Original Research Article

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

By

Abstract

The convergence to a D-optimum measure is possibly obtained using a continuous search technique while relying on the weighted variance approach. This search technique compares favourably well with other known sequential search technique.

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

1 Introduction

The weighted variance approach that obtains the optimum of a response function defined within a feasible region follows a sequential method. The aim of this research is to show that the continuous search technique Onukogu et al. (2002) relying on the weighted variance approach obtains the D-optimum exact design for a feasible region.

In the literature, the best D-optimal design can be selected within a design region with the aid of a combinatorial algorithm Iwundu et al. (2012). However, Wynn (1970) defined a sequential addition of points to a given initial design in approaching a D-optimum design measure. Wynn 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 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 $\overline{N}$ support points and subject the design measure to the weighted variance approach to obtain a D-optimum design for a response function.

2 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,

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

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 $\underline{\bar{x}}_A = (\underline{\bar{x}}_1, \underline{\bar{x}}_2, \cdots, \underline{\bar{x}}_n); \underline{\bar{x}}_i = \sum_{j=1}^{N} x_{ij} / N$

$M^{-1}(\xi_N)$ and the 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,

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

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

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

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

2.8) Set the starting point,

$\underline{\bar{x}} = \underline{\bar{x}}_A$; which corresponds to the mean arithmetic vector

2.9) Compute the step-length,

$$\rho = \rho_A = \min_{\rho} d(\underline{a}' (\underline{\bar{x}}_A + \rho \underline{d}_A^{*})) / d\rho = 0$$

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

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

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

2.11) Is $\left\| f(\underline{x}_{Aj}) - f(\underline{x}_{Aj-1}) \right\| \le \delta$?

Yes: Set $\underline{x}_{Aj-1} = \underline{x}^{*}$, 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)

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

We shall rely on the search technique developed Onukogu et al. (2002) 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

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

Such that all the support points in ξ_p 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 $\xi_N{}^0$ 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

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

d) Estimation of Regression Function

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

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

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

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

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

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.

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

Solution

a. Initial Design:

Let $\xi_3 = \begin{bmatrix} 1 & -1 \\ -1 & -1 \\ 0 & 0 \end{bmatrix}$, 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 $\widetilde{X}$.

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,

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

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

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

$$2.7575 + 2.9962\,\rho_A = 0 \Rightarrow \rho_A = -0.9203$$

$$\underline{x}_A = \underline{\bar{x}}_A + \rho_A \underline{d}_A^{\,*}$$

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

Yes: set $\underline{x}^* = \underline{x}_m$, $w^* = w_m$ in $\xi_4^{(0)}$ and return to step (**e**) above.

5 Conclusion

A continuous search technique has been shown to have the capacity to obtain 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.

6 Reference

Atkinson, A.C and Donev, A.N (1989): The Construction of Exact D-optimum Experimental Designs with Application to Blocking Response Surface Designs. *Biometrika, 76(3), 515-526.*

Flecher, R. (1981): Practical Methods of Optimization: Vol.2, Constrained Optimization; *Practical Hall Inc.,* U.S.A.

Iwundu, M. and Chibu, P. (2012): A Hill-Climbing Combinatorial Procedure for Constructing D-optimal Designs. *J.Stat. Appl. Pro 1, No 2,133-146.*

Onukogu, I.B. and Chigbu, P.E (2002): Super Convergent Line Series (In Optimal Design of Experiments and Mathematical Programming) *AP Express Publishing Company,* Nsukka.

Wynn, H.P (1970): The Sequential Generation of D-optimum Experimental Designs. *Annuals Mathematical Statistics, Vol 41, No. 5, 1655-1664.*

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