Proceedings of the 1989 COSMEX Meeting

STOCHASTIC METHODS in EXPERIMENTAL SCIENCES

Szklarska Poreba, Poland 8 – 14 September 1989

Editors
W. Kasprzak
A. Weron

World Scientific
Singapore • New Jersey • London • Hong Kong

COSMEX – 89
w. Kasprzak & A. Weron (Eds)
World Scientific, Singapore 1990

COMPUTATIONAL ALGORITHM FOR GENERATING OPTIMUM EXPERIMENTAL DESIGNS ON A HYPERCUBE

Ewaryst Rafajłowicz* Wojciech Myszka†

Abstract

In the paper the computational algorithm for finding optimum experimental design is proposed. The algorithm is based on authors' recent result indicating that for Hoel's type regression on a hypercube and for a large class of optimality criterions one can find the optimum design which is a product of designs for partial regressions. Consequently, the Wynn–Fedorov algorithm is applied to find optimal designs for partial regressions and then they are composed into the optimal design for all variables. It is also shown that the upper bound for the number of support points in the design generated by the algorithm is essentially smaller than provided by the commonly used upper bound.

1 INTRODUCTION

We consider the optimum experimental design problem for regression function $EY(x) = a^T f(x)$ in the classical setting (see, e.g., Fedorov (1972), Pazman (1986), Silvey (1980)). Above, $a \in R^m$ is the column vector of unknown parameters, while $f(x)$ is the vector of prescribed functions, which are continuous and linearly independent over a compact set $X \subset R^p$. For simplicity we assume $\mathrm{var}(Y(x)) = 1$. Our aim is to choose an experimental design, which is from the class $\Xi(X)$ of all probability measures on X including all discrete measures. This choice is based on the information matrix

$$M(\xi) = \int_X f(x) f^T(x) \xi(dx), \tag{1}$$

as well as on an optimality criterion $\Phi(M(\xi); m)$, which is a real valued function of the information matrix $M(\xi)$ and its dimension m.

*Technical University of Wrocław, Institute of Engineering Cybernetics, 50-370 Wrocław, Janiszewskiego 11, POLAND

†Technical University of Wrocław, Institute of Material Sciences and Technical Mechanics

Every design $\hat{\xi} \in \Xi(X)$ such that

$$\sup_{\xi \in \Xi(X)} \Phi(M(\xi); m) = \Phi(M(\hat{\xi}); m) \tag{2}$$

is called Φ – optimum. The problem (2) will be further called (Φ, m, f, X) optimum experimental design problem.

Numerical algorithms for generating a sequence ξ_n, $n = 1, 2, \ldots$ convergent to $\hat{\xi}$ have been proposed by Wynn (1970), Fedorov (1972), Wu (1978), Wu and Wynn (1978) and Atwood (1978). Available algorithms are summarized in Silvey (1980) and Pazman (1986), where also bibliography on generating exact N–point design can also be found. Here we concentrate our attention on generating the so called continuous Φ – optimum designs, which approximate N–point designs when N is large.

Roughly speaking, existing algorithms can be classified as follows.

I Algorithms operating in X, which generate both support points and weights of a design. This class of methods is based on ideas of Wynn and Fedorov.

II Methods updating only weights of a design, which is supported in a finite number of prescribed points (see, e.g.,Pazman (1986) Chapter V for their description).

In practice a set of support points used by the algorithms of class II is identified by a method from class I. For this reason we concentrate our attention on improvements of algorithms of type I. In order to simplify further references we shortly describe such an algorithm for differentiable and concave global optimality criterion. Let $F(M; m)$ be an $m \times m$ matrix with elements $\frac{\partial \Phi(M;m)}{\partial m_{ij}}$, where m_{ij} are elements of M. Define

$$d(x; M, m) = f^T(x) F(M; m) f(x), \tag{3}$$

and let $\xi_0 \in \Xi(X)$ be an arbitrary starting design with $M(\xi_0)$ nonsingular.

Let $\xi_n \in \Xi(X)$ be a design at n–th iteration. Then generalised algorithm of Wynn and Fedorov, further denoted by $AWF(\Phi, m, f, X)$, runs as follows (see Pazman (1986) page 157).

At the (n+1) – st step

$$\xi_{n+1} = (1 - \alpha_n)\xi_n + \alpha_n \delta(x_n) \tag{4}$$

where $\delta(x_n)$ is the measure lumped at x_n, which is defined by

$$\sup_{x \in \mathcal{X}} d(x; M(\xi_n), m) = d(x_n; M(\xi_n), m). \tag{5}$$

The sequence of numbers α_n, $n = 1, 2, \ldots$ is such that

$$\begin{array}{c} \alpha_n \in (0,1), \\ \sum_{n=1}^{\infty} \alpha_n = \infty, \\ \lim \alpha_n = 0, \end{array} \tag{6}$$

or it is chosen so as to maximise the function

$$\Phi(M\{(1-\alpha)\xi_n + \alpha\delta(x_n)\}; m) \tag{7}$$

with respect to $\alpha \in [0,1]$.

Sufficient conditions for convergence of this algorithm can be found in Pazman (1986) pages 157-161.

The most difficult and time consuming step in $AWF(\Phi, m, f, X)$ is maximisation (5). This step usually requires to use an iterative procedure for finding the global maximum of $d(x; M(\xi_n), m)$ constrained to X. As is known, it is one of the most difficult tasks of the numerical analysis and its difficulty grows nonlinearly with the growth of dimension X. These facts motivate our study, which is directed to overcoming difficulties of multidimensional optimization by decomposing (2) into a number of experimental design problems for partial regression functions.

This approach is feasible for sufficiently broad class of criterions and regression functions on a hypercube, as it will be indicated in the next section, using results from Rafajłowicz and Myszka (1988). In section 3 the proposed algorithm is described, while in section 4 bounds for a number of support points in the optimal design are derived.

2 PRELIMINARIES

We consider the regression function $EY(x) = a^T f(x)$ of a special structure, which was introduced by Hoel (1965) for two independent variables. Namely, the column vector $f(x)$ of continuous and linearly independent functions on a compact set X is of the form:

$$f(x) = g_1(x(1)) \otimes g_2(x(2)) \otimes \cdots \otimes g_r(x(r)) \hat{=} \bigotimes_{i=1}^{r} g_i(x(i)), \tag{8}$$

where $x(i)$, $i = 1, 2, \ldots, r$ are subvectors of the vector x. Furthermore, every $x(i) \in X_i$, where X_i is compact set and $X = X_1 \times X_2 \times \cdots \times X_r$. Column vectors $g_i(x(i))$ of continuous and linearly independent functions are defined on X_i with values in R^{m_i}. By $\otimes$ and $\bigotimes$ we denote Kronecker's or the direct product of matrices or vectors (see, e.g., Marcus and Minc (1964), Lankaster (1969), Graham (1981) for definition). Now $a \in R^m$ denotes the vector of unknown parameters with $m \hat{=} \prod_{i=1}^{r} m_i$ elements.

In order to indicate that the class of regression functions spanned by $f(x)$ of the form (8) is very large we consider two examples.

Example 2.1 Let $g_i(x(i)) = [1, x(i)]^T$; $i = 1, 2, 3$ where $x(i) \in X_i \subset R$. Then,

$$EY(x) = a_1 + a_2x(2) + a_3x(1) + a_4x(1)x(2) + a_5x(3) + \\ + a_6x(2)x(3) + a_7x(1)x(3) + a_8x(1)x(2)x(3). \tag{9}$$

Let us note that we listed parameters a_i, $i = 1, 2, \ldots, 8$ in the same order as elements of $f(x)$.

Example 2.2 Let $g_1(x(1)) = [1, x(1)]^T$, $g_2(x(2)) = [1, \sin(x(2)), \cos(x(2))]^T$, $x(i) \in [0, 2\pi]$, $i = 1, 2$. In this case

$$EY(x) = a_1 + a_2 \sin(x(2)) + a_3 \cos(x(2)) + a_4x(1) + \\ + a_5x(1)\sin(x(2)) + a_6x(1)\cos(x(2)). \tag{10}$$

When part of independent variables is fixed at prescribed levels $x(1) = x^*(1)$, $x(2) = x^*(2), \ldots, x(s) = x^*(s)$, say, then one may consider $a^T f(x)$ as a regression function of $x(s+1), \ldots, x(r)$, spanned by $g_i(x(i))$, $i = (s+1), \ldots, r$.

Such a function is further called a partial regression function.

The reader may also consult Kasprzak and Lysik (1974), where the methodology of step-by-step model building is proposed and illustrated by many examples from different disciplines.

Note that the orthogonal projection of the graph of the function $a^T f(x)$ onto X_j is of the form $b^T g_j(x(j))$. Thus, one can choose suitable base functions constituting $g_i(x(i))$ by looking at orthogonal projections of the data and following ideas of projection pursuit regression.

Our main interest in this paper is to propose an efficient numerical algorithm for solving the problem (1), (2) for regression functions spanned by $f(x)$ of the form (8). We confine our attention to the design optimality criterions for which the following conditions hold.

A1) **Φ is concave and differentiable in S^+, which is the subset of non-singular matrices in $S \hat{=} \{M(\xi) : \xi \in \Xi\}$.**

Let $F(M; m)$ be an $m \times m$ matrix with elements $\frac{\partial \Phi[M;m]}{\partial m_{ij}}$, where m_{ij} are elements of M.

A2) **The matrix function $F(M; m)$ is defined on S^+ and for every $c > 0$ one can find $0 < K_c < \infty$ so that for every $M, \hat{M} \in \{M \in S : \Phi(M; m) < c\}$**

$$\|F(M; m) - F(\hat{M}; m)\| \leq K_c \|M - \hat{M}\|, \tag{11}$$

where $\|\cdot\|$ denotes the matrix norm.

Discussion of conditions A1), A2) can be found in Pasman (1986) pages 158-161. They are used to ensure convergence of the Wynn–Fedorov algorithm and hold for a large class of optimality criterions. The next conditions are somewhat more restrictive but they allow to split the design problem into lower dimensional subproblems of finding designs in the sets $\Xi(X_i)$ of all probability measures on X_i. Note that the product measures

$$\xi(dx) = \prod_{i=1}^{r} \xi^{(i)}(dx(i)),\ \xi^{(i)} \in \Xi(X_i) \tag{12}$$

form a subset of $\Xi(X)$, denoted by $\Xi\Pi$. Elements of $\Xi\Pi$ are further called product type experimental designs. We also need the following algebraic results: for every $\xi \in \Xi\Pi$

$$M(\xi) = \bigotimes_{i=1}^{r} M^{(i)}(\xi^{(i)}), \tag{13}$$

where, for $\xi^{(i)} \in \Xi(X_i)$,

$$M^{(i)}(\xi^{(i)}) \doteq \int_{X_i} g_i(x(i)) g_i^T(x(i)) \xi^{(i)}(dx(i)). \tag{14}$$

Equation (13) can be directly verified using well known properties of Kronecker's product (see, e.g., Lankaster (1969)). This equality motivates our assumption.

A3) If $M = \otimes_{i=1}^{r} M^{(i)}$, then

$$F(M; m) = \bigotimes_{i=1}^{r} F(M^{(i)}; m^{(i)}), \tag{15}$$

where $M, M^{(i)}$ are $m \times m$ and $m^{(i)} \times m^{(i)}$ matrices, respectively, and $m = \prod_{i=1}^{r} m^{(i)}$.

Taking (13) into account, one can express assumption A3) as follows. We consider criterion functions with gradients factorisable (in the Kronecker sense) for all product type experimental designs.

Our last assumption is that

A4) **$F(M; m)$ is nonnegative definite matrix for all $m \times m$ matrices $M \in S^+$ and every $m = 1, 2, \ldots$.**

The following examples shows that A3), A4) hold for many commonly used optimality criterions (see, e.g., Fedorov (1982), Pasman (1986), Silvey (1980) for definitions and formulas for gradients).

D – optimality criterion. For $M \in S^+$, $\Phi(M;m) = \ln\det(M)$, where $\det(\cdot)$ denotes the determinant, while $F(M;m) = M^{-1}$. Thus, A3) follows from the known fact that for nonsingular matrices A, B we have

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}, \tag{16}$$

see, e.g., Lankaster (1969).

L_p – optimality. For positive integer p and $M \in S^+$,

$$\Phi(M;m) = -[\frac{1}{m}\mathrm{tr}(M^{-p})]^{\frac{1}{p}}$$

with the gradient

$$F(M;m) = m^{-\frac{1}{p}}[\mathrm{tr}(M^{-p})]^{\frac{(1+p)}{p}} M^{-(p+1)}. \tag{17}$$

Now A3) follows from (16) and the equality

$$\mathrm{tr}[\bigotimes_{i=1}^{r} M^{(i)}]^{-p} = \prod_{i=1}^{r} \mathrm{tr}[M^{(i)}]^{-p}.$$

Let us note that L_p – class contains the following criteria:

A – optimality for $p = 1$,

D – optimality for $p \to 0+$,

E – optimality for $p \to \infty$.

Q–optimality. For $M \in S^+$ $\Phi(M;m) = -\int_X f^T(x)M^{-1}f(x)dx$ and $F(M;m) = M^{-1}\bigotimes_{i=1}^{r} G_i M^{-1}$, where $G_i \hat{=} \int_{X_i} g_i(x(i))g_i^T(x(i))dx(i)$. Now A3) follows from the following property

$$(A \otimes B)(C \otimes D) = AC \otimes BD, \tag{18}$$

valid for matrices of appropriate dimensions.

Extrapolation at a point. Let $z \in X$ be the point at which prediction with the largest accuracy is desired and let $z = [z(1), z(2), ..., z(r)]^T$ be its partition $z(i) \in X_i, i = 1, 2, \ldots, r$. Then, $\Phi(M,m) = -f^T(z)M^{-1}f(z)$ with the gradient

$$F(M;m) = M^{-1}\bigotimes_{i=1}^{r} g_i(z(i))g_i^T(z(i))M^{-1}$$

and the verification of A3) is the same as above.

L–optimality. The last two examples suggest that A3) holds also for L – optimality criterion $\Phi(M;m) = -\mathrm{tr}(M^{-1}W)$ if the user defined weight matrix W is factorisable as follows $W = \otimes_{i=1}^{r} W(i)$, where $W(i)$ are nonnegative definite $m^{(i)} \times m^{(i)}$ matrices. In this case $F(M;m) = M^{-1} \otimes_{i=1}^{r} W(i) M^{-1}$ and A3) follows from (16) and (18).

The computational algorithm proposed in the next section is based on the following result proved in Rafajlowicz and Myszka (1988).

Theorem 1 *Let the regression function be spanned by $f(x)$ of the form (8). Consider the problem (2) with the criterion, for which A1), A3), A4) hold. Denote by $\hat{\xi}^{(i)} \in \Xi(X_i)$ the optimum experimental designs for partial regressions spanned by $g_i(x(i))$, which are defined by*

$$\sup_{\xi^{(i)} \in \Xi(X_i)} \Phi(M^{(i)}(\xi^{(i)}); m_i) = \Phi(M^{(i)}(\hat{\xi}^{(i)}); m_i) \; i = 1, 2, ..., r. \tag{19}$$

Then, for the whole regression, spanned by $f(x)$ one can find (Φ, m, f, X) – optimum experiment design $\hat{\xi} \in \Xi(X)$ of the form

$$\hat{\xi}(dx) = \prod_{i=1}^{r} \hat{\xi}^{(i)}(dx(i)). \tag{20}$$

In other words, there exists the optimal design, which can be composed from optimal designs for all partial regressions.

3 COMPUTATIONAL ALGORITHM

Theorem 1 suggests the following algorithm

Step 0. Choose the criterion Φ for which A1) – A4) hold. For $i = 1, 2, \ldots, r$ perform steps 1–4.

Step 1. Define X_i and $g_i(x(i))$.

Step 2. Choose a starting point $\xi^{(i)} \in \Xi(X_i)$ with $M^{(i)}(\xi^{(i)})$ nonsingular.

Step 3. Execute $AWF(\Phi, m_i, g_i, X_i)$ according to (4),(5) and (6) or (7).

Step 4. Store the (near-)optimal design $\hat{\xi}^{(i)} \in \Xi(X_i)$ obtained in Step 3.

Step 5. Compose the (near-)optimal design for the whole regression according to (20).

The above algorithm has been implemented in FORTRAN 77 on IBM PC computers by the present authors. Few remarks concerning implementation are in place.

Remark 3.1 In the implementation X_i are chosen to be user defined intervals, while $g_i(x(i))$ can be chosen from the following systems of functions:

a) the classical polynomials,

b) the Tschebyschew polynomials,

c) the Legendre polynomials,

d) the Laguerre polynomials,

e) the trigonometric system.

Note that each partial regression can be spanned by its own system a), b), c), d) or e).

All the criterions listed in Section 2 are available for the program user. It is clear, that some combinations of criterions and regressions are redundant, e.g., we obtain the same partial design for polynomials a), b), c) with D–optimality criterion. However not all criterions are invariant to nonsingular reparametrizations.

Remark 3.2 In the implementation of Step 3 the following subtasks are involved:

1. Global, one-dimensional optimisation, for finding the next point to be added to the currently available design. It is performed in two steps. Namely, a rough search on a uniform grid is made in order to locate the interval containing the global maximum, then the golden search method is applied to locate it with user's defined accuracy.
2. In order to avoid designs supported in too many unnecessary points, the following measures have been taken.
   (a) At each iteration of Wynn–Fedorov algorithm we check the possibility of improving current design by rejecting a point from its support. This is done according to the known method described in Pazman (1986) page 144, but we allow to reject a point even if it can lead to some increase of the criterion function in a current iteration. This occures to be an efficient tool in speeding up the algorithm.
   (b) Before adding a point to the support of existing design we check whether it is possible to "glue" it with earlier added points. By this we mean that very close points are not distinguished (with user's defined resolution) and their weights are added.
   (c) Before performing Step 4 the user can remove points from the design.

The standard stopping rule is applied at Step 3. It is based on checking whether the optimality condition is fulfilled (with a specified accuracy) by currently available design (see, e.g.,Pazman (1986) Theorem (V.28)).

Remark 3.3 Design for partial regressions, obtained at Step 3, are stored in the form $(x_j(i), p_j(i)), j = 1, 2, \ldots, J_i, i = 1, 2, \ldots, r$. Thus, the final step consists of generating all possible tuples $(x_{j_1}(1), x_{j_2}(2), \ldots, x_{j_r}(r))$, and the corresponding weights given by $p_{j_1}(1)p_{j_2}(2)\ldots p_{j_r}(r)$.

4 BOUNDS FOR THE NUMBER OF SUPPORT POINTS

At a first glance, one may think that the algorithm proposed in Section 3 generates designs supported in excessively large number of points. In fact the things are not so bad. Namely, we shall show that above strategy of composing multidimensional designs allows to reduce an upper bound for the cardinality of the optimal design support in comparison to the known bound. As it is well known (see, e.g., Fedorov (1982)), for every design $\xi \in \Xi(X)$ one can find a design with the same information matrix, which is lumped at a finite number of N points, say, and such that

$$m \le N \le N_u \doteq \frac{m(m+1)}{2} + 1. \tag{21}$$

The above statement holds under the assumptions of Section 1, where m is defined as $\dim a = \dim f(x)$. Furthermore, if the optimality criterion is such that $M(\xi)$ lies on the boundary of the set $\{M(\xi) : \xi \in \Xi(X)\}$, then

$$m \le N \le \hat{N}_u \doteq \frac{m(m+1)}{2}. \tag{22}$$

We remark that inequalities in (21), (22) are sharp, i.e., for a general experiment design problem examples are known where equalities hold.

Let us assume that all the assumptions of Theorem 1 hold. Denote by $L(i)$ the number of support points in the design problem (Φ, m_i, g_i, X_i). Applying inequalities (21), (22) to this problem we obtain, respectively

$$m_i \le L(i) \le L_u(i) = \frac{m_i(m_i+1)}{2} + 1, \tag{23}$$

$$m_i \le L(i) \le \hat{L}(i) = \frac{m_i(m_i+1)}{2}. \tag{24}$$

Let L denotes the number of support points in the design obtained at the Step 5 of our algorithm. According to Remark 3.3 we obtain from (23) and (24), respectively

$$m \le L \le L_u = \prod_{i=1}^{r} L(i), \tag{25}$$

Table 1: Ratio $q(k,r) = \frac{L_u}{N_u}$, defined by (27), (28)

r	k								
	2	3	4	5	6	7	8	9	10
2	1.45	1.06	0.88	0.785	0.726	0.686	0.659	0.637	0.621
3	1.73	0.90	0.64	0.520	0.454	0.413	0.386	0.366	0.351
4	1.86	0.72	0.44	0.335	0.279	0.245	0.223	0.208	0.197
5	1.93	0.56	0.30	0.215	0.170	0.145	0.129	0.118	0.110
10	2.00	0.16	0.05	0.010	0.000	0.000	0.000	0.000	0.000

Table 2: Ratio $\hat{q}(k,r) = \frac{\hat{L}_u}{\hat{N}_u}$, defined by (27), (28)

	k								
	2	3	4	5	6	7	8	9	10
2	0.90	0.80	0.74	0.69	0.66	0.64	0.62	0.61	0.60
3	0.75	0.57	0.48	0.43	0.40	0.38	0.36	0.34	0.33
4	0.60	0.39	0.30	0.26	0.24	0.22	0.20	0.19	0.18
5	0.46	0.26	0.19	0.15	0.14	0.13	0.12	0.11	0.10
10	0.11	0.03	0.02	0.01	0.00	0.00	0.00	0.00	0.00

$$m \leq L \leq \hat{L}_u = \prod_{i=1}^{r} \hat{L}(i), \tag{26}$$

where $m = \prod_{i=1}^{r} m_i$.

In order to compare the known upper bounds N_u and $\hat{N}_u$ with bounds L_u and $\hat{L}_u$ in a transparent way, let us assume that all partial regressions are spanned by the same number of functions, i.e. we temporarily assume $m_i = k \geq 2; i = 1, 2, \ldots, r$. In this case $m = k^r$, what yields

$$N_u = \frac{k^r(k^r+1)}{2} + 1, \ \hat{N}_u = \frac{k^r(k^r+1)}{2} \tag{27}$$

$$L_u = \lfloor \frac{k(k+1)}{2} + 1 \rfloor^r, \ \hat{L}_u = \lfloor \frac{k(k+1)}{2} \rfloor^r \tag{28}$$

Let us define the ratios $q(k,r) = \frac{L_u}{N_u}$ and $\hat{q}(k,r) = \frac{\hat{L}_u}{\hat{N}_u}$. Their values are tabulated for different k, r in Tables 1 and 2, respectively.

Analysis of these tables leads to the following conclusions.

1. For $k > 2$ and every $r \geq 2$ upper bounds $L_u, \hat{L}_u$ are closer to the lower bound m than N_u and $\hat{N}_u$, respectively. Even for moderate numbers of independent

variables r and numbers of functions spanning partial regressions k reduction of the largest necessary points in the optimal design support can be essential, e.g., by half for $k = 4$ and $r = 3$ in Table 2.

2. The only exception is the case $k = 2$ when $q(k, r) > 1$ for all $r \geq 2$, but $q(2, r) \leq 2$. Besides, also in this case $\hat{q}(2, r) < 1$ for all $r \geq 2$.

That is easy to verify that if all the designs for partial regressions are saturated, then also the global design is saturated, i.e. its support contains the same number of points as the number of parameters.

5 EXAMPLES

Before presenting numerical results let us note that the proposed algorithm may also be used in calculations by hand when a closed form solutions for one-dimensional experimental design subproblems are known. This is illustrated by the following examples, in which we only sketch multidimensional optimal designs.

Example 5.1 Let $g_i(t) = [1, t, t^2, \ldots, t^{m_i-1}]^T$, $t \in [-1, 1]$, $i = 1, 2, \ldots, r$. Consider the optimum extrapolation experiment design to the point $x = [x(1), x(2), ..., x(r)]^T$, $|x(i)| > 1$, $i = 1, 2, \ldots, r$.

Then, the optimal design on the hypercube $[-1, 1]^r$ can be composed on the grid of Tschebyschew points, using the result from page 184 in Pazman (1986).

Example 5.2 Let $g_i(t) = [1, \sin t, \cos t, ...]^T$ with $m_i = k$ components, $i = 1, 2, \ldots, r$, $t \in [-\pi, \pi]$. Then, A-optimum experiment design on the hypercube $[-\pi, \pi]^r$ is supported at the knots of the uniform grid with the step size $\frac{1}{k}$. This design puts equal masses equal to k^{-r} at every support point. Optimality of this design follows from Theorem 1 and the result from page 186 in Pazman (1986).

Example 5.3 Let $g_i(t) = [1, t, t^2]^T$, $t \in [-1, 1]$, $i = 1, 2$. Using the above proposed algorithm the following design is obtained for L_p–optimality criterion with $p = 12$: 4 points $[\pm 1, \pm 1]$ enter with equal weigts $\frac{1}{25}$, while the central point $[0, 0]$ enters with the weight 0.36. The rest of the design is supported by the points $[0, \pm 1]$ and $[\pm 1, 0]$ with the weights 0.12.

Let us note that L_p criterion with $p = 12$ is in practice equivalent to E–optimality criterion.

Example 5.4 Let us consider the same regression function as in Example 5.3 on the unit square but with A–optimality criterion. Then, the same support points of the design are obtained with the following weights: $\frac{1}{16}$ at all the corners of the square, $\frac{1}{4}$ at $[0, 0]$ and $\frac{1}{8}$ at the middle points of all the edges.

Example 5.5 For the same regression as in Example 5.3 with Q–optimality criterion we obtain the same optimal design as in Example 5.4.

Let us note that having generated designs for one dimensional regression by AWF algorithm we can easily generalize Examples 5.3, 5.4 and 5.5 to more than two dimensions.

6 CONCLUDING REMARKS

As it can be seen from the above examples, the proposed algorithm is able to find approximate optimal experiment designs for regression with a relatively large number of independent variables in a reasonable computational time. The algorithm uses Wynn–Fedorov method for one-dimensional regression designs. Thus, any improvement of their method will make the proposed algorithm more efficient.

The computer code in FORTRAN 77 is available from the authors on request (for non – commercial use).

References

[1] Atwood C.L., Convergent design seqnences for sufficiently regular optimality criteria. Ann. Statist., vol. 4, pp 1124 - 1138, (1978).

[2] Fedorov V.V., Theory of Optimal Experiments. Academic Press, New York. (1982).

[3] Graham A., Kronecker Products and Matrix Calculus with Applications. Horwood, Chichester (1981).

[4] Hoel P., Minimax designs in two dimensional regression. Ann. Math. Statist., vol. 36, pp 1097 - 1106, (1965).

[5] Lankaster P., Theory of Matrices, Academic Press, NewYork, (1969).

[6] Marcus M, Minc H.. A Survey of Matrix Theory and Matrix Inequalities. Allyn and Bacon, Boston (1964).

[7] Pazman A., Foundations of Optimum perimental Designs, Reidel, Dordrecht, (1986).

[8] Rafajłowicz E., Myszka W., Optimum experimental design for a regression on a hypercube – Generalization of Hoel's result. Annals of the Institute of Statistical Mathematics, vol.40, No 4, pp 821 - 827 (1988).

[9] Silvey S.D., Optimal Design. Chapman & Hall, London (1980).

[10] Wu C.F., Some iterative procedures for generating nonsingular optimal designs. Comm. Statist. A - Theory Methods, vol.7, pp 1399 - 1412, (1978).

[11] Wu C.F., Wynn H.P., The convergence of general step–length algorithms for regular optimum design criteria. Ann. Statist., vol.6, pp 1273 - 1285 (1978).