Kybernetika

Jaroslava Mikulecká

On a hybrid experimental design

Kybernetika, Vol. 19 (1983), No. 1, 1--14

Persistent URL: http://dml.cz/dmlcz/124466

KYBERNETIKA — VOLUME 19 (1983), NUMBER 1

ON A HYBRID EXPERIMENTAL DESIGN

JAROSLAVA MIKULECKÁ

A hybrid design simultaneously provides information about both the supposed regression model and the hypothesis that the model is adequate. The hybrid design criterion is defined and an algorithm for the construction of hybrid optimal design is given.

1. INTRODUCTION

In statistical data processing, the basic problem is the correct choice of the mathematical model of the experiment and the problem of optimal parameter estimation in this model. Designing experiments, optimal (in some sense) for simultaneous model discrimination and parameter estimation, present thus an actual problem. This joint interest in discrimination and estimation was treated in several papers ([1], [2], [3], [4], [5], [7], [8], [17]) where the authors suggested various hybrid optimality criteria. In our paper we give an extension of the criterion defined by Stigler [17] and we give an algorithm for the construction of hybrid optimal designs.

We shall consider the following structure of regression experiment. On a compact metric space X, $m + 1$ linearly independent continuous functions $f_1, f_2, \ldots, f_{m+1}$ are given. For each $x \in X$ can be performed an elementary experiment whose outcome is a random variable $y(x)$ with the mean

$$\mathrm{E}(y(x)) = \sum_{i=1}^{m+1} \alpha_i f_i(x) \equiv \boldsymbol{\alpha}' \boldsymbol{f}(x) \tag{1.1}$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_{m+1})'$, $\boldsymbol{f}(x) = (f_1(x), f_2(x), \ldots, f_{m+1}(x))'$ and the variance $\sigma^2(y(x)) \equiv 1$.

The parameters $\alpha_1, \alpha_2, \ldots, \alpha_{m+1}$ are unknown and they have to be estimated from uncorrelated measurements performed in different points of X.

A design is a probability measure ξ on X supported by a finite set. ($\xi(x)$ is proportional to the number of repeated uncorrelated observations performed in x.)

An information matrix of the design ξ is

$$M(\xi) = \sum_{x \in X} f(x) f'(x)\, \xi(x) \tag{1.2}$$

The set of all information matrices we call $\mathfrak{M}$, the set of all designs we call Ξ.

Let N independent measurements of $y(x)$ be made at the points $x_1, x_2, \ldots, x_N$ in X. Assuming that M is nonsingular, it is well known that the covariance matrix of the best linear unbiased estimator of α is

$$\operatorname{cov}(\hat{\alpha}) = \frac{\sigma^2}{N} M^{-1}(\xi) \tag{1.3}$$

For a fixed total number of measurements, N, it would be desirable to allocate the measurements so that the covariance matrix $\operatorname{cov}(\hat{\alpha})$ is small in terms of some preference function Φ. The choice of the function of M^{-1} to minimize has important practical implications and depends on the aim of the experiment.

Throughout the paper we assume that Φ is convex and bounded below on $\mathfrak{M}$ and that it is differentiable in the space of all nonnegative definite $(m + 1) \times (m + 1)$ matrices. The $(m + 1) \times (m + 1)$ matrix $\nabla\Phi$ is defined as

$$(\nabla\Phi)_{ij} = \frac{\partial\Phi(M)}{\partial(M)_{ij}} \tag{1.4}$$

If Φ satisfies the condition

$$\Phi(M) < \infty \Leftrightarrow M \text{ is nonsingular} \tag{1.5}$$

then it is said to be a *global* criterion function.

2. A HYBRID DESIGN

Let the model for the expected response be

$$\mathsf{E}(y(x)) = \sum_{i=1}^{m} \alpha_i f_i(x) \equiv \alpha^{(1)\prime} f^{(1)}(x) \tag{2.1}$$

where

$$\alpha^{(1)} = (\alpha_1, \alpha_2, \ldots, \alpha_m)'$$
$$f^{(1)} = (f_1, f_2, \ldots, f_m)'$$

The model (2.1) adequacy can be tested by embedding (2.1) in the more general model

$$\mathsf{E}(y(x)) = \sum_{i=1}^{m+1} \alpha_i f_i(x) \equiv \alpha' f(x) \tag{2.2}$$

where

$$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_{m+1})'$$

$$f = (f_1, f_2, \ldots, f_{m+1})'$$

and testing the hypothesis

$$H_0 : \alpha_{m+1} = 0 \quad \text{against the alternative hypothesis}$$

$$H_1 : \alpha_{m+1} \neq 0$$

The power of the F-test of the hypothesis H_0 depends on the value $[\text{var}(\hat{\alpha}_{m+1})]^{-1}$ (see [2]). In order to detect departures from the model (2.1), experiments should be planned to give small values of the quantity var $(\hat{\alpha}_{m+1})$. These designs, providing a good check on the adequacy of the model, may give poor estimates of the oiiginal parameters $\alpha^{(1)}$ if the model (2.1) is actually adequate. Therefore it is appropriate to ask the hybrid design to optimize the parameter's estimate in the original model (2.1) subject to an upper bound on the value of var $(\hat{\alpha}_{m+1})$, e.g. var $(\hat{\alpha}_{m+1}) \leqq c$. Let c_0 be the minimal possible value of the var $(\hat{\alpha}_{m+1})$. Then the efficiency of some design ξ for estimating α_{m+1} is (see [1], [2])

$$\text{(2.3)} \qquad e(\hat{\alpha}_{m+1} \mid \xi) = \frac{c_0}{\text{var}(\hat{\alpha}_{m+1} \mid \xi)}$$

The hybrid optimal design then minimizes some optimality criterion in the model (2.1) subject to lower bound on the value of the efficiency for estimating α_{m+1}.

The matrix $M \in \mathfrak{M}$ can be partioned into submatrices (as in [11])

$$M = \begin{pmatrix} M_{\text{I}} & M_{\text{II}} \\ M'_{\text{II}} & M_{\text{III}} \end{pmatrix}$$

M_{I} being a $m \times m$ matrix belonging to the model (2.1).

Let $\tilde{\Phi}(M_{\text{I}})$ be some convex global optimality criterion in the model (2.1) (see [12]). Let us define the functions

$$\Phi(M) = \tilde{\Phi}(M_{\text{I}})\,; \quad M \in \mathfrak{M}$$

and

$$\Psi(M) = p'M^{-1}p - c\,; \quad p' = (0, 0, \ldots, 1) \in \mathbb{R}^{m+1}$$

Then we have the following nonlinear programming problem:

(P) minimize $\Phi(M)$ over $M \in \mathfrak{M}_+ = \{M \in \mathfrak{M} : \det M > 0\}$
subject to $\Psi(M) \leqq 0$

Definition 2.1. If M^* is a solution of the problem (P), then the design ξ^*, corresponding to M^*, is called a *hybrid optimal design.*

Theorem 2.1. The problem (P) has the following properties:

a) (P) is a convex programming problem,

b) the solution of (P) exists,
c) a Kuhn - Tucker vector exists for (P).

Proof. From the definition of (P) it is evident that all properties of a convex programming problem are satisfied (a detailed proof is given in [9], Theorem 3.6.2).
Let us define the following problem:

(A) minimize $\Phi(M)$ over $M \in \mathfrak{M}$
subject to $\det M_I \leqq c \det M$

which can be reformulated into the problem

$$\text{minimize } \Phi(M) \text{ over } M \in \mathfrak{M}^c \equiv \mathfrak{M} \cap \{M : \det M_I \leqq c \det M\}$$

We know that

1) $\mathfrak{M}$ is compact (see [11], proposition III. 9),
2) $\det M$ is a continuous function of matrix elements, therefore the set $\mathfrak{M}^c$ is closed;
3) Φ is a continuous function on $\mathfrak{M}$ (see [11]);

and we can conclude that a solution of (A) exists. Let M^A be a solution of (A), then $\det M_I^A > 0$, because $\tilde{\Phi}$ is a global optimality criterion in the model (2.1). On the other hand, $M^A \in \mathfrak{M}^c$, therefore the solution of (A) lies in $\mathfrak{M}_+$, hence the solution of (P) exists.

Assuming that the optimal value of (P) is finite, it is well known (see [6]) that a Kuhn - Tucker vector exists if the Slater condition is satisfied: there is a point $M \in \mathfrak{M}_+$ where $\Psi(M) < 0$. The formulation of (P) implies that $c > c_0$ where

$$c_0 = \min_{\xi \in \Xi} \{p' M^{-1}(\xi)\, p\} \; ; \quad p' = (0, 0, \ldots, 1) \in \mathbb{R}^{m+1}$$

The design $\xi_T^* : p' M^{-1}(\xi_T^*)\, p = \min_{\xi \in \Xi} p' M^{-1}(\xi)\, p$ can be found according to the algorithm, given in [11], chapter V. 2. This algorithm formulates such a sequence of information matrices $M_i \in \mathfrak{M}_+$, where the inequality $\Psi(M) < 0$ is fulfilled for all $i < i_0$. □

Theorem 2.2. Let λ be a Kuhn-Tucker vector for (P). The criterion $\Phi_\lambda : M \in \mathfrak{M} \to \Phi(M) + \lambda\, \Psi(M)$ is a global optimality criterion for the model (2.1).

Proof. According to (1.5) we shall prove that

$$\{M \in \mathfrak{M} : \Phi(M) + \lambda\, \Psi(M) < \infty\} \equiv \mathfrak{M}_+$$

But

$$\Phi(M) = \tilde{\Phi}(M_I) < \infty \Leftrightarrow \det M_I > 0 \Leftrightarrow \text{the parameters } \alpha_1, \alpha_2, \ldots, \alpha_m \text{ are estimable},$$

and

$$\Psi(M) = \operatorname{var}(\hat{\alpha}_{m+1} \mid M) - c < \infty \Leftrightarrow \text{the parameter } \alpha_{m+1} \text{ is estimable.}$$

Because $\lambda \geqq 0$, $\Psi(M) \geqq 0$ and the function $\Phi(M)$ is bounded below for all $M \in \mathfrak{M}$, we see that

$$\Phi(M) + \lambda\, \Psi(M) < \infty \Leftrightarrow \Phi(M) < \infty$$

and

$$\Psi(M) < \infty \Leftrightarrow (\alpha_1, \alpha_2, \ldots, \alpha_{m+1}) \quad \text{are estimable} \Leftrightarrow M \in \mathfrak{M}_+ . \qquad \square$$

Theorem 2.3. Assuming that ξ_c^* is a hybrid optimal design, then either ξ_c^* is Φ-optimal for the model (2.1), or var $(\hat{\alpha}_{m+1} \mid \xi_c^*) = c$.

Proof. The proof is obvious, if we take into account [14], Theorem 28.3. □

3. THE ALGORITHM

The hybrid optimal design was stated as a convex programming problem (P), so it is possible to use the Rockafellar's algorithm (see [15]). This algorithm converts the constrained problem to a sequence of unconstrained optimization problems, having the property that the successive solutions of the unconstrained problems converge to the solution of (P). Then some of the known optimal design algorithms ([11]) can be used in every unconstrained optimization. This is a very useful property, since we get directly the optimal design. Usual convex-programming algorithms would determine only the optimal matrix M^* corresponding to the optimal design ξ^*. And having the matrix $M^* = M(\xi^*)$ we would have to find a corresponding design by solving the following problem.

Seek a function $\xi^*(x)$, $x \in X$ subject to

$$\sum_{x \in X} f(x) f'(x)\, \xi^*(x) = M^*$$

$$\sum_{x \in X} \xi^*(x) = 1$$

$$\xi^*(x) \geqq 0 \quad \text{for all} \quad x \in X .$$

This problem can be solved, but evidently it complicates the determination of the optimal design.

Let us define the penalty function for (P) as

$$\text{(3.1)} \qquad L_r(M, \lambda) = \Phi(M) + \frac{1}{4r}\left[\theta^2(\lambda + 2r\, \Psi(M)) - \lambda^2\right]$$

where

$$\text{(3.2)} \qquad \theta(t) = \max\{t, 0\}$$

The basic procedure is that

1) In the internal iterations, given $\lambda^{(k)}$ and r ($r > 0$), we determine $M^{(k)}$ minimizing

$L_r(M, \lambda^{(k)})$ on $\mathfrak{M}_+$. The function $L_r(M, \lambda^{(k)})$ is convex in $M \in \mathfrak{M}_+$ (see [16]) and it can be considered as a global optimality criterion (according to Theorem 2.2). If some other requirements on $L_r(M, \lambda^{(k)})$ are fulfilled (we shall prove them later), minimum of $L_r(M, \lambda^{(k)})$ can be found according to the general optimal design algorithm (see [11], chapter V. 5).

2) In the *external iterations* we set (see [15])

$$\lambda^{(k+1)} = \lambda^{(k)} + 2r\, \Psi(M^{(k)}) \tag{3.3}$$

The transformation of this procedure into a locally convergent algorithm is patterned after Powell [13] and it is given in Fig. 1. A few comments on the flow diagram

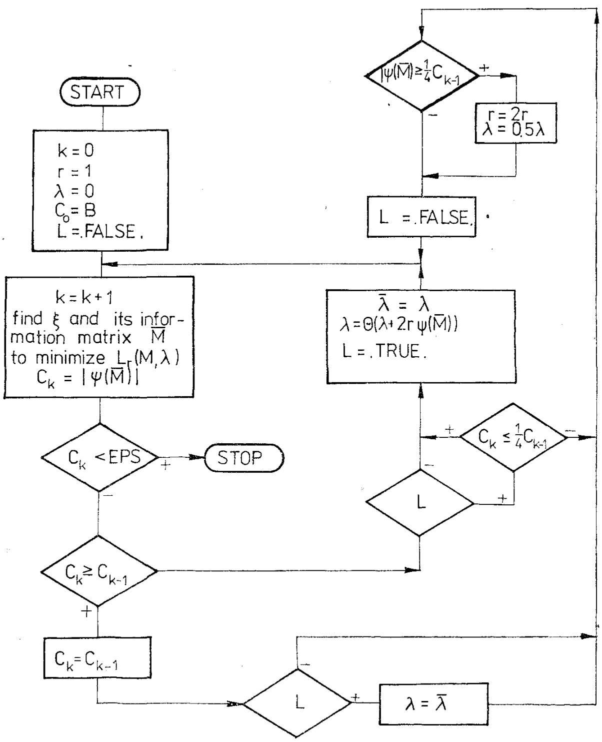

Fig. 1. Flow chart of a hybrid optimal design algorithm.

(Fig. 1) are needed, k is the number of an external iteration, and C_k is usually set to the last value of $|\Psi(M)|$, which has been calculated. To start the iterative process, we set $C_0 = B$, where B is some large positive number exceeding the magnitude of $|\Psi(M)|$. L is a logical variable used as a switch. If L is. FALSE., it indicates that we have just chosen a new value of r. If L is .TRUE., we have applied the correction of λ in the previous iteration. We continue to apply this correction on λ provided that it gives the required convergence, namely $C_k \leqq \frac{1}{4}C_{k-1}$.

The most difficult operation in the previous algorithm is the minimization of $L_r(M, \lambda)$. Under the following properties (A), (B), (C) of $L_r(M, \lambda)$, $\min L_r(M, \lambda)$ (with min considered over $M \in \mathfrak{M}_+$) can be found according to the general optimal design algorithm ([11], chapter V.5).

Let $\mathfrak{S}^{(m+1)\times(m+1)}$ be the set of all symmetric $(m + 1) \times (m + 1)$ matrices, and let $\mathscr{L}(\mathfrak{M})$ be the linear space spanned by $\mathfrak{M}$. The conditions which are to be checked are:

(A) There exists such a set $\mathfrak{U}_L \subset \mathfrak{S}^{(m+1)\times(m+1)}$ which
  a) $\mathfrak{M}_+ \subset \mathfrak{U}_L \subset \mathscr{L}(\mathfrak{M})$;
  b) $\mathfrak{U}_L$ is open in $\mathscr{L}(\mathfrak{M})$;
  c) $L_r(M, \lambda)$ is defined and finite on $\mathfrak{U}_L$ for every admissible r, λ;
  d) $L_r(M, \lambda)$ is convex in M on $\mathfrak{U}_L$.

(B) If $M_n \in \mathfrak{M}_+$ $(n = 1, 2, \ldots)$ and $\lim_{n\to\infty} M_n = M \in \mathfrak{M} - \mathfrak{M}_+$ then $\lim_{n\to\infty} L_r(M_n, \lambda) = \infty$;

(C) For every $M \in \mathfrak{M}_+$ there exists the gradient $\nabla_M L_r(M, \lambda)$.

To prove this properties, let $\mathfrak{U}_L = \{A : A \in \mathscr{L}(\mathfrak{M}), \det(A) > 0\}$. Then obviously $\mathfrak{M}_+ \subset \mathfrak{U}_L \subset \mathscr{L}(\mathfrak{M})$. Because the determinant is a continuous function of the matrix elements, $\mathfrak{U}_L$ is open in $\mathscr{L}(\mathfrak{M})$.

The property c) is obvious from the definition of $L_r(M, \lambda)$ and d) follows from the convexity of $\Phi(M)$ and $\Psi(M)$. Let $M_n \in \mathfrak{M}_+$ and let

$$\lim_{n\to\infty} M_n = M \in \mathfrak{M} - \mathfrak{M}_+$$

$$\liminf_{n\to\infty} L_r(M_n, \lambda) = \liminf_{n\to\infty} \left\{\Phi(M_n) + \frac{1}{4r}[\theta^2(\lambda + 2r\,\Psi(M_n)) - \lambda^2]\right\}$$

Because $\Phi(M)$ is a global optimality criterion, $\lim_{n\to\infty} \Phi(M_n) = \infty$ and

$$\liminf_{n\to\infty} \theta^2(\lambda + 2r\,\Psi(M_n)) - \lambda^2 \geqq -\lambda^2 \geqq -\infty$$

Hence $\lim_{n\to\infty} L_r(M_n, \lambda) = \infty$.

For every $M \in \mathfrak{M}_+$ there exists $\nabla_M L_r(M, \lambda)$ and

(3.4) $$\nabla_M L_r(M, \lambda) = \nabla\Phi(M) + \theta(\lambda + 2r\,\Psi(M))\,\nabla\Psi(M)$$ □

To guarantee the convergence of the general optimal design algorithm we have to prove the following condition (see [11], chapter V.5):

(D) For every $d > 0$ there exists such a $K_d > 0$ that

$$\frac{\|\nabla_M L_r(M, \lambda) - \nabla_M L_r(\bar{M}, \lambda)\|}{\|M - \bar{M}\|} \leqq K_d \quad \text{where} \quad \|M\| = \operatorname{Tr} M'M \tag{3.5}$$

$$(M, \bar{M} \in \mathfrak{M}_d \equiv \{M : M \in \mathfrak{M}, L_r(M, \lambda) \leq d\})$$

Lemma 3.1. If

$$-\ln \det M_I + \frac{1}{4r}\{\theta^2(\lambda + 2r\,\Psi(M)) - \lambda^2\} \leqq d \tag{3.6}$$

then

$$\|M_I^{-1}\|^2 \leqq m\left[A^{m-1} \cdot \exp\left\{d + \frac{\lambda^2}{4r}\right\}\right]^2 \tag{3.7}$$

where

$$A = \max_{x \in X} \|f^{(1)\prime}(x)\| < \infty \tag{3.8}$$

Proof. For every matrix M_I is

$$\|M_I^{-1}\|^2 = \operatorname{Tr} M_I^{-2} = \sum_{i=1}^{m} \varrho_i^{-2} \tag{3.9}$$

$$\det M_I^{-1} = \prod_{i=1}^{m} \varrho_i^{-1} \tag{3.10}$$

where ϱ_i are eigenvalues of M_I.

Because M_I is an information matrix $\varrho_i \geqq 0$ for all i. To prove the lemma, it is in view of (3.9) enough to prove that eigenvalues of M_I^{-1} are bounded. According to (3.6) is

$$\det M_I^{-1} \cdot \exp\left\{\frac{1}{4r}\left[\theta^2(\lambda + 2r\,\Psi(M)) - \lambda^2\right]\right\} \leqq \exp(d) \tag{3.11}$$

Because $\theta^2(\lambda + 2r\,\Psi(M)) - \lambda^2 \geqq -\lambda^2$, we have

$$\det M_I^{-1} \exp\left\{-\frac{\lambda^2}{4r}\right\} \leqq \det M_I^{-1} \exp\left\{\frac{1}{4r}\left[\theta^2(\lambda + 2r\,\Psi(M)) - \lambda^2\right]\right\} \leqq \exp(d)$$

i.e.

$$\det M_I^{-1} \leqq \exp\left\{d + \frac{\lambda^2}{4r}\right\} \tag{3.12}$$

Because M_I is an information matrix, there exists a number $D > 0$ such that

$$D = \min_{M \in \mathfrak{M}} \det M_I^{-1} \leqq \det M_I^{-1} \tag{3.13}$$

From (3.12) and (3.13) follows the inequality

$$0 < \mathrm{D} \leqq \det M_{\mathrm{I}}^{-1} \leqq \exp\left\{d + \frac{\lambda^2}{4r}\right\} \tag{3.14}$$

that is

$$\exp\left\{-d - \frac{\lambda^2}{4r}\right\} \leqq \det M_{\mathrm{I}} \leqq \frac{1}{\mathrm{D}}$$

respectively. M_{I} is an information matrix, therefore

$$M_{\mathrm{I}} = \sum_{x \in X} f^{(1)}(x) f^{(1)\prime}(x)\, \xi(x)$$

$$\max_i \varrho_i = \max_{\|u\|=1} u' M_{\mathrm{I}} u = \max_{\|u\|=1} \sum_{x \in X} [u' f^{(1)}(x)]^2\, \xi(x) \leqq \sum_{x \in X} \max_{\|u\|=1} [u' f^{(1)}(x)]^2\, \xi(x)$$

But

$$\max_{\|u\|=1} [u' f^{(1)}(x)]^2 = \left[\frac{f^{(1)}(x)}{\|f^{(1)}(x)\|} f^{(1)}(x)\right]^2 = \|f^{(1)}(x)\|^2$$

hence

$$\max_i \varrho_i \leqq \sum_{x \in X} \|f^{(1)}(x)\|^2\, \xi(x)$$

The mapping $f : X \to \mathbb{R}^m$ is continuous, X is a compact set, therefore such a number A exists that

$$\max_{x \in X} \|f^{(1)}(x)\|^2 = \mathrm{A} < \infty$$

This implies that

$$\max_i \varrho_i \leqq \sum_{x \in X} \|f^{(1)}(x)\|^2\, \xi(x) \leqq \mathrm{A} \sum_{x \in X} \xi(x) = \mathrm{A} \tag{3.15}$$

that is, the eigenvalues of M_{I} are bounded.

From the inequality (3.15) follows

$$\min_i \varrho_i^{-1} \geqq 1/\mathrm{A} > 0$$

According to (3.10) and (3.12) is

$$\prod_{i=1}^{m} \varrho_i^{-1} \leqq \exp\left\{d + \frac{\lambda^2}{4r}\right\}$$

Then for all j's is (according to (3.15))

$$\varrho_j^{-1} \leqq \mathrm{A}^{m-1} \exp\left\{d + \frac{\lambda^2}{4r}\right\}$$

and according to (3.9)

$$\|M_{\mathrm{I}}^{-1}\| \leqq \sqrt{(m)}\left(\mathrm{A}^{m-1} \exp\left\{d + \frac{\lambda^2}{4r}\right\}\right) \qquad \square$$

Corollary 3.1. If (3.6) is satisfied by matrices M and $\overline{M}$, then there exists a number H such that

$$(3.17) \qquad \frac{\|\overline{M}^{-1}pp'\overline{M}^{-1} - M^{-1}pp'M^{-1}\|}{\|\overline{M} - M\|} \leqq \mathrm{H}$$

Proof.

$$\|\overline{M}^{-1}pp'\overline{M}^{-1} - M^{-1}pp'M^{-1}\| \leqq \|(\overline{M}^{-1} - M^{-1})\,pp'(\overline{M}^{-1} + M^{-1})\| \leqq$$
$$\leqq \|\overline{M}^{-1} - M^{-1}\|\;\|pp'\|\;\|\overline{M}^{-1} + M^{-1}\| \leqq$$
$$\leqq \|\overline{M}^{-1}\|\;\|\overline{M} - M\|\;\|M^{-1}\|\;(\|\overline{M}^{-1}\| + \|M^{-1}\|)$$

The previous lemma implies then the existence of H. □

Lemma 3.2. Let $v(M)$ be a function defined on $\mathfrak{M}_+$ and let Z be some real number. If

$$\frac{|v(M) - v(\overline{M})|}{\|M - \overline{M}\|} < \mathrm{Z}\,; \quad (M, \overline{M} \in \mathfrak{M}_+)$$

then

$$\frac{|\theta(v(M)) - \theta(v(\overline{M}))|}{\|M - \overline{M}\|} < \mathrm{Z}$$

where $\theta(v(M)) = \max\{v(M), 0\}$.

Proof. The proof is evident if we count the value of $|\theta(v(M)) - \theta(v(\overline{M})|$. □

Theorem 3.1. The condition (D) is fulfilled for the function

$$\mathrm{L}_r(M, \lambda) = -\ln\det M_\mathrm{I} + \frac{1}{4r}\{\theta^2(\lambda + 2r(p'M^{-1}p - c)) - \lambda^2\}$$

Proof. For all M, $\overline{M} \in \mathfrak{M}_+$ is according to (3.5)

$$(3.18) \qquad \frac{\|\nabla\mathrm{L}_r(M, \lambda) - \nabla\mathrm{L}_r(\overline{M}, \lambda)\|}{\|M - \overline{M}\|} \leqq \frac{\left\|\begin{pmatrix}\overline{M}_\mathrm{I}^{-1} & 0\\ 0 & 0\end{pmatrix} - \begin{pmatrix}M_\mathrm{I}^{-1} & 0\\ 0 & 0\end{pmatrix}\right\|}{\|M - \overline{M}\|} +$$

$$+ \frac{\theta[\lambda + 2r(p'\overline{M}^{-1}p - c)]\,(\overline{M}^{-1}pp^1\overline{M}^{-1}) - \theta[\lambda + 2r(p'M^{-1}p - c)]\,(M^{-1}pp'M^{-1})}{\|M - \overline{M}\|}$$

Now, we shall find the upper bound for both members in the righthand side of (3.18).

Let M, $\overline{M} \in \mathfrak{M}_d$. Then

$$1)\ \left\|\begin{pmatrix}\overline{M}_\mathrm{I}^{-1} & 0\\ 0 & 0\end{pmatrix} - \begin{pmatrix}M_\mathrm{I}^{-1} & 0\\ 0 & 0\end{pmatrix}\right\|^2 = \|\overline{M}_\mathrm{I}^{-1} - M_\mathrm{I}\|^2 \leqq \|\overline{M}_\mathrm{I}^{-1}\|^2\,\|\overline{M}_\mathrm{I} - M_\mathrm{I}\|^2\,\|M_\mathrm{I}^{-1}\|^2 \leqq$$
$$\leqq m\left(\mathrm{A}^{m-1}\exp\left\{d + \frac{\lambda^2}{4r}\right\}\right)^2\|\overline{M}_\mathrm{I} - M_\mathrm{I}\|^2$$

the last inequality being the consequence of Lemma 3.1. Therefore

$$\frac{\|\bar{M}_I^{-1} - M_I^{-1}\|}{\|\bar{M} - M\|} \leqq \sqrt{(m)}\, A^{m-1} \exp\left\{d + \frac{\lambda^2}{4r}\right\} \tag{3.19}$$

2) Let

$$l(M) = \theta[\lambda + 2r(p'M^{-1}p - c)]$$
$$h(M) = M^{-1}pp'M^{-1}$$

then

$$\frac{\|l(\bar{M}) \, . \, h(\bar{M}) - l(M) \, . \, h(M)\|}{\|M - \bar{M}\|} =$$

$$= \frac{\|l(\bar{M})\, h(\bar{M}) - l(\bar{M})\, h(M) + l(\bar{M})\, h(M) - l(M)\, h(M)\|}{\|M - \bar{M}\|} \leqq$$

$$\leqq |l(\bar{M})| \frac{\|h(\bar{M}) - h(M)\|}{\|M - \bar{M}\|} + |l(\bar{M}) - l(M)| \, . \, \frac{\|h(M)\|}{\|M - \bar{M}\|}$$

According to the Corollary 3.1 there exists a number H such that

$$\frac{\|h(\bar{M}) - h(M)\|}{\|M - \bar{M}\|} \leqq \mathrm{H}$$

Now, we shall find the upper bound for the expıession

$$|l(\bar{M}) - l(M)| \frac{\|h(M)\|}{\|M - \bar{M}\|}$$

Because $M^{-1}pp'M^{-1}$ is a symmetric matrix, Lemma 3.1 implies the existence of a number Q such that for all $M \in \mathfrak{M}_d$ is $\|M^{-1}pp'M^{-1}\| \leqq \mathrm{Q}$, i.e., $\|h(M)\| \leqq \mathrm{Q}$.

Moreover

$$\|p'M^{-1}p - p'M^{-1}p\| = \|\mathrm{Tr}\, pp'M^{-1} - \mathrm{Tr}\, pp'\bar{M}^{-1}\| =$$
$$= |\mathrm{Tr}\, pp'(M^{-1} - \bar{M}^{-1})| = \langle pp', M^{-1} - \bar{M}^{-1}\rangle \leqq$$
$$\leqq \|pp'\| \, . \, \|M^{-1} - \bar{M}^{-1}\| = \|M^{-1} - \bar{M}^{-1}\|$$

that is

$$\frac{|p'M^{-1}p - p'\bar{M}^{-1}p|}{\|M - \bar{M}\|} \leqq \frac{\|M^{-1} - \bar{M}^{-1}\|}{\|M - \bar{M}\|}$$

and according to Lemma 3.1 and the relation (3.17) there exists a number R′ such hat

$$\frac{|l(M) - l(\bar{M})|}{\|M - \bar{M}\|} \leqq \mathrm{R}' \quad \text{for all} \quad M, \bar{M} \in \mathfrak{M}_d$$

We have proved the existence of the numbers H, Q, and R′ that

$$\frac{\|l(\bar{M})\, h(\bar{M}) - l(M)\, h(M)\|}{\|M - \bar{M}\|} \leqq l(M) \, . \, \mathrm{H} + \mathrm{Q} \, . \, \mathrm{R}' \, .$$

□

Example 1. (See [17])

Let $X = \{x \in R : -1 \leqq x \leqq 1\}$ and let the original model be

$$E(y(x)) = \alpha_1 + \alpha_2 x \quad (3.20)$$

The model adequacy will be tested by embedding (3.20) in the more general model

$$E(y(x)) = \alpha_1 + \alpha_2 x + \alpha_3 x^2 \quad (3.21)$$

and we shall test the hypothesis

$H_0 : \alpha_3 = 0$ against the alternative hypothesis

$H_1 : \alpha_3 \neq 0$.

We set

$$\Phi(M) = -\ln \det M_1$$
$$\Psi(M) = p'M^{-1}p - c$$

where $p' = (0, 0, 1)$. We demand the 80 per cent efficiency for estimating α_3 in (3.21). According to [4] we determine $c_0 = 4$ and according to (2.3) we get $c = 5$. The hybrid optimal design was obtained in two steps. The results of our algorithm are tabulated below.

Initial data:

R = 1·00, λ = 0·00, c = 5, EPS = 0·001
$\xi(-1) = 0{\cdot}33$, $\xi(0) = 0{\cdot}33$, $\xi(1) = 0{\cdot}33$

Iteration 1.

R = 1·00, λ = 0·00
$\xi(-1) = 0{\cdot}36443$. $\xi(0) = 0{\cdot}27111$, $\xi(1) = 0{\cdot}364443$

Iteration 2.

R = 1·00, λ = 0·121006
$\xi^*(-1) = 0{\cdot}361758$, $\xi^*(0) = 0{\cdot}276480$, $\xi^*(1) = 0{\cdot}361756$.

Stigler [17] has found the hybrid optimal design for model (3.20) according to completely different method, and his results were $\xi^*(-1) = 0{\cdot}3618$, $\xi^*(0) = 0{\cdot}2763$, $\xi^*(1) = 0{\cdot}3618$. Stigler's method cannot be used for other models. In view of Theorem 3.1 it is obvious that for the case of D-optimality our algorithm will work well for various models.

Example 2.

Let $X = \{x \in R : -1 \leqq x \leqq 1\}$ and let the original model be

$$E(y(x)) = \alpha_1 + \alpha_2 x + \alpha_3 x^2 \quad (3.22)$$

The model adequacy will be tested by embedding (3.22) in the more general model

$$\mathsf{E}(y(x)) = \alpha_1 + \alpha_2 x + \alpha_3 x^2 + \alpha_4 x^3 \tag{3.23}$$

and we shall test the hypothesis

$H_0 : \alpha_4 = 0$ against the alternative hypothesis
$H_1 : \alpha_4 \neq 0$.

We set again

$$\Phi(M) = -\ln \det M_1$$
$$\Psi(M) = p' M^{-1} p - c$$

where $p' = (0, 0, 0, 1)$, and we demand the 80 per cent efficiency for estimating α_4 in (3.23). The results are given below.

Initial data:

$R = 1{\cdot}00$, $\lambda = 0{\cdot}00$, $c = 20$, EPS $= 0{\cdot}05$
$\xi(-1) = 0{\cdot}25$, $\xi(-0{\cdot}45) = 0{\cdot}25$, $\xi(0{\cdot}45) = 0{\cdot}25$, $\xi(1) = 0{\cdot}25$

Iteration 1.

$R = 1{\cdot}00$, $\lambda = 0{\cdot}00$
$\xi(-1) = 0{\cdot}2707$, $\xi(-0{\cdot}45) = 0{\cdot}1874$, $\xi(0) = 0{\cdot}0833$, $\xi(0{\cdot}45) = 0{\cdot}2707$,
$\xi(1) = 0{\cdot}1874$

Iteration 2.

$R = 1{\cdot}00$, $\lambda = 0{\cdot}00$
$\xi^*(-1) = 0{\cdot}2737$, $\xi^*(-0{\cdot}529) = 0{\cdot}3007$, $\xi^*(0{\cdot}529) = 0{\cdot}2407$, $\xi^*(1) = 0{\cdot}1874$

ACKNOWLEDGMENT

I am deeply indebted to my advisor, RNDr. Andrej Pázman, DrSc., for his help and many useful discussions.

(Received March 1, 1982.)

REFERENCES

[1] A. C. Atkinson: Planning experiments to detect inadequate regression models. Biometrika *59* (1972), 275—293.

[2] A. C. Atkinson: Planning experiments for model testing and discrimination. Math. Operationsforsch. Statist. *6* (1975), 252—267.

[3] D. M. Borth: A total entropy criterion for the dual problem of model discrimination and parameter estimation. J. Roy. Statist. Soc. Ser. B *37* (1975), 77—87.

[4] V. V. Fedorov: Theory of Optimal Design. Nauka, Moskva, 1971. In Russian.

[5] V. V. Fedorov and A. Pázman: Design of physical experiments. Fortschr. Phys. *16* (1968), 325—355.

[6] M. Hamala: Nonlinear Programming. ALFA, Bratislava 1972. In Slovak.

[7] W. J. Hill, W. G. Hunter, and D. W. Wichern: A joint design criterion for the dual problem of model discrimination and parameter estimation. Technometrics *10* (1968), 145—160.
[8] E. Läuter: Experimental design in a class of models. Math. Operationsforsch. Statist. *5* (1974), 379—398.
[9] J. Mikulecká: A Hybrid Optimal Design of Experiments. Ph. D. Dissertation, Comenius University, Bratislava 1981. In Slovak.
[10] A. Pázman: A convergence theorem in the theory of D-optimum experimental designs. Ann. Statist. *2* (1974), 216—218.
[11] A. Pázman: Foundations of Optimization of Experiments. Veda, Bratislava 1980. In Slovak.
[12] A. Pázman: Some features of the optimal design theory — A survey. Math. Operationsforsch. Statist. Ser. Statist. *11* (1980), 415—446.
[13] M. J. D. Powell: A method for nonlinear constrains in minimization problems. In: Optimization (R. Fletcher, ed.), Academic Press, London—New York 1969.
[14] R. T. Rockafellar: Convex Analysis. Princeton University Press, Princeton 1970.
[15] R. T. Rockafellar: The multiplier method of Hestens and Powell applied to convex programming. J. Optim. Theory Appl. *12* (1973), 555—561.
[16] R. T. Rockafellar: A dual approach to solving nonlinear programming problems by unconstrained optimization. Math. Programming *5* (1973), 354—373.
[17] S. M. Stigler: Optimal experimental design for polynomial regression. J. Amer. Statist. Assoc. *66* (1971), 311—318.

RNDr. Jaroslava Mikulecká, CSc., Katedra kybernetiky SVŠT (Department of Cybernetics — Slovak Technical University), Gottwaldovo nám. 17, 812 43 Bratislava. Czechoslovakia.