D and G Optimality and Efficiency Concerning Response Surface Designs

Shaun S. Wulff
May 7, 1994

Masters Writing Project
Montana State Univerisity
Mathematics Department
supervised by Dr. John J. Borkowski

Experimental designs define an arrangement for producing data in which a treatment or treatments are applied. Criteria have been constructed to compare experimental designs and to find *best* designs among a class of designs used for an experiment. Optimal design theory is a methodology that finds these *best* designs. Even if an applicable optimal design cannot be found, the optimality criteria can provide a measure of the adequacy of any design under study.

The experimental design and response model can be characterized by matrix notation. A design matrix X displays k factors $(x_1, x_2, ..., x_k)$ for n experimental runs. This matrix is given by

$$X_{n\times k} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1k} \\ x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \\ x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix}$$

where the i^{th} row $\underline{x}_i$=$[x_{i1}, x_{i2}, ..., x_{ik}]$ represents one experimental run. [19] The set of values the factor can assume define an experimental design region $\mathcal{X}$. For convenience, the factors are centered and scaled. The centered and scaled j^{th} factor in the i^{th} experimental run is $z_{ij} = \frac{(x_{ij}-\bar{x}_i)}{\frac{R}{2}}$ where R= range of the x_i and $\bar{x}$= mean of the x_i $(i = 1, 2, ..., k)$. When the factors are centered and scaled in this manner, the following experimental design regions can be defined:

- Hypercube in $\mathcal{R}^k$: $\mathbf{H_k}$= $\{(x_1, x_2, ..., x_k) \mid |x_i| \leq 1\}$
- Hypersphere in $\mathcal{R}^k$(of radius r): $\mathbf{S_{k,r}} = \{(x_1, x_2, ..., x_k) \mid \sum_{i=1}^{k} x_i^2 \leq r^2\}$.

For many designs, the centered and scaled points can be identified as barycenters. A barycenter of depth ω for $1 \leq \omega \leq k$ is a point with ω coordinates equal to 0 and $k - \omega$ coordinates equal to ± 1. The set of all barycenters of depth ω is denoted $J_k(\omega)$. The set of factorial points is given by $J_k(0)$ with each point of the form $(x_1, x_2, ..., x_k) = (\pm 1, \pm 1, ..., \pm 1)$

and the set of center points is given by $J_k(k)$ with each point of the form $(x_1, x_2, ..., x_k) = (0, 0, ..., 0)$. [10] Representing an experimental design as a design matrix X, defining the experimental design region $\mathcal{X}$, and considering sets of barycenters of depth ω $J_k(\omega)$ will be useful when studying the design criteria.

The response model for the experimental design is given by $\underline{Y}_{n\times1}=X_{n\times k}\underline{\beta}_{k\times1} + \underline{\epsilon}_{n\times1}$ where $\underline{Y}$ is a vector of responses, X is the design matrix $\ni X \subset \mathcal{X}$, $\underline{\beta}$ is a vector of unknown parameters, and $\underline{\epsilon}$ is a random vector of errors $\ni \underline{\epsilon} \sim (0, \sigma^2 I)$. [18] The experimental design problem consists of finding estimates of the unknown model parameters. It will be assumed that the design matrix is full column rank ($r(X) = k$). By the Gauss-Markov Theorem, the best linear unbiased estimator of $\underline{\beta}$ is the least-squares estimator $\underline{b}$. [18] This gives the fitted model $\hat{\underline{Y}}=X\underline{b}$. It is desirable to aquire estimates which have small variance, and which minimize the variance of the predicted values $\hat{\underline{Y}}$. The variance/covariance matrix is given by $\text{Var}(\underline{b})=\sigma^2(X'X)^{-1}$, the prediction variance is given by $\text{Var}(\hat{\underline{Y}})=\sigma^2 X(X'X)^{-1}X'$, and the standardized prediction variance is given by $d(X)=\frac{n}{\sigma^2}\text{Var}(\hat{\underline{Y}})$. [1] Criteria will be defined that focus on minimizing $\text{Var}(\underline{b})$ and $d(X)$.

A variety of criteria have been defined for choosing a design. The choice of criterion will affect the type of design that is determined to be optimal. The criteria developed in this analysis are related to the $X'X$ matrix which characterizes $\text{Var}(\underline{b})$ and $\text{Var}(\hat{\underline{Y}})$. Supposing $\underline{b}$ is the vector of least-square parameter estimates, then the $100(1-\alpha)\%$ confidence ellipsoid for $\underline{\beta}$ is given by $(\underline{\beta}-\underline{b})'(X'X)(\underline{\beta}-\underline{b}) = K_\alpha$ where K_α is a constant that depends on α. [18] Box and Draper consider D,A,and E optimality in terms of the eigenvalues $(\lambda_1, \lambda_2, ..., \lambda_k)$ of $(X'X)^{-1}$. These criteria are:

- $D = |X'X| = \prod_{i=1}^{k} \lambda_i^{-1}$ which measures the volume of the confidence ellipsoid.
- $A = trace(X'X)^{-1} = \sum_{i=1}^{k} \lambda_i$ which measures the nonsphericity of the confidence ellipsoid.
- $E = \max_{1 \le i \le k} \lambda_i$ which measures the nonsphericity of the confidence ellipsoid.

The G-criterion deals with the maximum prediction variance over the experimental region $\mathcal{X}$ where G= $\max_{x \in \mathcal{X}}$ Var($\hat{\underline{Y}}$). [6] In this study, the D and G criteria are of primary interest and will be further developed.

In general, optimum designs do not have equal weighting of points in the design. The weighting scheme for the points form a probability measure ξ on the design space $\mathcal{X}$. The design space is a class of subsets of S where S is an arbitrary set of points. The design space is a **field** if

1) $S \in \mathcal{X}$

2) $A \in \mathcal{X} \Rightarrow A^c \in \mathcal{X} \quad \forall$ classes $A \ni A \in S$.

3) $A, B \in \mathcal{X} \Rightarrow A \cup B \in \mathcal{X} \quad \forall$ classes A and $B \ni A, B \in S$.

A **set function** is a real-valued function defined on the field $\mathcal{X}$. A set function ξ can be defined as a **probability measure** on $\mathcal{X}$ if

1) $0 \le \xi(A) \le 1 \quad \forall A \in \mathcal{X}$

2) $\xi(\emptyset) = 0, \ \xi(S) = 1$

3) If $A_1, A_2, ...$ is a disjoint sequence of sets in $\mathcal{X}$ and if $\bigcup_{i=1}^{\infty} A_i \in \mathcal{X}$, then

$$\xi(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \xi(A_i).$$

If $\mathcal{X}$ is a field and ξ is a probability measure, then $\mathcal{X}$ is a **support** for ξ if $\xi(\mathcal{X}) = 1$. [2]

A continuous measure assigns probabilities to a continuum of points whereas the discrete measure assigns probabilities to a countable number of points. For example, consider a discrete measure ξ. If A is the class of points $x_1, x_2, ..., x_n$, then the probability measure for A is given by $\xi(A) = \sum_{i=1}^{n} \omega_i$ where $A \in \mathcal{X}$. [2] Continuous and discrete measures will both be of interest for chosen supports. However, an applicable experimental design requires a finite set of points.

The discrete measure ξ can be expressed as

$$\xi = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ \omega_1 & \omega_2 & \dots & \omega_n \end{pmatrix}$$

where $x_1, x_2, ..., x_n$ are distinct points in $\mathcal{X}$ with associated weights $\omega_1, \omega_2, ..., \omega_n$. Since ξ is a discrete probability measure $\int_{\mathcal{X}} \xi(dx) = \sum_{i=1}^{n} \omega_i = 1$ and $0 \leq \omega_i \leq 1 \ \forall\ i$. [1] A **continuous design** is one in which $\omega_i \in [0,1] \ \ \forall\ i$. An **exact design** is one in which $\omega_i \in [0,1] \cap \mathcal{Q} \ \forall\ i$. For instance, an exact design for n trials on t distinct points can be expressed as

$$\xi_n = \begin{pmatrix} x_1 & x_2 & \dots & x_t \\ \frac{r_1}{n} & \frac{r_2}{n} & \dots & \frac{r_t}{n} \end{pmatrix}$$

where $r_i \in \mathcal{Z}$ for $i = 1, 2, ..., n$ and $\sum_{t=1}^{n} r_t = n$. [1] For an exact design, the proportion of points at x_i can be collected, but for a continuous design the proportion of points at x_i may be impossible to collect. For example, it is not possible to collect a proportion of points at x_i if the associated weight is $\omega_i = \frac{\sqrt{2}}{n}$. Such a value for ω_i is possible for a continuous design, but not for an exact design. Continuous designs are useful since they make the calculation of optimum designs possible. If one or more $\omega_i's$ are obtained such that $\omega_i \notin [0,1] \cap \mathcal{Q}$ then $\gamma_i \in [0,1] \cap \mathcal{Q}$ is selected which approximates ω_i in the continuous design. This approximation yields an exact design, for which observations can be collected at the specified proportions of γ_i.

The moment matrix $M(\xi)$ for a design ξ is a generalization of the matrix $X'X$. The $X'X$ matrix assumes each point in the design has equal weight, while the $M(\xi)$ matrix does not assume an equal weighting scheme. The matrix $M(\xi)$ has elements $m_{ij}(\xi) \ni m_{ij}(\xi) = \int_{\mathcal{X}} f_i(x) f_j(x) \xi(dx)$ where $f_i(x)$ is a function of the design points. Let f represent a known $k \times 1$ vector of continuous functions defined on ξ. Let $f(x)$ denote the column k-vector with components $f_i(x)$ with the q-variable vector $(x_1, x_2, ..., x_q)$. The normalized prediction variance is given by $d(x,\xi) = (f(x))' M^{-1}(\xi)(f(x))$. [11] The moment matrix allows the optimality criteria to be generalized to unequal weighting of the design points which may be necessary to obtain an optimum experimental design.

Using the design measure ξ and the moment matrix $M(\xi)$, an optimal design ξ^* can be defined in terms of the D and G criteria. A D-optimum design ξ^* maximizes $|M(\xi)|$ or, equivalently, minimizes $|M^{-1}(\xi)|$ for nonsingular $M(\xi)$. Thus, the D-optimum design minimizes the volume of the confidence ellipsoid for the parameter vector $\underline{\beta}$. For any arbitrary design ξ, a D-efficiency is defined by $D_{eff} = \left(\frac{|M(\xi)|}{|M(\xi^*)|}\right)^{\frac{1}{q}}$. The D-efficiency measures how efficient any point in ξ is relative to any point in ξ^*. [1] Secondly, a design ξ^* is G-optimal if $\min_{\xi} \max_{x \in \mathcal{X}} d(x,\xi) = \max_{x \in \mathcal{X}} d(x,\xi^*)$ where $d(x,\xi)$ is the standardized prediction variance with respect to the design measure ξ. Hence, a G-optimal design will minimize the maximum prediction variance. For any arbitrary design ξ, the G-efficiency is defined by $G_{eff} = \frac{d(x,\xi^*)}{d(x,\xi)}$. [1] Therefore, D and G optimality have been defined from the criteria using the design measure ξ and the moment matrix $M(\xi)$.

The Keifer-Wolfowitz Theorem demonstrates the equivalency of D and G optimal designs. For nonsingular $M(\xi)$, the Kiefer-Wolfowitz Theorem states conditions i), ii), iii) are equivalent:

i) ξ^* is D-optimal

ii) ξ^* is G-optimal

iii) $\max_{x \in \mathcal{X}} d(x, \xi^*) = k$

where k=number of parameters in the model. [13]

The results of the Kiefer-Wolfowitz theorem can be used to establish a procedure for finding an optimal design. This procedure can be outlined as follows:

1) Numerically find weights to assign sets of points that will maximize $|M(\xi)|$.

2) Construct a design (ξ') with these weights.

3) Show that the chosen weights for (ξ') satisfy $\max_{x \in \mathcal{X}} d(x, \xi') = k$.

By the Keifer-Wolfowitz Theorem, ξ' is G-optimal by condition (ii) and ξ' is D-optimal by condition (i). The difficulty in this procedure is the determination of the weights that maximize $|M(\xi)|$. A general method for finding these weights will be considered for the design regions of the hypercube and hypersphere in $\mathcal{R}^q$.

For the hypercube $\mathbf{H_q}$, the design weights can be determined by considering a support consisting of subsets of barycentric points. From the Kiefer-Wolfowitz Theorem, $\max_{x \in \mathcal{X}} d(x, \xi^*) = k$ for an optimal design ξ^* where k is the number of parameters in the model. The maximum value of the standardized prediction variance must occur at some subset of points in the design region which can be represented by barycentric points. Farrell, Kiefer, and Walbran [9] establish this with the following theorem.

Theorem. The set of barycentric points $\mathcal{J}(j_1, j_2, j_3) = \bigcup_{i=1}^{3} J_q(j_i)$ support a symmetric optimum design for quadratic regression on $\mathbf{H_q}$ iff

- for $2 \leq q \leq 5$: $j_1 = 0$, $j_2 = 1$, $2 \leq j_3 \leq q$
- for $q \geq 6$: $j_1 = 0$, j_2= 1 or 2 , $3 \leq j_3 \leq q$.

Among the sets specified by the theorem, the smallest number of points is $1 + q2^{q-1} + 2^q$ which is given by the set $\mathcal{J}(0,1,q)$. [9] The requirement of a symmetric optimum design can be satisfied by finding a group of transformations G for which the class of design measures is G-invariant $\Rightarrow \xi(gA) = \xi(A)\ \forall g \in G,\ A \in \mathcal{X}$. This transformation can be used to express the function $d(x, \xi^*)$ as a symmetric polynomial in terms of $x_1^2, x_2^2, ..., x_k^2$. [9]

Thus, it is necessary to find three weights ω_1, ω_2, ω_3 to assign to the sets of barycentric points $J_q(q), J_q(1), J_q(0)$ where w_i is constant for all points in $J_q(i)$. For a D-optimal design, these three weights need to maximize $|M(\xi)|$. Assuming a symmetric design, consider $F(u,v) = |M(\xi)|$ where

$$u(\omega) = \int x_1^2 \xi(dx) = \int x_1^4 \xi(dx) = 1 - \sum_{i=0}^{q} \frac{i\omega_i}{q}$$

$$v(\omega) = \int x_1^2 x_2^2 \xi(dx) = \sum_{i=0}^{q} \frac{\omega_i(q-i)(q-i-1)}{q(q-1)}$$

where x_i=0,1,-1. [10] The values of u, v need to be solved for numerically using the process:

1. consider $log|F(u,v)| = L$
2. find $\frac{\partial L}{\partial u}$, $\frac{\partial L}{\partial v}$,
3. set $\frac{\partial L}{\partial u} = 0$, $\frac{\partial L}{\partial v} = 0$
4. solve the system of 2 nonlinear equations defined in (3)
5. verify a maximum is obtained.

The values of $u(\omega)$ and $v(\omega)$ are substituted into $M(\omega)$ which has the general form

$$M(\omega) = \begin{bmatrix} 1 & h & 0 & 0 \\ h' & H & 0 & 0 \\ 0 & 0 & uI_q & 0 \\ 0 & 0 & 0 & vI_{\frac{q(q-1)}{2}} \end{bmatrix}$$

where

- I_r is the identity matrix of order r

- h is the row q-vector with all entries u

- H has diagonal entries u and off-diagonal entries v, ie. $H = (u-v)I + vJJ'$. [10]

Using $M(\omega)$, $\max_{x \in \mathcal{X}} d(x, \xi^*)$ is calculated. If $\max_{x \in \mathcal{X}} d(x, \xi^*) = k$, then an optimal design has been obtained as indicated by the Kiefer-Wolfowitz Theorem.

For the hypercube $\mathbf{H_q}$ using a second-order polynomial in q factors, the optimal weights $\omega_1, \omega_2, \omega_3$ assigned to the sets of barycentric points $J_q(q)$, $J_q(1)$, $J_q(0)$ are given in the following table. [1]

Design Weights			
q	ω_1	ω_2	ω_3
2	0.096	0.321	0.583
3	0.066	0.424	0.510
4	0.047	0.502	0.451
5	0.036	0.562	0.402

Thus, each factorial point in $J_q(0)$ has weight $\frac{\omega_3}{2^q}$, each facecenter point in $J_q(1)$ has weight $\frac{\omega_2}{q2^q}$, and weight ω_1 is given to the center points in $J_q(0)$.

In order to find the optimal weights for ξ^* on the unit hypersphere $\mathbf{S_{q,1}}$ a continuous measure ξ must be examined. Reasonable approximations to this design ξ will be considered which are implementable and near optimal. For a model of degree 2 (m=2), weight $\alpha = \frac{2}{(q+1)(q+2)}$ is given to the center point, and weight $\beta = 1 - \alpha$ is given to the boundary of $\mathcal{X}$. [9] For m=3, the optimal weighting scheme considers two parameters ρ=radius of the sphere ($\rho \leq 1$) and β= measure given to the points spread uniformly over the sphere. Using these parameters, the optimum weights β and $1-\beta$ can be found by maximizing $\log|M(\xi)|$ where

$$\log|M(\xi)| = C_q + 2q\log\rho + (q+1)\log[\beta(1-\beta)(1-\rho^2)^2] + \tfrac{(q+2)(q-1)}{2}\log[(1-\beta)+\beta\rho^4] + \tfrac{q(q+4)(q-1)}{6}\log[(1-\beta)+\beta\rho^6]$$

and C_q is a constant depending on q. [9] A maximum may be found by simultaneously solving $\frac{\partial \log|M(\xi)|}{\partial\beta} = 0$ and $\frac{\partial \log|M(\xi)|}{\partial\rho} = 0$ for the parameters β and ρ. For $m \geq 4$, the same approach can be used, but it is more difficult. [9]

Near optimal discrete measures ξ can be defined on the factorial and star points on the unit-sphere. By the Kiefer-Wolfowitz Theorem, the maximization of the standardized variance must occur at points in the design region. This discrete design region can be represented by factorial points $(\pm\frac{1}{\sqrt{q}}, \pm\frac{1}{\sqrt{q}}, ..., \pm\frac{1}{\sqrt{q}})$ and star points $(0, ..., \pm 1, ..., 0)$. Every factorial point has weight $\frac{\alpha}{2^q}$ and every star point has weight $\frac{\beta}{2q}$. Thus, the total weight for the factorial points is α and the total weight for the star points is β where $\alpha+\beta=1$. [9] Center points can be added by the experimenter with the design weights adjusted accordingly. The moments for $M(\xi)$ are given by

$$\int x_1^2\xi(dx) = \frac{1}{q}, \quad \int x_1^4\xi(dx) = \frac{3}{q(q+2)}, \quad \int x_1^2x_2^2\xi(dx) = \frac{1}{q(q+2)}$$

where all other odd moments less than four are zero. The weights $\alpha = \frac{q}{q+2}$ and $\beta = \frac{1}{q(q+2)}$ satisfy the above equations and meet the constraint $\alpha+\beta=1$. [9] The number of points for this design is $2q$ (star points)+ n_q (factorial points).

Optimality will be considered for two classes of designs which are commonly used: the Central Composite designs and the Box-Behnken designs. Efficiencies based on a quadratic model with k terms and q scaled variables $-1 \leq x \leq 1$ on the hypersphere or hypercube will be analyzed. This analysis will demonstrate how close these designs are to the optimal design. All points in these designs have equal weight, so the design ξ can be represented by

$$\xi = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ \frac{1}{n} & \frac{1}{n} & \dots & \frac{1}{n} \end{pmatrix}.$$

The Central Composite design consists of points from two level factorial or fractional factorial designs, star points, and center points. The standard composite design for three factors ($q = 3$) with one center point has the design matrix

$$X_{15\times 3} = \begin{bmatrix} +1 & +1 & +1 \\ +1 & +1 & -1 \\ +1 & -1 & +1 \\ +1 & -1 & -1 \\ -1 & +1 & +1 \\ -1 & +1 & -1 \\ -1 & -1 & +1 \\ -1 & -1 & -1 \\ +\alpha & 0 & 0 \\ -\alpha & 0 & 0 \\ 0 & +\alpha & 0 \\ 0 & -\alpha & 0 \\ 0 & 0 & +\alpha \\ 0 & 0 & -\alpha \\ 0 & 0 & 0 \end{bmatrix} \quad [8].$$

Different composite designs can be constructed by varying the number of center points, changing the star point distance (α), using different fractional factorial designs, and replicating design points. Among implementable designs, there is a design that maximizes $|X'X|$. This best implementable Central Composite design will be compared to both the optimal design and the optimal Central Composite design for a specified number of factors.

For $\mathbf{H_q}$, the star point distance is 1 ($\alpha = 1$). The best implementable Central Composite designs are quite close to the D-optimum design. For $q = 3$, the 14 point Central Composite design (no center point) has a D-efficiency of .976 and a G-efficiency of .893. For $q = 4$, the 24 point Central Composite design (no center point) has D-efficiency of .936 and G-efficiency of .811. [17] As the number of center points increase, the G-efficiencies decrease. [3] For

$q \geq 5$, fractional factorial points can be used to reduce the number of points in the design without reducing the efficiencies. However, the D and G efficiencies decrease as q increases. [17] (see Table 1)

For $\mathbf{S_{q,1}}$, the star point distance is $\alpha = \sqrt{q}$. For $q = 2$, the D-efficiencies are .9862, .9964, .9691 and the G-efficiencies are .6667, .9600, .8727 with 1,2, and 3 center points respectively. For $q = 3$, the D-efficiencies are .9914, .9961, .9763 and the G-efficiencies are .6667, .9459, .8903 with 1,2, and 3 center points respectively. [17] Using two center points often resulted in large increases in G-efficiencies since this proportion of center points is closer to the weight given to the center of the design region in the optimal design. The D-efficiencies are quite close to one for all q, while the G-efficiences vary widely depending on the number of center points. [17] (see Table 2)

Implementable Central Composite Designs can also be compared to optimal Central Composite designs. This comparison will be restricted to $\mathbf{H_q}$, because optimal Central Composite Designs on $\mathbf{S_{q,1}}$ lack practical significance, particularly for $k \geq 4$. [16] Optimal weights $\omega_1, \omega_2, \omega_3$ assigned to the barycentric points $J_q(q)$, $J_q(1)$, $J_q(0)$ can be found by numerically solving

$$|X'X| = \left(\omega_3 + \frac{\omega_2}{q}\right)^q \omega_3^{\binom{q}{2}} \left(\frac{\omega_2}{q}\right)^{q-1} \left(\left(q\omega_3 + \frac{\omega_2}{q}\right) - q\left(\omega_3 + \frac{\omega_2}{q}\right)^2\right) n$$

subject to the constraints $\omega_i \geq 0$ $i = 1, 2, 3$ and $\sum_{i=1}^{3} \omega_i = 1$ where n is the number of points in the design. The solutions to this equation are given in the following table up to 5 factors. [16]

Design Weights CCD			
q	ω_1	ω_2	ω_3
2	0.096	0.321	0.583
3	0	0.345	0.655
4	0	0.292	0.708
5	0	0.253	0.747

The D-efficiences between the best implementable and the optimal Central Composite designs are quite similiar. (see Tables 1 and 3) Although the D-efficiencies of the optimal Central Composite designs decrease as q increases, these efficiencies remain relatively high. This indicates that optimal Central Composite designs and implementable Central Composite designs are near optimal on $\mathbf{H_q}$. [16]

The Box-Behnken design consists of barycentric points. For $q = 3$, the barycentric points are $J_q(1)$ and $J_q(q)$. The standard Box-Behnken design for three factors and one center point has the design matrix

$$X_{13\times 3} = \begin{bmatrix} +1 & +1 & 0 \\ +1 & -1 & 0 \\ -1 & +1 & 0 \\ -1 & -1 & 0 \\ +1 & 0 & +1 \\ +1 & 0 & -1 \\ -1 & 0 & +1 \\ -1 & 0 & -1 \\ 0 & +1 & +1 \\ 0 & +1 & -1 \\ 0 & -1 & +1 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \quad [5].$$

Different Box-Behnken designs are constructed by using barycentric points of different depths for larger q, varying the number of center points, and replicating the design points. Among these designs, there is one design that maximizes $|X'X|$ among all implementable Box-Behnken designs. This best implementable Box-Behnken design will be compared to the optimal design for a specified number of factors.

For $\mathbf{H_q}$, the Box-Behnken design points do not fill in the extremes of the design region since it does not contain the set of factorial points ($J_q(0)$). These points lie on the corners of the design region. Thus, on the hypercube, the D and G efficiencies of Box-Behnken designs would not be as close to optimal as those for the Central Composite designs. Therefore, these designs should not be used in practice for this region.

For $\mathbf{S_{q,1}}$, two center points are included yielding the best Box-Behnken design. However, the addition of center points does not significantly lower D and G efficiencies. For $q = 3$, the fourteen point Box-Behnken design has a D-efficiency of .9653 and a G-efficiency of .7143. For $q = 4$, the 26 point Box-Behnken has a D-efficiency of .9992 and a G-efficiency of .9890. [17] The efficiences are quite close to one for most values of q, but the G-efficiencies do vary for some values of q. [17] (see Table 2)

The Central Composite and Box-Behnken designs are quite efficient. If the design region is the hypercube, the efficiencies decrease as q increases, but if the design region is the hypersphere, the efficiencies remain relatively high. Thus, these designs are good designs to use in practice according to the D and G optimality criteria. [17] However, variations of these designs or other designs may be of interest to experimenters. These experimenters may value other design criteria such as orthogonality, rotatability, replication, lack of fit, and internal estimates of error. Whatever design is chosen, the D and G optimality criteria can be used to assess how close the chosen design comes to minimizing the volume of the confidence ellipsoid for the parameter estimates and minimizing the maximum prediction variance for the specified model.

TABLE 2—*Design Comparison on a Hypercube*

No. of Factors	No. of Terms	No. of Points	Design	$M(\xi)$	D-eff	G-eff	d_{max}/n
3	10	--	D-Optimum	$.578 \times 10^{-3}$	100.0	100.0	--
		14	Best Composite	$.453 \times 10^{-3}$	97.6	89.3	.80
		13	Hoke D_6	$.432 \times 10^{-3}$	97.1	65.1	1.18
		10	Hoke D_2	$.104 \times 10^{-3}$	84.3	14.3	7.00
		10	Box-Draper	$.185 \times 10^{-3}$	89.2	54.3	1.84
		11	2^3 + 3 Star Pts.	$.323 \times 10^{-3}$	94.4	60.6	1.50
4	15	--	D-Optimum	$.216 \times 10^{-4}$	100.0	100.0	--
		24	Best Composite	$.804 \times 10^{-5}$	93.6	81.1	.77
		42	Pesotchinsky DP	$.13 \times 10^{-4}$	96.7	82.4	.43
		34	Pesotchinsky DB	$.44 \times 10^{-5}$	89.9	67.6	.65
		19	Hoke D_6	$.170 \times 10^{-5}$	84.4	31.9	2.48
		15	Hoke D_2	$.795 \times 10^{-6}$	80.2	39.1	2.56
		15	Box-Draper	$.389 \times 10^{-6}$	76.5		
5	21	--	D-Optimum	$.635 \times 10^{-6}$	100.0	100.0	--
		42	Best Composite	$.684 \times 10^{-7}$	89.9	61.4	.81
		26	Best Composite	$.328 \times 10^{-7}$	86.8	77.8	1.04
		50	Pesotchinsky DP	$.40 \times 10^{-6}$	97.8	87.1	.48
		84	Pesotchinsky DB	$.18 \times 10^{-6}$	94.2	77.8	.32
		26	Hoke D_6	$.702 \times 10^{-7}$	90.0	71.7	1.13
		21	Hoke D_2	$.505 \times 10^{-7}$	88.7	52.5	1.91
		21	Box-Draper	$.339 \times 10^{-10}$	62.6		
6	28	--	D-Optimum	$.154 \times 10^{-7}$	100.0	100.0	--
		44	Best Composite	$.288 \times 10^{-9}$	86.7	63.9	1.00
		88	Nalimov	$.21 \times 10^{-8}$	93.1	65.7*	.48
		66	Pesotchinsky DP	$.14 \times 10^{-7}$	99.7	94.9	.45
		100	Pesotchinsky DB	$.54 \times 10^{-8}$	96.3	83.8	.33
		34	Hoke D_6	$.583 \times 10^{-10}$	81.9	59.2	1.39
		28	Hoke D_2	$.473 \times 10^{-10}$	81.3	46.0	2.17
		28	Box-Draper	$.932 \times 10^{-16}$	50.9		
7	36	--	D-Optimum	$.317 \times 10^{-9}$			
		78	Best Composite	$.105 \times 10^{-11}$	85.3	44.7*	1.03
		113	Pesotchinsky DB	$.13 \times 10^{-9}$	97.6	87.6	.36
		43	Hoke D_6	$.122 \times 10^{-14}$	70.7	46.9	1.79
		36	Hoke D_2	$.104 \times 10^{-14}$	70.4	39.0	2.56
		36	Box-Draper	$.648 \times 10^{-23}$	41.7		
8	45	--	D-Optimum	$.568 \times 10^{-11}$	100.0	100.0	--
		80	Best Composite	$.248 \times 10^{-14}$	84.2	47.4	1.19
		53	Hoke D_6	$.593 \times 10^{-21}$	60.0	36.8	2.31
		45	Hoke D_2	$.503 \times 10^{-21}$	59.8	38.0	2.63
		45	Box-Draper	$.922 \times 10^{-32}$	34.5*		
9	55	--	D-Optimum	$.900 \times 10^{-13}$	100.0	100.0	--
		146	Best Composite	$.294 \times 10^{-17}$	82.9	44.1*	.85
		64	Hoke D_6	$.531 \times 10^{-29}$	50.7*	28.9	2.98
		55	Hoke D_2	$.463 \times 10^{-29}$	50.6	29.4	3.40
		55	Box-Draper	$.219 \times 10^{-42}$	28.9*		
10	66	--	D-Optimum	$.128 \times 10^{-14}$	100.0		
		148	Best Composite	$.566 \times 10^{-20}$	83.0		
		76	Hoke D_6	$.780 \times 10^{-39}$	43.0*		
		66	Hoke D_1	$.683 \times 10^{-39}$	42.9		
		66	Box-Draper	$.732 \times 10^{-55}$	24.6*		
11	78	--	D-Optimum	$.164 \times 10^{-16}$	100.0		
		151	Best Composite	$.681 \times 10^{-23}$	82.8		
		89	Hoke D_6	$.154 \times 10^{-50}$	36.6*		
		78	Hoke D_1	$.136 \times 10^{-50}$	36.6*		
		78	Box-Draper	$.279 \times 10^{-69}$	21.1*		

*Indicates a design that did not achieve an efficiency as high as the lower bound given in formula 2.

Table 2

TABLE 1—*Design Comparison on A k-sphere*

No. of Factors	No. of Terms	No. of Points	Design	M(ξ)	D-eff	G-eff	d_{max}/n	Number of Center Points
2	6	--	D-optimum	2.62×10^{-4}	100.00	100.00	-	
		7	Uniform shell	2.58×10^{-4}	99.78	85.71	1.00	1
		8		2.31×10^{-4}	98.00	90.00	.83	2
		9		1.71×10^{-4}	93.20	80.00	.83	3
		9	Composite	2.41×10^{-4}	98.62	66.67	1.00	1
		10		2.56×10^{-4}	99.64	96.00	.62	2
		11		2.17×10^{-4}	96.91	87.27	.62	3
3	10	--	D-optimum	2.52×10^{-9}	100.00	100.00	-	-
		13	Uniform shell	1.86×10^{-9}	97.00	76.92	1.00	1
		14	≡ Box-Behnken	1.77×10^{-9}	96.53	71.43	1.00	2
		15		1.33×10^{-9}	93.82	66.67	1.00	3
		15	Composite	2.31×10^{-9}	99.14	66.67	1.00	1
		16		2.42×10^{-9}	99.61	94.59	.66	2
		17		1.98×10^{-9}	97.63	89.03	.66	3
4	15	--	D-optimum	7.50×10^{-17}	100.00	100.00	-	-
		25	+Composite	6.68×10^{-17}	99.23	60.00	1.00	1
		26	Box-Behnken	7.42×10^{-17}	99.92	98.90	.58	2
		27		6.32×10^{-17}	98.86	95.24	.58	3
		21	Uniform shell	2.80×10^{-17}	93.64	71.43	1.00	1
		22		2.79×10^{-17}	93.61	75.76	.90	2
		23		2.15×10^{-17}	91.99	72.46	.90	3
5	21		D-optimum	4.44×10^{-27}	100.00	100.00	-	-
		43	Composite	3.30×10^{-27}	98.60	48.84*	1.00	1
		44		4.08×10^{-27}	99.60	87.63	.54	2
		45		3.82×10^{-27}	99.28	85.96	.54	3
		27	Composite (Half Rep)	3.19×10^{-27}	98.43	77.78	1.00	1
		28		2.97×10^{-27}	98.10	87.64	.86	2
		29		2.13×10^{-27}	96.57	84.62	.86	3
		31	Uniform shell	5.32×10^{-28}	90.39	67.74	1.00	1
		32		5.46×10^{-28}	90.50	67.02	.98	2
		33		4.29×10^{-28}	89.47	64.99	.98	3
		41	Box-Behnken	2.87×10^{-27}	97.95	51.22	1.00	1
		42		3.47×10^{-27}	98.83	90.91	.55	2
		43		3.17×10^{-27}	98.41	88.79	.55	3
		46		1.54×10^{-27}	95.08	83.00	.55	6
6	28		D-optimum	3.63×10^{-40}	100.00	100.00	-	-
		45	Composite (Half Rep)	3.20×10^{-40}	99.55	62.22	1.00	1
		46		3.46×10^{-40}	99.83	96.95	.63	2
		47		2.84×10^{-40}	99.13	94.89	.63	3
		43	Uniform shell	8.46×10^{-42}	87.44	65.12	1.00	1
		44		8.89×10^{-42}	87.59	65.18	.98	2
		45		7.11×10^{-42}	86.90	63.73	.98	3
		49	Box-Behnken	6.77×10^{-41}	94.18	57.14	1.00	1
		50		7.69×10^{-41}	94.61	67.20	.83	2
		51		6.62×10^{-41}	94.11	65.88	.83	3
		54		2.67×10^{-41}	91.11	62.22	.83	6
7	36		D-optimum	3.01×10^{-56}	100.00	100.00	-	-
		79	Composite (Half Rep)	2.01×10^{-56}	98.88	45.57*	1.00	1
		80		2.56×10^{-56}	99.55	84.72	.53	2
		81		2.45×10^{-56}	99.43	83.68	.53	3
		95 †		1.89×10^{-56}	98.72	91.29	.42	3
		57	Uniform shell	8.05×10^{-59}	84.83	63.16	1.00	1
		58		8.61×10^{-59}	84.99	62.07	1.00	2
		59		6.98×10^{-59}	84.49	61.02	1.00	3
		57	Box-Behnken	2.74×10^{-56}	99.74	63.16	1.00	1
		58		2.93×10^{-56}	99.93	99.31	.62	2
		59		2.38×10^{-56}	99.35	97.63	.62	3
		62		7.98×10^{-57}	96.38	92.90	.62	6
8	45		D-optimum	1.94×10^{-75}	100.00	100.00	-	-

- make design robust to worst possible case

- don't have optimal designs for subsets of models

- robust designs

optimality of efficiency

optimum design theory

- D efficiency upper bound for G efficiency

- at an optimum design D efficiency = G efficiency

- the more center points added, the lower the D+G criterion

- might want to give up some D efficiency for gains in lack of fit

Table 3

Optimum Composite Designs

alias structures such as $I = ABC$ and $I = A$. $|X'X|/n^{t} = .21 \times 10$ and $a_{max} = 40$, while out

TABLE 2—*Design Comparison*

Factors k	Number of Coefficients t	Number of* Observations n	Design	One Center Point** $\lvert X'X\rvert/n^t$	D-eff.	G-eff.	No Center Points** $\lvert X'X\rvert/n^t$	D-eff.	G-eff.
2	6	-	D-Optimum	$.114 \times 10^{-1}$					
		-	Optimum Composite	$.114 \times 10^{-1}$	1.000				
		9	Composite p = 0	$.975 \times 10^{-2}$	.974	.628	$.879 \times 10^{-2}$	.957	.600
		7	Composite p = 1 (I = A)	$.163 \times 10^{-2}$	.723	.202	$.137 \times 10^{-2}$	.702	.200
		6	Box-Draper best 6-Point Design	$.574 \times 10^{-2}$	.892	.536			
		7	Box-Draper best 7-Point Design	$.634 \times 10^{-2}$	.949	.612	—		
		8	Box-Draper best 8-Point Design	$.901 \times 10^{-2}$	.962	.665			
3	10	-	D-Optimum	$.578 \times 10^{-3}$					
		-	Optimum Composite	$.511 \times 10^{-3}$	.988				
		15	Composite p = 0	$.320 \times 10^{-3}$	.942	.836	$.453 \times 10^{-3}$	.976	.693
		11	Composite p = 1 (I = ABC)	$.363 \times 10^{-5}$	.602	.130	$.655 \times 10^{-5}$	.639	.143
		10	Box-Draper 10-Point Design	$.185 \times 10^{-3}$	.892	.543			
4	15	-	D-Optimum	$.216 \times 10^{-4}$					
		-	Optimum Composite	$.848 \times 10^{-5}$	.940				
		25	Composite p = 0	$.536 \times 10^{-5}$	.911	.780	$.804 \times 10^{-5}$	.936	.611
		17	Composite p = 1 (I = ABC)***	$.104 \times 10^{-7}$	.601	.127	$.210 \times 10^{-7}$	.630	.134
5	21	-	D-Optimum	$.635 \times 10^{-6}$					
		-	Optimum Composite	$.692 \times 10^{-7}$	.900				
		43	Composite p = 0	$.483 \times 10^{-7}$	.885	.602	$.684 \times 10^{-7}$	.899	.614
		27	Composite p = 1 (I = ABCDE)	$.173 \times 10^{-7}$	.842	.749	$.328 \times 10^{-7}$	.868	.778
6	28	-	D-Optimum	$.154 \times 10^{-7}$					
		-	Optimum Composite	$.339 \times 10^{-9}$	.873				
		77	Composite p = 0	$.177 \times 10^{-9}$	.853	.417	$.228 \times 10^{-9}$	.860	.421
		45	Composite p = 1 (I = ABCDEF)	$.172 \times 10^{-9}$	.852	.625	$.288 \times 10^{-9}$	.867	.639
		29	Composite p = 2 (I = ABC = DEF)	$.240 \times 10^{-16}$	.485	.074	$.570 \times 10^{-16}$	.500	.077
7	36	-	D-Optimum	$.317 \times 10^{-9}$					
		-	Optimum Composite	$.111 \times 10^{-11}$	.855				
		143	Composite p = 0	$.185 \times 10^{-12}$	.813	.264	$.217 \times 10^{-12}$	.818	.266
		79	Composite p = 1 (I = ABCDEFG)	$.728 \times 10^{-12}$	.845	.442	$.105 \times 10^{-11}$	.853	.447
		47	Composite p = 2 (I = ABCDE = AFG)	$.448 \times 10^{-16}$	.645	.108	$.885 \times 10^{-16}$	.658	.110

*For composite designs, the n-value for a design with one center point is given.

**Number of center points refers to finite composite designs only.

***I=ABCD is not used as this alias structure confounds some two-factor interactions, so all terms in a second-order model are not estimable.

BIBLIOGRAPHY

[1] Atkinson, A.C. and Donev, A.N. Optimum Experimental Designs. Clarendon Press, Oxford, 1992.

[2] Billingsley, Patrick. Probability and Measure. John Wiley and Sons, New York, 1986.

[3] Borkowski, John J. Finding Maximum G-criterion Values for Central Composite Designs on the Hypercube.

[4] Borkowski, John J. The Evaluation of Mixed Resolution Designs. PhD Thesis, University of Delaware, Jan. 1992.

[5] Box, G.E.P. and Behnken, D.W. Some new three-level desgns for the study of quantitative variables. Technometrics, 2:455-475, 1960.

[6] Box, G.E.P. and Draper, N.R. Empirical Model Building and Response Surfaces. John Wiley and Sons,New York, 1987.

[7] Box, G.E.P. and Draper, N.R. Factorial designs, the $|X'X|$ criterion, and some related matters. Technometrics, 13:731-743, 1987.

[8] Box, G.E.P. and Wilson, K.B. On the experimental attainment of optimum conditions. Journal of the Royal Statistical Society, B13:1-45, 1951.

[9] Farell, R.H., Kiefer, J. and Walbran, A. Optimum Multivariate designs. The Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1:113-138, 1967.

[10] Galil, Z. and Kiefer, J. Comparison of design for quadratic regression on cubes. Journal of Statistical Planning and Inference, 1:121-131, 1977.

[11] Kiefer, J. Optimum designs in regression problems, II. Annals of Mathematical Statistics, 32:298-325, 1961.

[12] Kiefer, J. Optimum experimental designs. Journal of the Royal Statistical Society, B21:272-304, 1959.

[13] Kiefer, J. and Wolfowitz, J. The equivalence of two extremum problems. Canadian Journal of Mathematics, 12:363-366, 1960.

[14] Kiefer, J. and Wolfowitz, J. Optimum designs in regression problems. Annals of Mathematical Statistics, 30:271-294, 1959.

[15] Lucas, James M. Design efficiencies for varying number of centre points. Biometrika, 64:145-147, 1977.

[16] Lucas, James M. Optimum composite designs. Technometrics, 16:561-567, 1974.

[17] Lucas, James M. Which respone surface is best. Technometrics, 18:411-417, 1976.

[18] Myers, Raymond H. and Milton, Janet S. A First Course in the Theory of Linear Statistical Models, PWS-KENT Publishing Co., Boston, 1991.

[19] Myers, Raymond H. Response Surface Methodology. Allyn and Bacon, Inc., Boston, 1971.

[20] Wadrop, Daniel M. and Myers, Raymond H. Some response surface designs for finding optimal conditions. Journal of Statistical Planning and Inference, 25:7-28, 1990.