisize the set V (and therefore $V \times V'$ also), say with interval Δ;
- collect all points v_j on the corresponding grid which positively contribute to the sum

$$d(\xi_s) = \sum_{j,j'} f(v_j)M^{-1}(\xi_s)f(v_{j'}) \tag{18}$$

- put $h_s(v_j) = 1$ if v_j was chosen on the previous stage; otherwise $h_s(v_j) = 0$
- the fulfillment of the inequality (which is corollary to Theorem 1) $d(\xi_s h_s) \geq m$ tests that h_s can be used in (16).

This procedure is admissible for applications, for instance, when V is one-dimensional and that is the case for a number of applied problems (see Section 2).

Note 1.

Iterative procedure (16) will converge in sense (15) (for both original and discrete versions) if instead of (16) only the inequality

$$d(h_s, \xi_s) > m$$

will take place on every step.

Note 2.

The sequences $\{j\}$ and $\{j'\}$ must be identical. Therefore if $j = j_0$ is included in sum (18), then $j' = j_0$ must be included also.

Note 3.

When the slit function $h(v)$ can change its value at the modes of Δ-grid, then one can tell about Δ-optimal designs (ξ_Δ^*) which can be considered as some approximation of optimal designs defined by (12). The iterative procedure (16), (17) and (18) guarantees that

$$\lim_{s \to \infty} |M(\xi_s)| = M(\xi_\Delta^*).$$

The idea of Δ-optimal designs can be advanced further. As it was observed in Example 2, the widths of slit function windows were related to the intervals of variation of basic functions. Therefore, it is reasonable to decompose the set V into a comparatively moderate number of subsets Δ_j $(j = \overline{1,k})$, for instance, coinciding with the most typical fluctuations of basic functions. Assume that integrals

$$F_j = \int_{\Delta_j} f(v)dv \tag{19}$$

can be calculated (numerically or analytically). Then the operability region X can be approximated by X_Δ with elements

$$x = Fu ,$$

where $F=(F_1,\ldots,F_k)$, $u^\tau = (u_1,\ldots,u_k)\in U_\Delta$ and $u_j = 1$, if $h(v)=1$, $v\in\Delta_j$, and $u_j=0$; otherwise, $j=\overline{1,k}$. Observing that $\vartheta^\tau x = \vartheta^\tau Fu = \omega_\tau u$ and the information matrix equals

$$M = F\int_{U_\Delta} uu^\tau \xi(du)F^\tau = FM_\omega(\xi_u)F^\tau$$

where $\xi(du)$ describes a design ξ_u with supporting points in U_Δ, one can conclude that rank F must be equal to m (number of estimated parameters). Therefore, the decomposition of V should contain at least m subsets Δ_j ($k\geq m$).

When $k=m$ and of course $|F|\neq 0$, then $|M|=|F|^2|M_\omega|$ and the design problem is reduced to the maximization of $|M_\omega|$. The latter problem coincides with the routine problem of "optimal weighting." (See Ermakov 1983)

If $k>m$ then iterative procedure (13),(14) can be used with the replacement of the vector $\lambda^{1/2}(x)x$ with the vector Fu [or $Fu/\sum_{j=1}^{k} u_j$; compare with comments to (16),(17)]:

$$u_s = Arg\max_u uF^\tau M^{-1}(\xi_s)Fu. \qquad (20)$$

The maximization problem (20) is a discrete one and at every s-th stage it demands no more than 2^k calculations of $uF^\tau M^{-1}(\xi_s)Fu$.

6. STRUCTURE OF SLIT FUNCTION

In the previous sections, it was assumed that the slit function can equal 1 or 0. Some "physical" arguments were behind this assumption. The compactness of operability region X (see Section 4) was also an essential assumption which was done to slimplify all final results. If one refuses this assumption, then instead of optimal designs, so-called optimal sequences (see, for example, Ermakov 1983) must be considered and that leads to some technical difficulties. The following results (which are straightforward corollaries of well-known results from classical approximation theory; see, for example, Karlin and Studden 1966, Chapter VIII) illuminate that both above mentioned assumptions are not very restrictive. For the sake of simplicity, we consider a one-dimensional case ($V\in R^1$):

Assume now that:

(a) $0\leq h(v)\leq 1$, for any $v\in(a,b)$

(b) Functions $f(v)$ constitute a Tchebysheff system on the open interval (a,b), where a and b are possibly infinite. This assumption requires that the functions $f(v)$ be continuous on (a,b) and the determinants

$$\begin{vmatrix} f_1(t_1) & f_1(t_2) & \ldots & f_1(t_m) \\ f_2(t_1) & f_2(t_2) & \ldots & f_2(t_m) \\ \cdot & \cdot & \ldots & \cdot \\ \cdot & \cdot & \ldots & \cdot \\ \cdot & \cdot & \ldots & \cdot \\ f_m(t_1) & f_m(t_2) & \ldots & f_m(t_m) \end{vmatrix}$$

$$t_i\in(a,b), \quad i=\overline{1,m}$$

are positive.

Theorem 2.

The operability region

$$X = \{x = \int_a^b f(v)h(v)dv:\ 0 \le h(v) \le 1\}$$

is a compact convex set in R^m.

From Theorem 1 it is clear that all supporting points of any optimal design must be boundary points of X. Therefore, only these points had to be considered in the previous sections, and for them the following result takes place:

Theorem 3.

The necessary and sufficient condition for x to be a boundary point of X is the fulfillment of the condition

$$h(v)\{1-h(v)\} = 0 \tag{21}$$

almost everywhere in (a,b).

Let $h(v)$ be a function satisfying (21) and let $I(x)$ be the number of separate nondegenerate intervals (windows of a slit function) 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 $x \in X$, $I^*(x)$ stands for the least possible $I(x)$ (it could be several different functions $h(v)$ giving the same x).

Theorem 4.

A necessary and sufficient condition that x belongs to the boundary of X is that $I^*(x) \le (m-1)/2$. Moreover, every boundary point corresponds to a unique $h(v)$ with $I(x)=I^*(x)$.

Theorems 3 and 4 allow for the development of a comparativeiy simple algorithm of optimal design construction.

Let $\bar{v}=(a,v_1,\ldots,v_{m-1},b)$, where $a \le v_1 \le \cdots \le v_{m-1} \le b$. According to Theorem 4, 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 allows for the modification of the iterative procedure (16),(17) to the procedure with maximization in space which dimension is less or equal $(m-1)$, where m is a number of basic functions:

$$\xi_{s+1} = (1-\alpha_s)\xi_s + \alpha_s \xi(h_s), \tag{22}$$

$$h_s = Arg \max_{\gamma,x} d(x_\gamma, \xi_s), \quad \gamma=1,2, \tag{23}$$

where $a \le v_2 \le \cdots \le v_{m-1} \le b$,

$$x_1 = \int_a^b f(v)\bar{h}(v)dv$$

and

$$x_2 = \int_a^b f(v)\underline{h}(v)dv.$$

Procedure (22),(23) in a computational sense coincides with iterative procedures used for traditional design problems and can be handled with software developed for the latter one (see Fedorov, Uspensky 1975).

REFERENCES

Condratjev, K.J. and Timofeev (1970) The Thermal Remote Sounding of Atmosphere from Satellites. Leningrad, 410.

Ermakov, S.M., editor (1983) Mathematical Theory of Experimental Design. Moscow: Nauka, 391 (In Russian).

Fedorov, V.V. (1972) Optimal Design of Experiments. New York: Academic Press, 292.

Fedorov, V.V. and A.B. Uspensky (1975) Numerical Aspects of the Least Squares Methods. Moscow: Moscow State University, 168.

Fedorov, V.V. and A.B. Uspensky (1977) On the Optimal Condition Choice of Spectroscopic Measurements. Moscow: Proc. of State Scient. and Research Center for the Study of Earth's Environment and Natural Resources, 4:42-53 (In Russian).

Fedorov, V.V. (1980) Convex Design Theory. Math. Oper. Stat., 11:403-413.

Karlin, S. and W.J. Studden (1966) Tchebycheff Systems: With Applications in Analysis and Statistics. New York: Wiley & Sons, 586.

Kozlov, V.P. (1981) Design of Regression Experiments in Functional Spaces, in "Mathematical Methods in the Design of Experiments", ed. Penenko, V.V. Moscow: Nauka, Novosibirsk, 74-101 (In Russian).

Krein, M.G. and A.A. Nudelman (1973) Markov Moment Problem and Extremal Problems. Moscow: Nauka, 552.

Mehra, R.K. (1976) Optimal Input Signals for Parameter Estimation in Dynamic System. IEEE Trans. on Autom. Control AC-19:753-768.

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

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

Strand, O.N. and E.R. Westwater (1968) Minimum RMS-Estimation of the Numerical Solution of a Fredholm Integral Equation of the First Kind. SIAM J. Numer. Anal. 5:287-295.

Twomey, S. (1966) Indirect Measurements of Atmospheric Temperature Profiles from Satellites: Mathematical Aspects of the Inversion Problem. Monthly Weather Review, 99:363-366.