D-optimal Exact Design: Weighted VarianceApproach and a Continuous Search Technique

Otaru, O.Paul[1*], Enegesele Dennis[2]

[1]Department of Mathematics and Statistics, University of Port Harcourt, Nigeria
Email:-paultex2005@yahoo.com,otarupaul@gmail.com. Tel:-+2348060563933

[2]Department of Mathematics and Statistics, University of Port Harcourt, Nigeria
Email:-den4_js@yahoo.com Tel:-+2348057708900

ABSTRACT

The aim of this study is to obtain an N-point exact D-optimal design for a feasible region defined on a polynomial model. The weighted variance approach and a continuous search technique were used to obtain a D-optimum measure. The illustrative example buttresses the effectiveness of the method. The minimum variance was obtained in the first iteration and needs no further improvement.

Keywords: Weighted Variance, D-optimum, Gradient Vector and Variance Function

1. INTRODUCTION

In any experimental work it is important to choose the best design in a class of existing designs. However, the choice is solely dependent on the interest of the experimenter and the adequacy of an experimental design can be determined from the information matrix [1]. The most commonly used criteria for choosing experimental design is the D-optimality criteria. D-optimal designs are basically generated by iterative search algorithms, and they seek to minimize the variance and co-variances of the parameter estimates for a specified model [2]. According [3] a D-optimal design is a computer aided design which contains the best subset of all possible experiments, depending on a selected criterion and a given number of design runs. [4]defined that the optimum D-optimal design can be selected within a design region with the aid of a combinatorial algorithm. This combinatorial algorithm requires grouping the support points in the design regions as a measure of their distance from the centre of the design.[5]defined a sequential addition of points to a given initial design in approaching a D-optimum design measure. [5]further explained that a design is D-optimum if the determinant of the information matrix is maximized. Thus, rather than a sequential search from an initial

* Tel.: +2348060563933;
E-mail address: paultex2005@yahoo.com.

design, the search technique reported in this research simply uses a non-sequential approach to obtain an N-point trial D-optimal design from the N-support points and subject the design measure to the weighted variance approach to obtain a D-optimum design.

2. METHODOLOGY

2.1 Weighted Variance Approach

2.1) Let $f(x)$ be an n-variate, p-parameter polynomial of degree m, given by

$$f(x) = \underline{a}\underline{x} + e;$$

$\underline{a}$ is a *p*-component vector of known coefficients, $\underline{x} \in \widetilde{X}$ (feasible region)

2.2) Define the n-component gradient vector,

$$\underline{g} = \{\partial f(x) / \partial x_i\} = \begin{pmatrix} g_1(x) \\ \vdots \\ g_n(x) \end{pmatrix}; g_i(x) = \underline{q}\underline{x} + u;$$

where $g_i(x)$ is an $(m-1)$ degree polynomial, $i = (1, n)$

$\underline{q}$ is r-component vector of known coefficients.

Then compute the gradient vectors, $\{\underline{g}_i\}_{i=(1,n)}$

2.3) From $\widetilde{X}$, define the design measure,

$$\xi_N^0 = \begin{pmatrix} \underline{x}_1 & 1/N \\ \vdots & \\ \underline{x}_N & 1/N \end{pmatrix}; \underline{x}_i = (x_{i1}, \dots, x_{ir});$$

where the N points are spread evenly in $\widetilde{X}$.

2.4) From (2.2) and (2.3) obtain the design matrix

$$X(\xi_N) = \begin{pmatrix} x_{11} \cdots x_{1r} \\ \vdots \\ x_{Ni} \cdots x_{Nr} \end{pmatrix} = \begin{pmatrix} \underline{x}'_1 \\ \vdots \\ \underline{x}'_N \end{pmatrix}$$

and compute, the arithmetic mean vector $\overline{\underline{x}} = (\overline{\underline{x}}_1, \overline{\underline{x}}_2, \cdots, \overline{\underline{x}}_n); \overline{\underline{x}}_i = \sum_{j=1}^{N} x_{ij} / N$

$M^{-1}(\xi_N)$ andthe variances $\{V_i\}_{i=(1,N)}$; $V_i = \underline{x}_i' M^{-1}(\xi_N)\underline{x}_i, M(\xi_N) = X'(\xi_N)X(\xi_N)$

2.5) Define the direction vector,

$$\underline{d} = \sum_{i=1}^{N} \theta_i \underline{g}_i ; \theta_i \in (0,1), \sum_{i=1}^{N} \theta_i = 1$$

and its variance $V(\underline{d}) = \sum_{i=1}^{N} \theta_i^{\,2} V_i$

2.6) Solve for $\{\theta_i\}_{i=(1,N-1)}$ from the partial derivatives

$$\partial V(\underline{d}_A) / \partial \theta_i = 0; \theta_N = 1 - \sum_{i=1}^{N-1} \theta_i$$

and normalize to $\theta_i^{*} = \theta_i \left(\sum_{i=2}^{N} \theta_i^{\,2} \right)^{-\frac{1}{2}} ; \sum_{i=1}^{N} \theta_i^{*2} = 1$

2.7) Define the vector

$\underline{d} = \sum_{i=1}^{N} \theta_i^{*} \underline{g}_i$ andthen normalize $\underline{d}$ to $\underline{d}^{*}; \underline{d}^{*\prime} \underline{d}^{*} = 1$

2.8) Set the starting point,

$\underline{\bar{x}}^{*}$; which corresponds to the mean arithmetic vector

2.9) Compute the step-length,

$$\rho^{*} = \min_{\rho} d(\underline{a}' (\underline{\bar{x}} + \rho \underline{d}^{*})) / d\rho$$

then equate to zero and solve for ρ^{*}

2.10) move to $\underline{x}^{*} = \underline{\bar{x}}^{*} + \rho^{*} \underline{d}^{*}$, and at the jth step, to

$$\underline{x}_j = \underline{\bar{x}}_{j-1} + \rho_{j-1} \underline{d}_{j-1}^{\;*}$$

and compute, $f(\underline{x}_j)$

2.11) Is $\left\| f(\underline{x}_j) - f(\underline{x}_{j-1}) \right\| \leq \delta$?

Yes: Set $\underline{x}_{j-1} = \underline{x}^{*}$, the maximize and stop,

No: Define, $\xi_{N+1}^{\;(0)} = \begin{pmatrix} \xi_N^{\;(0)} \\ \cdots \\ \underline{x}_{j-1} \end{pmatrix}$ and return to (2.3)

2.2 A Continuous Search Technique for an N-Point D-Optimal Exact Design

The continuous search techniquerelies on the search technique developed [6] as follows: The space of possible trials is defined by

$$\widetilde{X} = \{x_i ; a_i \le x_i \le b_i \forall i = 1,2,....,n\}$$

and the sequence of steps required to obtain the design are as follows:

a) **Initial Design:** Assuming f(.) to be a p-parameter m-degree polynomial surface, using a non-sequential method, obtain a non-singular p-point design

$$\xi_p^{(0)} = \begin{Bmatrix} \underline{x}_1, \underline{x}_2, ..., \underline{x}_p \\ w_{1,} w_2, ..., w_p \end{Bmatrix}$$

Such that all the support points in $\xi_p^{(0)}$ fall within the feasible region $\widetilde{X}$.

b) Regression Model of Variance Function

Define a p-parameter polynomial regression function of degree 2m,

$$y(\underline{x}) = b_{00} + \sum_{i=1}^{n} b_{10} x_i + \sum b_n x_i x_j +,.......,+ \sum_{i=m}^{n} \overline{b}_{mn} x_i^{2m}$$

where $y(\underline{x})$ = $d(\underline{x}_p, \xi_p) = \underline{x}'_j M^{-1}(\xi_p)\underline{x}_j ; \underline{x}_j \in \widetilde{X} : j = 1,2,....,\overline{N}, \overline{N} \ge q \ge p$

$$\underline{b} = (b_{00}, b_{10}, ..., b_{n0}, ..., \overline{b}_{mo}, ..., \overline{b}_{mn})$$

The $\overline{N}$ support points are normally inclusive of the initial p-points and are well-spread out as to be representative of $\widetilde{X}$.

c) Trial D-optimal Exact Design

The design ξ_N^0 is achieved if

a. $\sum_{i=j}^{p} x_{ij}^2$ is maximum $\forall j = 1,2,...., p$

b. $\left\| \sum_{i=1} x_{ij} \right\|, \left\| \sum_{i=1} x_{ij} x_{ij} \right\|$, etc . . . are respectively minimized $\forall j, j < j$

Thus an N-point design from the $\overline{N}$ support points and designated as

$$\xi_N^{(0)} = \begin{Bmatrix} \underline{x}_1, \underline{x}_2, ..., \underline{x}_m, ..., \underline{x}_N \\ w_{1,} w_2, ..., w_m, ..., w_N \end{Bmatrix}$$

d) Estimation of Regression Function

By the method of least squares applied to the data in step (b) above, compute the estimates $\underline{\hat{b}}$ and $\underline{\hat{y}} = X\ \underline{\hat{b}}$.

e) Global Maximum of $\underline{\hat{y}}$

Obtain the global maximum $\underline{x}^*$ of $\underline{\hat{y}}$ using the weighted variance approach and compute

$$d(\underline{x}^*, \xi_N^{(0)}) = \underline{x}^{*\prime} M^{-1}(\xi_N^{(0)})\underline{x}^* \text{ and}$$

$$d(\underline{x}_m, \xi_N^{(0)}) = \underline{x}_m{}' M^{-1}(\xi_N^{(0)})\underline{x}_m = \min_x\{\underline{x}' M^{-1}(\xi_N^{(0)})\underline{x}\};\ \underline{x} \in \xi_N^{(0)}$$

f) A Check for Optimality

Is $d(\underline{x}^*, \xi_N^{(0)}) \geq d(\underline{x}_m, \xi_N^{(0)})$?

No: Stop $\xi_N^{(0)}$ is D-optimal

Yes: set $\underline{x}^* = \underline{x}_m$, $w^* = w_m$ in $\xi_N^{(0)}$ and return to step (**e**) above.

3. RESULTS AND DISCUSSION

3.1 Illustrative Example

Obtain a 4-point D-optimal exact design for the response function

$$f(x_1 x_2) = b_0 + b_1 x_1 + b_2 x_2 + \varepsilon$$

Subject to $\widetilde{X} = \{(x_1, x_2) = (-1,1), (-1,-1), (1,-1), (0,0), (1,1), (2,2)\}$

3.2 Solution

a. Initial Design:

Let $\xi_3{}^0 = \begin{bmatrix} 1 & -1 \\ -1 & -1 \\ 0 & 0 \end{bmatrix}$, and the design matrix is thus $X = \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & -1 \\ 1 & 0 & 0 \end{bmatrix}$

b. Regression Model of the Variance Function

The regression function for the variance functions is a quadratic, since m=1

$$y(\underline{x}) = b_{00} + b_{10}x_1 + b_{20}x_1 + b_{12}x_1x_2 + b_{11}x_1^2 + b_{22}x_2^2 + \varepsilon$$

and in addition to the three points in (a) above, y($\underline{x}$) is evaluated at arbitrarily chosen points{(-1,1),(1,3/2),(2,2),(3/2,1),(-1,0),(1,0),(1,1)}, such that the support points chosen are generously spread over $\tilde{X}$.

c. Trial D-optimal Exact Design

Based on the criteria (3c) above, a good trial design is

$$\xi_4^0 = \begin{bmatrix} 2 & 2 \\ -1 & 1 \\ -1 & -1 \\ 1 & -1 \end{bmatrix}$$

d. Estimation of Regression Coefficients

$\hat{\underline{b}} = (X'X)^{-1}X'y = (1,0,2,0,1/2,3/2)$ and

$\hat{y} = 1 + 2x_2 + 1/2x_1^2 + 3/2x_2^2$.

e. Global Maximum of $\hat{y}$

To obtain the global maximum $\underline{x}^*$ of $\underline{\hat{y}}$, we will really on the variance modulated technique (**2**) above. Thus,

$\underline{\bar{x}}^* = \begin{pmatrix} 0.25 \\ 0.25 \end{pmatrix}$ then

$$\underline{d} = \sum_{i=1}^{4} \theta_i^* \underline{g}_i = \begin{pmatrix} 0.1835 \\ 4.4955 \end{pmatrix} \text{and normalize } \underline{d} \text{ to } \underline{d}^* = \begin{pmatrix} 0.0407 \\ 0.9991 \end{pmatrix}; \underline{d}^{*\prime} \underline{d}^* = 1$$

$$\rho^* = \min_{\rho} d(\underline{a}'(\underline{\bar{x}}^* + \rho\underline{d}^*))/d\rho \Rightarrow d\left(\underline{a}'\left(\begin{pmatrix} 0.25 \\ 0.25 \end{pmatrix} + \rho\begin{pmatrix} 0.0407 \\ 0.9991 \end{pmatrix}\right)\right)/d\rho = 0$$

$2.7575 + 2.9962\rho = 0 \Rightarrow \rho^* = -0.9203$

$\underline{x}^* = \underline{\bar{x}}^* + \rho^*\underline{d}^*$

f) A Check for Optimality

The minimum variance in table 1, is

$\underline{x}_m = (1,-1,-1), d(\underline{x}_m, \xi_4^{(0)}) = 0.5789$

while $d(\underline{x}^*, \xi_4^{(0)}) = 0.3954$; therefore

Is $d(\underline{x}^*, \xi_4^{(0)}) \geq d(\underline{x}_m, \xi_4^{(0)})$?

No: Stop $\xi_4^{(0)}$ is D-optimal

Thus, an exact N-point D-optimal design was obtained in the first iteration.

4. CONCLUSION

This continuous search technique has the capacity to obtain an N-point D-optimum exact design of a response function, relying on the weighted variance approach within a feasible region. This technique is very effective for obtaining optimal design in both block and non-block experiments for a feasible region.

AUTHORS' CONTRIBUTIONS

"'Otaru.O.P' designed the study, performed the theoretical analysis and managed the literature searches,'Enegesele D.' managed the illustrative study and also assisted in the theoretical analysis. All authors read and approved the final manuscript."

REFERENCES

1. Atkinson AC, Donev AN. The Construction of Exact D-optimum Experimental Designswith Application to Blocking Response Surface Designs. Biometrika,1989;76(3), 515-526.
2. Eriksson .L.,Johansson .E. , Kettaneh-Wold .N. , Wikstraom.C., Wold. S. Designof Experiments-Principles and Applications. Learnways AB, Umea, 2000;pp 56-98.
3. Iwundu M., Chigbu P. A Hill-Climbing Combinatorial Procedure for Constructing D-optimal Designs. J.Stat. Appl. 2012; Pro 1, No 2,133-146.
4. Wynn HP. The Sequential Generation of D-optimum Experimental Designs. AnnualsMathematical Statistics, 1970; Vol 41, No. 5, 1655-1664.
5. Onukogu IB, Chigbu PE. Super Convergent Line Series (In Optimal Design of Experiments and Mathematical Programming) 2nd ed; AP Express Publishing Company:Nsukka; 2002.

APPENDIX

Table 1: The table showing design points, gradient vector, variance and weighting factor at *N*= 4 design points in the feasible region.

Serial	Design Points (x_{1i}, x_{2i})	Gradient Vector $\underline{g}_i = (g_{1i}, g_{2i})$	Variance (V_i)	Weighting Factor (θ_i^*)
1	2 , 2	2 , 8	0.8947	0.4034
2	-1 , 1	-1 , 5	0.7631	0.4731
3	-1 , -1	-1 , -1	0.5789	0.6237
4	1 , -1	1 , -1	0.7631	0.4735