A Combinatorial Search Technique for Constructing
Optimal Designs for Biased Models

Akpan S. S.

Department of Maths, Stats and Computer Science

University of Calabar, Calabar, Nigeria

Udoh, M. E

Department of Maths, Stats and Computer Science

University of Calabar, Calabar, Nigeria

Nduka, E. C.

Department of Mathematics & Statistics University of Port Harcourt

Port Harcourt, Nigeria

Abstract

A new algorithm has been proposed to locate the global minimum of the alias matrix for a biased response function defined in a set of distinct support points. The search begins by classifying all the support points in the experimental region of interest into groups. Then starting from an arbitrary set of N support points, the algorithm obtains an N-point optimal design by systematically adding and dropping support points from the various groups in such a way as to continuously reduce the determinant of the alias matrix of the design at each step of the sequence. Numerical demonstrations confirm the effectiveness of this algorithm.

Keywords: Combinatorics, Biased, Optimal designs, Construction, Optimality criterion

1. Introduction

Very often in design of experiments, experimenters are faced with the problem of choosing among a number of good competing designs based on certain criteria (Cook, R. D. 1980). Most of the criterion functions are functional form of the information matrix. For biased models, the optimality criterion will not be suitable because of its dependability on the model. One criterion for making this choice is by determining the determinant of the square of the bias or AL optimal designs from the class of all possible designs with the same characteristics. Hence given the triplet {,, ?x}, where is the space of all possible trials of an experiment, F_X is the space of finite dimensional continuous function defined on and is the space of non negative continuous random observation error defined on Allan, J. (1969), the problem is to obtain AL design measure ξN from for a given n-variate of a biased model such that the determinant of $A^T A$ of the associated design measure is minimized, where $A = (X_1^T X_1) X_1^T X_2$, X_1 is an N x n design matrix with the column space spanned by the vector x_1 and X_2 is an N x n matrix with column space spanned by vector $\underline{x}2$. We hereby introduce a sequential approach on the combinatorics of the possible design available for a given problem.

2. Methodology

For a given biased function f(x) and support point $\{x_1, x_2, \ldots, x_N\}$ $\varepsilon \tilde{X}$ where $\bar{N}$ represents the number of support points in a finite space of trials or the number of grid of points in an infinite space of trials, the support points are arranged into H concentric balls g1, g2, . . . , gn. (Notz, W. T. 1989) The jth ball contains nj support points and ? nj = j=1, 2,. . . , H. the support points in the jth ball are of equal distance dj each, from the center of and the distances of the support points in the H balls are such that d1>d2>. . . >dH. By some selection techniques, the non-promising designs are eliminated and the method moves along a steepest decent path to reach the required design. (Federov, V. V. 1972)

The search begins with an initial distribution of N support points amongst the balls such that g1 has N11 support points, g2 has N12 support points and so on. (Atkinson, A. C. 1988) By this distribution a baseline class has been defined where its determinant value is obtained. Starting from ball one in an increasing or decreasing order the direction of the search is defined by altering the number of support points in the other H-1 balls. For each alteration a determinant value is obtained. The system has converged when the determinant value is no longer decreasing.

3. The Algorithm

The $\bar{N}$ support points in $\tilde{X} = \{\underline{X}_1, \underline{X}_2, \underline{X}_3, ..., \underline{X}_{\bar{N}}\}$ are arranged into H concentric balls g1, g2, g3, . . . , gH such that

$$g_j = \left\{ \begin{array}{c} \underline{x}_{j1} \\ \underline{x}_{j2} \\ \vdots \\ \underline{x}_j n_j \end{array} \right\} ; \underline{x}'_{jk} = (x_{1jk} \; x_{2jk} \cdots x_{njk})$$

where

$\bar{N} = \sum n_j$ and

$$\left\{ \underline{x}'_{jk} \, \underline{x}_{jk} \right\}^{\frac{1}{2}} = s_j , \quad k = 1, \, 2, \, \cdots , \, n_j$$

i.e. $s_1 > s_2 >, \ldots, > s_H$

(a) For H = 2, the algorithm can be described in the following ways:

(1) Select n_{11} support points from g_1 and n_{12} support points from g2 and thus define the 2-tuple $\underline{U}_0 = (n_{11} \; n_{12})$, and note that there are a_0 available designs; where

$$a_0 = \left(\begin{array}{c} n_1 \\ n_{11} \end{array} \right) \left(\begin{array}{c} n_2 \\ n_{12} \end{array} \right) = a_{10}a_{20},$$

$n_1 > n_{11}, n_2 > n_{12}$ and $n_{11} + n_{12} = N$; $n_{11}, n_{12} \geq 0$;

if $n_1 < n_{11}, n_2 < n_{12}$ and $n_{11} + n_{12} = N$; $n_{11}, n_{12} \geq 0$.

Define $b_{ii} = n_{11} - \alpha n_1, b_{i2} = n_{12} - \alpha n_2$ where b_{11}, b_{12} are the residues and α is an integer such that $b_{ii} < n_1, b_{i2} < n_2$, then
$a_0 = \left(\begin{array}{c} n_1 \\ b_{11} \end{array} \right) \left(\begin{array}{c} n_2 \\ b_{12} \end{array} \right)$

(2) Define $T \left(\xi_N^{uv} \right) = A' \left(\xi_N^{uv} \right) A \left(\xi_N^{uv} \right), T \left(\xi_N^{uv} \right) \in S_w^{p \times p}$;

where $S_w^{p \times p}$ is the set of all non-singular matrices; $u = 1, \, 2, \, ..., \, a_{10}, \, v = 1, \, 2, \, ..., \, a_{20}$ and $w = 1, \, 2, \, ..., \, a_0$.

$$A \left(\xi_N^{uv} \right) = \xi_N^{11}, \xi_N^{21}, \cdots, \xi_N^{a_{10}1}, \xi_N^{12}, \xi_N^{22}, \cdots, \xi_N^{a_{10}2}, \xi_N^{13}, \xi_N^{23}, \cdots, \xi_N^{a_{10}3}, \cdots \xi_N^{a_{10}} \xi_N^{a_{20}}$$

where $\xi_N^{11} = \left(\begin{array}{c} \xi_{n_{11}}^{(1)} \\ \xi_{n_{12}}^{(1)} \end{array} \right), \xi_N^{21} = \left(\begin{array}{c} \xi_{n_{11}}^{(2)} \\ \xi_{n_{12}}^{(1)} \end{array} \right), \cdots, \xi_N^{10^1} = \left(\begin{array}{c} \xi_{n_{11}}^{(a_{10})} \\ \xi_{n_{12}}^{1} \end{array} \right), \cdots$

(3) Compute det $T \left(\xi_N^{uv} \right) = dw, w = 1, 2, \cdots, a_0$ and define $d_0 = \min\{d_w\}$.

(4) Refer to Table 1 for the rest of the steps leading to $d_1, \, d_2, \, ..., \, d_{t+1}$;

where m = k+t+2 and

$$d_1 \geq d_2 \geq \cdots \geq d_k < d_{k+1}; \quad d_{k+2} \geq d_{k+3} \geq \cdots \geq d_t < d_{t+1}... \tag{1}$$

(5) Define $d_c = \min\{d_0, d_k, \, d_t\}$ and $\xi_N^{(c)}$; where dc is the value of the determinant when the algorithm converges and $\xi_N^{(c)}$ is the corresponding optimum design measure when $H = 2$.

The above sequence shall be referred to as S_2 sequence.

(b) For $H = 3$

(1) Select n_{11} support from points g_1, n_{12} support points from g_2 and n_{13} support points from g_3 and define the 3-tuple $\underline{U}_0 = (n_{11} \; n_{12} \; n_{13})$.

(2) Holding n_{13} fixed in g_3, conduct an S_2 search with g_1 and g_2 and obtain

$d_{0(n_{13}+0)} = \min\{d_w\}, w = 1, 2, \, ..., \, a_0$.

where $a_0 = \left(\begin{array}{c} n_1 \\ n_{11} \end{array} \right) \left(\begin{array}{c} n_2 \\ n_{12} \end{array} \right) \left(\begin{array}{c} n_3 \\ n_{13} \end{array} \right),$

$n_1 > n_{11}, n_2 > n_{12}, n_3 > n_{13}$ and $n_{11} + n_{12} + n_{13} = N$; $n_{11}, n_{12}, n_{13} \geq 0$.

if $n_1 < n_{11}, n_2 < n_{12}, n_3 < n_{13}$ and $n_{11} + n_{12} + n_{13} = N$; $n_{11}, n_{12}, n_{13} \geq 0$,

Define $b_{i1} = n_{11} - \alpha n_1$, $b_{i2} = n_{12} - \alpha n_2$, $b_{i3} = n_{13} - \alpha n_3$ where b_{11}, b_{12}, b_{13} are the residues and α is an integer such that $b_{ii} < n_1$, $b_{i2} < n_2$ and $b_{i3} < n_3$ then $a_0 = \begin{pmatrix} n_1 \\ b_{11} \end{pmatrix} \begin{pmatrix} n_2 \\ b_{12} \end{pmatrix} \begin{pmatrix} n_3 \\ b_{13} \end{pmatrix}$.

(3) Set $n_{13} = n_{13} + 1$ and conduct an S_2 search to obtain $d_{(1/n_{13}+1)}$.

(4) Repeat the process in step 3 above for $r = 2, 3, ..., k, k+1, ..., t, t+1$, obtain $d_{(k/n_{13}+k)}$ and $d_{(t/n_{13}+t)}$; where $d_{(k/n_{13}+k)}$ is the minimum determinant value of the alias matrix in increasing order of r, and $d_{(t/n_{13}+t)}$ is the minimum determinant value of the alias matrix in decreasing order of r.

(5) Define $d_c = \min\{d_{(0/n_{13}+0)}, d_{(k/n_{13}+k)}, d_{(t/n_{13}+t)}\}$ and $\xi_N^{(c)}$

where dc is the value of minimum determinant when the sequence converges and $\xi_N^{(c)}$ is the corresponding optimum design measure.

(c) For any H

(1) Select n_{11} support points from g_1, n_{12} support points from g_2, ..., $n_{1(H-1)}$ support points from $g_{1(H-1)}$ and n_{1H} support points from g_H and define the H-tuple $\underline{U}_0 = (n_{11} \ n_{12} \cdots n_{1H})$.

(2) Holding n_{1H} fixed in g_H, conduct an $S_{(H-1)}$ search with the remaining (H-1) balls and obtain

$d_0 = \min\{d_w\}, w = 1, 2, ..., a_0$.

where $a_0 = \begin{pmatrix} n_1 \\ n_{11} \end{pmatrix} \begin{pmatrix} n_2 \\ n_{12} \end{pmatrix} \cdots \begin{pmatrix} n_H \\ n_{1H} \end{pmatrix}$,

$n_1 > n_{11}, n_2 > n_{12}, \cdots, n_H > n_{1H}$ and $n_{11} + n_{12} + ... + n_{1H} = N$; $n_{11}, n_{12}, ..., n_{1H} \geq 0$.

if $n_1 < n_{11}, n_2 < n_{12}, n_3 < n_{13} \cdots, n_H < n_{1H}$ and

$$n_{11} + n_{12} + n_{13} + \cdots + n_{1H} = N; \ n_{11}, n_{12}, n_{13} \cdots, n_{1H} \geq 0,$$

Define $b_{i1} = n_{11} - \alpha n_1$, $b_{i2} = n_{12} - \alpha n_2$, ..., $b_{1H} = n_{1H} - \alpha n_H$ where b_{11}, b_{12}, ..., b_{1H} are the residues and α is an integer such that $b_{ii} < n_1$, $b_{i2} < n_2$, ..., $b_{iH} < n_H$ then $a_0 = \begin{pmatrix} n_1 \\ b_{11} \end{pmatrix} \begin{pmatrix} n_2 \\ b_{12} \end{pmatrix} \cdots \begin{pmatrix} n_H \\ b_{1H} \end{pmatrix}$.

(3) Set $n_{1H} = n_{1H} + 1$ and conduct an $S_{(H-1)}$ search to obtain $d_{(1/n_{1H}+1)}$.

(4) Repeat the process in step 3 above for r = 2, 3, ..., k, k+1, ..., t, t+1, obtain $d_{(k/n_{1H}+k)}$ and $d_{(t/n_{1H}+t)}$; where $d_{(k/n_{1H}+k)}$ is the minimum determinant value of the alias matrix in increasing order of r, and $d_{(t/n_{1H}+t)}$ is the minimum determinant value of the alias matrix in decreasing order of r.

(5) Define $d_c = \min\{d_{(0/n_{1H}+0)}, d_{(k/n_{1H}+k)}, d_{(t/n_{1H}+t)}\}$ and $\xi_N^{(c)}$

where dc is the value of minimum determinant when the sequence converges and $\xi_N^{(c)}$ is the corresponding optimum design measure.

Properties of S_H Sequence

Any search algorithm is made up of a starting point, the direction in which the sequence moves, the step-length in that direction, and where it terminates. Here are the properties of the SH search:

(1) Starting Point: The sequence starts at an arbitrary point $\underline{U}_0 = (n_{11} \ n_{12} \cdots n_{1H})$, $n_{11}, n_{12} > 0$, $n_{11} + n_{12} + ... + n_{1H} = N$, $n_{1i} = 0, 1, ..., N$, $(i = 1)$,

(2) Direction of search: The sequence moves in the direction of decreasing value of the determinant of the alias matrix.

(3) Step length: The step length at which the sequence goes is

$\underline{U}_r = (n_{11} \pm r, \ n_{12} \pm r, ..., n_{1H} \pm r)$, r = 0, 1, ...

This is one step at a time both increasing and decreasing value of $(n_{11} \ n_{12}, ..., n_{1H})$.

(4) The stopping rule: The sequence stops at the minimum value of the determinant of the alias matrix, i.e. where r = t+1

(5) The set $\{\underline{U}_r\}$ is completely exhaustive of all possible steps; where r = 1, 2, ... This means that the total number of steps required for the sequence $\{\underline{U}_r\}$ to converge to minimum will take m = k+t+2 steps. For instance, if a fresh sequence is started at

$\underline{U}_0^* = (n_{11}^* n_{12}^*)$; where $n_{11}^* = n_{11} + s$ and $n_{12}^* = n_{12} - s, s > 0$

then at r = 1, $\underline{U}_1^* = (n_{11} + s + 1 \quad n_{12} - s - 1)$

This implies that it will take k+1 steps to get d_k, where d_k is the minimum determinant value of the alias matrix in the increasing direction of r.

Similarly, it will take t+1 steps to get d_t, where d_t is the minimum determinant value of the alias matrix in the decreasing direction of r. Hence the total number of steps required for the sequence $\{\underline{U}_r^*\}$ to converge at minimum will be m = k+1+s+t-s+1 = k+t+2, i.e., regardless of any starting point, the 2-tuple will traverse in the same direction.

4. Illustration Using Biased First-Order Models with $\tilde{X}$ at Specified Values

REGULAR GEOMETRY: Here interest is in finding an N-point optimal design measure where N will be varied.

Given the response model,

$$f(x_1, x_2) = a_{00} + a_{01}x_1 + a_{02}x_2 + a_{12}x_1x_2 + \{b_1x_1^3 + b_2x_2^3\} + \varepsilon$$

defined in a regular geometric area shown below

In the above model, we have four linear/interactive terms and two biasing terms, so the total number of terms in the model are six. In the experimental region there are nine support points, $\tilde{N} = 9$ and these are arranged into ball according to their distances from the centre. The balls formed are

$$g_1 = \begin{pmatrix} -1 & -1 \\ -1 & 1 \\ 1 & -1 \\ 1 & 1 \end{pmatrix}; g_2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}; g_3 = \begin{pmatrix} 0 & 0 \end{pmatrix}.$$

The combinatorics table is given in table 2 below.

The optimal vector $\underline{U}1 = (4, 2, 0,; 1.296 \times 103)$, the best ball combination is (g_1, g_3) and this is used for the next line of search.

r* in stage one is + 2, the same principle is applied in stage two and is presented below.

Stage Two

Step (r)	Balls			Available designs (ao)	Minimum Determinant Value (d_r)	Local AL Optimum
	g1	g2	g3			
1	4	2	0	6	1.296 x 103	9.934 x 102
2	5	1	0	16	9.934 x 102	
3	3	3	0	16	1.2960e + 003	
4	3	2	1	24	1.2960e + 003	

The optimal vector in stage two is

$\underline{U}_2 = (5, 1, 0; 9.934 \times 10^2)$

The best ball combination is (g_1, g_2) with minimum determinant value of 9.934×10^2 and r* = + 1.

In this search, the sequence has converged in stage two because subsequent analysis brought values higher than 9.934×10^2. The optimal design measure is

$$\xi_6 = \begin{pmatrix} -1 & -1 \\ 1 & -1 \\ -1 & 1 \\ 1 & 1 \\ -1 & 1 \\ -1 & 0 \end{pmatrix}$$

5. Conclusion

A new method for constructing optimal biased designs is proposed. This method gives an optimal combination of balls and an optimal number of support points r* to be taken or added to each ball at each search sequence. This method is certain to converge to an AL design measure for a variety of experimental conditions. Numerical demonstrations indicate a good measure of success in the performance of the algorithm.

References

Allan, J. (1969). *Mathematics for Engineers and Scientists*. Great Britain: Thomas Nelson and Sons Ltd., Chap. 6, 146-198

Allen, T., Yu, L. and Schmitz, J. (2003). An Experimental Design Criterion for minimizing Meta-Model Prediction Errors Applied to die casting process Design. *J. Roy. Statist. Soc. C.*, 52, 103 – 117.

Anderson, V. L. and McLean, R. A. (1974). *Design of Experiments*: A Realistic Approach, Marcel Dekker, Chap. 3, 96-104

Atkinson, A. C. (1982). Development in the Design of Experiments. *International Statistical Review*, 50, 2, 161 – 177.

Atkinson, A. C. (1988). Recent Developments in the Methods of Optimum and related Experimental Designs. *International Statistical Review*, 56, 2, 99 – 115.

Atkinson, A. C. and Donew, A. N. (1992). *Optimum Experimental Design*. Oxford Science Publications, Chap. 9, 334-371

Box, S. E. P. and Draper, N. R. (1959). A Basis for the Selection of a Response Surface Design. *J. Amer. Statist. Ass.*, 54, 622 – 654.

Cook, R. D. and Nachtsheim, C. J. (1980). A Comparison of Algorithms for Constructing Exact D-Optimal Designs. *Technometrics*, 22, 315 – 324.

Ellerton,R.R. and Tsai, W.Y. (1979). Minimum bias estimation and the selection of polynomial terms for response surfaces. *Biometrics*, 35, 631-635 .

Federov, V. V. (1972). *Theory of Optimal Experiments*. Academic Press, New York, Chap. 11, 416-428

Notz, W. T. (1989). Optimal Designs for Regression Models with Possible Bias. *J. Statist. Plann. Inf.*, 22, 43 – 54.

Table 1. Combinatorics for Obtaining Global Minimum When H = 2

Step (r)	g1 (n1)	g2 (n2)	Minimum Determinant (dr)
0	N11	N12	do*
1	N11 + 1	N12 – 1	d1
2	N11 + 2	N12 – 2	d2
⋮	⋮	⋮	⋮
⋮	N11 + k	N12 – k	dk
	N11 + k + 1	N12 – k + 1	dk+1
⋮	N11 – 1	N12 + 1	dk+2
	N11 – 2	N12 +2	dk+3
	⋮	⋮	⋮
	N11 – t	N12 + t	dt
m	N11 – t – 1	N12 + t + 1	dt+1

Table 2. Combinatorics for Selection of 6 Support Points from 3 balls

Step (r)	Balls			Available designs (a0)	Minimum Determinant Value (d_r)	Local AL Optimum
	g1	g2	g3			
0	2	2	2	36	1.2960e + 003	1.296 x 103
1	4	0	2	1	1.2960e + 003	
2	4	2	0	6	1.296 x 103	
3	0	4	2	1	Singular	
4	0	2	4	6	Singular	