TECHNOMETRICS ©, VOL. 26, NO. 3, AUGUST 1984

Computer-Aided Design of Experiments for Response Estimation

William J. Welch

Faculty of Commerce and Business Administration
The University of British Columbia
Vancouver, British Columbia V6T 1Y8
Canada

In experiments for response estimation, algorithms developed for constructing D-optimal exact (integer replication) designs may be inappropriate when the number of observations is not large relative to the number of parameters. This article generalizes Mitchell's DETMAX algorithm to an arbitrary design criterion and describes an efficient implementation for minimizing either the maximum variance or the average variance of the response estimator over the design region. Some examples are given that demonstrate the possible improvements in response estimator variances.

KEY WORDS: Computer search; DETMAX; Mixture experiment; Optimal design; Response estimation; Second-order design.

1. INTRODUCTION

Although much attention has been given to the construction of D-optimal exact (integer replication) experiments, criteria concerned with the variance of the response estimator, such as G optimality, have received scant consideration in the literature. This is inappropriate because, in practice, many experiments are aimed at estimation of the response over the region of interest rather than parameter estimation. Moreover, D and G optimality are not necessarily equivalent for exact designs.

For example, the experiments motivating this research often involve the relationship between the yield of uranium leached from an ore and explanatory variables such as grind size, reaction time, reagent concentrations, and temperature. The conditions maximizing the yield are usually not the most economic, and the response should be estimated over the entire region of interest. Furthermore, observations tend to be expensive, necessitating small experiments.

In general, suppose the experimenter wishes to estimate a response y depending on a vector of explanatory variables x over a region X. In defining X we may differentiate between two types of explanatory variables: those that may be easily set anywhere in a specified range and those restricted to a finite set of values. The second type includes not only qualitative factors but also quantitative variables limited to a few levels for experimental convenience. Both kinds may occur in an experiment. Mitchell (1974) outlines various reasons for restricting attention to explanatory variables with a finite number of levels, and this is the approach adopted here. Continuous variables may still be approximated by a fine spacing of levels, though computational considerations limit the number of such factors. The adequacy of the approximation is illustrated in Section 3. Hence, X comprises r candidate points $x_1, \ldots, x_r$, with the vector x_j representing a combination of levels of the explanatory variables.

The design problem is to choose n not necessarily distinct points $x_{(1)}, \ldots, x_{(n)}$ from X at which to take observations $Y_{(1)}, \ldots, Y_{(n)}$ of the response according to an assumed linear model:

$$Y_{(i)} = f^T(x_{(i)})\,\beta + \varepsilon_i, \qquad i = 1, \ldots, n.$$

The $k \times 1$ vectors f and β are k specified linearly independent functions on X and the associated k unknown parameters ($k \leq n$). The random errors $\varepsilon_1, \ldots, \varepsilon_n$ are assumed to be uncorrelated, each with mean zero and constant variance σ^2. Typically, for response estimation the model will be a first- or second-order polynomial.

It is convenient to represent an n-point design with n_j replications at x_j in X by an $r \times 1$ vector p, where $p_j = n_j/n$ is the proportion of the total observations at x_j ($j = 1, \ldots, r$). The D-optimality criterion chooses p to maximize the determinant of the information matrix

$$M(p) = \sum_{j=1}^{r} p_j f(x_j) f^T(x_j)$$

or, equivalently, to minimize the determinant of $M^{-1}(p)$, the covariance matrix of the least squares estimators $\hat{\beta}$ normalized for n and σ^2. Throughout we shall assume that $M(p)$ is invertible; singularity is easily avoided by the adjustment of Mitchell (1974).

The D-optimality criterion, however, is inappropriate if we are primarily interested in the response estimators

$$\hat{y}(x_j) = f^T(x_j)\hat{\beta}, \qquad j = 1, \ldots, r.$$

For a design p the normalized variance of $\hat{y}(x_j)$ is

$$d(x_j, p) = f^T(x_j)M^{-1}(p)f(x_j), \qquad j = 1, \ldots, r,$$

and we might choose p to minimize the maximum variance,

$$d_{\max}(p) = \max_{j=1,\ldots,r} d(x_j, p), \tag{1}$$

or to minimize the average variance,

$$d_{\text{ave}}(p) = r^{-1} \sum_{j=1}^{r} d(x_j, p). \tag{2}$$

A method of computing the maximum or average variance over a grid that approximates one or more continuous variables, even if the design is restricted to a few convenient levels, is illustrated in Section 3. For a weighted average, only minor changes need be made. Criterion (1) is commonly called G optimality and (2) will be called V optimality, borrowing the terminology V from Box and Draper (1959). The mean squared error approach of Box and Draper is not pursued here, however, since the number of observations is assumed to be small and the importance of variance relative to bias is assumed to be large (see Kiefer 1975 and Galil and Kiefer 1977). Even when parameter estimation is the primary objective, preferred designs perform well when judged by further criteria such as $d_{\max}$ and d_{ave} in addition to D optimality.

The approximate design theory of Kiefer and Wolfowitz (1960), in which the design is not constrained to have integer replications, leads to equivalence of D and G optimality. Similarly, Fedorov (1972, Theorem 2.12.1) gives conditions for equivalence between V optimality (Fedorov's Q) and either D optimality or minimizing the trace of $M^{-1}(p)$. If n is not large relative to k, however, there is no such equivalence over the set of exact designs in general. Furthermore, the constraint of integer replication leads to considerable computational difficulties. For D optimality, the DETMAX algorithm of Mitchell (1974), with computational improvements by Galil and Kiefer (1980), provides designs at least close to optimality. The branch-and-bound algorithm of Welch (1982) guarantees, at greater cost, a catalog of all designs with $\det\{M(p)\}$ within a specified small tolerance of the optimal design, which can then be compared using other criteria of interest. The primary design criterion, however, is still D optimality. Computer time would also be prohibitively expensive for most of the examples attempted here.

In Section 2, Mitchell's DETMAX algorithm is adapted to find G- or V-optimal designs. The two series of examples presented in Section 3 illustrate that the G and V criteria may produce designs performing well by a range of criteria compared with the D-optimal design, without excessive execution time.

2. ALGORITHMS FOR G AND V OPTIMALITY

2.1 Excursions

The DETMAX algorithm can be generalized to minimize an arbitrary design criterion $c(p)$ over the set of exact designs subject to

$$l_j \le p_j \le u_j, \qquad j = 1, \ldots, r, \tag{3}$$

where l_j and u_j are specified constants constraining the replication at x_j. For D optimality $c(p) = \det\{M^{-1}(p)\}$ and DETMAX is recovered. In an unconstrained-replication problem,

$$l_j = 0, u_j = 1, \qquad j = 1, \ldots, r.$$

For design augmentation, though, where n_j observations have already been taken at x_j, the best n-point design subject to inclusion of the previous points is found by setting

$$l_j = n_j/n, u_j = 1, \qquad j = 1, \ldots, r.$$

Alternatively, suppose we wish to minimize $d_{\max}$ in (1) or d_{ave} in (2), evaluated over a fine grid of points in X, but restrict the design to a subset of X representing convenient values of the explanatory variables. We simply assign $u_j = 1$ for the allowed candidates and $u_j = 0$ otherwise. The augmentation and restriction facilities are illustrated in Section 3.

Like DETMAX, we attempt to improve an initial n-point design by a series of excursions. An excursion starts by adding or subtracting a point from the current n-point design $p^{(1)}$ and then performs a number of additions or subtractions of a single point, eventually returning to a possibly new n-point design $p^{(2)}$. If $c(p^{(2)}) < c(p^{(1)})$, then the excursion has succeeded and $p^{(2)}$ is used as the start for the next excursion, whereas if $c(p^{(2)}) \ge c(p^{(1)})$, a failure has occurred and $p^{(1)}$ is again the starting design. At each step within an excursion, two decisions are made: whether to add or subtract a point and which design point to add or subtract accordingly. The rules for the first decision follow DETMAX and are described by Mitchell (1974) and Galil and Kiefer (1980).

The algorithm to minimize $c(p)$ deviates from DETMAX in the choice of a promising point to add or subtract as required. If the adjustment involves candidate x_j, then element p_j of the current design is

increased or decreased by n^{-1}. The obvious generalization of DETMAX is to add a point x_{j_0}, satisfying

$$c(p + n^{-1}\delta_{j_0}) = \min_{j:p_j<u_j} c(p + n^{-1}\delta_j), \tag{4}$$

or when subtracting, select x_{j_0} to satisfy

$$c(p - n^{-1}\delta_{j_0}) = \min_{j:p_j>l_j} c(p - n^{-1}\delta_j), \tag{5}$$

where δ_j is the $r \times 1$ vector with one in position j and zero elsewhere. In (4) and (5) we normalize for n observations, even though the number of design points changes during an excursion.

In applying this to the G and V optimality criteria, we need to define the normalized covariance between $\hat{y}(x_i)$ and $\hat{y}(x_j)$,

$$d(x_i, x_j, p) = f^T(x_i)M^{-1}(p)f(x_j), \tag{6}$$

and use the result (Fedorov 1972, Lemma 3.3.1),

$$M^{-1}(p \pm n^{-1}\delta_j) = M^{-1}(p) - \frac{M^{-1}(p)f(x_j)f^T(x_j)M^{-1}(p)}{d(x_j, p) \pm n}. \tag{7}$$

From (6) and (7), the normalized variance of $\hat{y}(x_i)$, after a change at x_j, is

$$d(x_i, p \pm n^{-1}\delta_j) = d(x_i, p) - \frac{d^2(x_i, x_j, p)}{d(x_j, p) \pm n}. \tag{8}$$

Accordingly, for G optimality we would compute the minimum in (4) or (5) as appropriate using

$$c(p \pm n^{-1}\delta_j) = d_{\max}(p \pm n^{-1}\delta_j) = \max_{i=1,\ldots,r} \left\{ d(x_i, p) - \frac{d^2(x_i, x_j, p)}{d(x_j, p) \pm n} \right\}, \tag{9}$$

whereas for V optimality,

$$c(p \pm n^{-1}\delta_j) = c(p) - \frac{r^{-1}\sum_{i=1}^{r} d^2(x_i, x_j, p)}{d(x_j, p) \pm n}. \tag{10}$$

2.2 Some Computational Considerations

For both G and V optimality, it is necessary to maintain $M^{-1}(p)$ and $d(x_j, p)$ $(j = 1, \ldots, r)$. These can be efficiently updated when p changes as described by Galil and Kiefer (1980).

The G optimality computation of $d(x_i, x_j, p)$ $(i = 1, \ldots, r)$ in (9) is similarly facilitated by writing

$$d(x_i, x_j, p) = f^T(x_i)\{M^{-1}(p)f(x_j)\}$$

and computing the vector $M^{-1}(p)f(x_j)$ only once. If $r \gg k$, the maximization in (9) therefore requires approximately rk operations of one multiplication and one addition. When adding a design point, if (9) is calculated for every candidate x_j, a total of r^2k operations will be required, followed by a further rk operations to update $d(x_i, p)$ $(i = 1, \ldots, r)$ for the next step. This compares unfavorably with D optimality, where the only major computation is the updating process.

The calculation of (9), however, can be reduced further. When adding a point x_j, clearly

$$d(x_i, p) - \frac{d^2(x_i, x_j, p)}{d(x_j, p) + n} \le d(x_i, p). \tag{11}$$

Similarly, when subtracting a point x_j, the inequality

$$d(x_i, p) - \frac{d^2(x_i, x_j, p)}{d(x_j, p) - n} \le \frac{d(x_i, p)}{1 - n^{-1}d(x_j, p)} \tag{12}$$

follows from (a) the relationship

$$d^2(x_i, x_j, p) \le d(x_i, p)d(x_j, p)$$

between a squared covariance and variances and (b) the result $d(x_j, p) \le n$ (Atwood 1973, Theorem 1). The bounds supplied by the right sides of (11) and (12) are denoted by $b(x_i, p \pm n^{-1}\delta_j)$. They are exploited to find the new maximum variance in (9) without necessarily searching through the entire design region.

First, we need to sort X into the order $x_{\langle 1\rangle}, \ldots, x_{\langle r\rangle}$, such that

$$d(x_{\langle 1\rangle}, p) \ge \cdots \ge d(x_{\langle r\rangle}, p).$$

The $O(r \log r)$ sorting time (e.g., see Knuth 1973, Ch. 5) is trivial compared with up to r^2k operations without bounding. Next, for each candidate x_j, where the replication constraints (3) would allow $p \pm n^{-1}\delta_j$, we find the new maximum variance in (9) in four steps.

1. Set MAX = 0. MAX will ultimately hold the value of $d_{\max}(p \pm n^{-1}\delta_j)$. Now perform Steps 2, 3, and 4 for $i = 1, \ldots, r$.
2. For $x_{\langle i\rangle}$, compute the bound $b(x_{\langle i\rangle}, p \pm n^{-1}\delta_j)$ from (11) or (12).
3. If $b(x_{\langle i\rangle}, p \pm n^{-1}\delta_j) \le$ MAX, then $x_{\langle i\rangle}$ cannot increase the current value of MAX. Moreover, as the variances $d(x_{\langle i\rangle}, p)$ and hence the bounds $b(x_{\langle i\rangle}, p \pm n^{-1}\delta_j)$ are nonincreasing in i, neither can any $x_{\langle i'\rangle}$ for $i' > i$, and we terminate with the current MAX. If, however, $b(x_{\langle i\rangle}, p \pm n^{-1}\delta_j) >$ MAX, then continue to Step 4.
4. Compute $d(x_{\langle i\rangle}, p \pm n^{-1}\delta_j)$ from (8) and increase MAX if necessary. Continue at Step 2 for the next i.

The best candidate x_j has the minimum MAX value. Further refinements are possible: At Step 4, for example, if MAX becomes larger than $d_{\max}(p \pm n^{-1}\delta_{j'})$ for a candidate $x_{j'}$ already evaluated, then further consideration of x_j is unnecessary.

Even with this enhancement, G optimality remains expensive. The results can also be disappointing: Paradoxically the D- or V-optimal designs often possess smaller values of $d_{\max}$ than those generated by the

G-optimality algorithm intended to minimize $d_{\max}$. The G criterion is particularly prone to becoming trapped at a (poor) local optimum.

Fortunately, though, a simple modification produces substantially smaller values of $d_{\max}$ on the average and simultaneously reduces execution time yet further. Within an excursion, minimizing $d_{\max}(p \pm n^{-1}\delta_j)$ seeks the best change one step ahead—an optimal strategy for the final step. At an earlier stage, however, suppose that $x_{(1)}, \ldots, x_{(m)}$ would be the best set of $m > 1$ points to collectively introduce. It appears that selecting a single additional point to minimize $d_{\max}(p + n^{-1}\delta_j)$ can lead to a poor compromise not included in $x_{(1)}, \ldots, x_{(m)}$. The modification retains the criterion $c(p) = d_{\max}(p)$ to determine whether a G-optimality excursion has been successful and in selecting the final point (where the preceding computational improvements are still relevant). The criterion $d_{\max}$ is temporarily ignored at all other steps of an excursion. At these steps, follow the much cheaper DETMAX strategy and either add or subtract a point at the candidate where the current design achieves either maximum or minimum $d(x_j, p)$, subject to (3). A wide range of empirical studies have clearly demonstrated the superior performance of this variant, and it is adopted for the examples of Section 3, where empirical execution times are also reported.

Turning to V optimality, an efficient computational procedure has already been described by Welch (1983), where details may be found. Briefly, in (10) one may write

$$r^{-1}\sum_{i=1}^{r} d^2(x_i, x_j, p) = f^T(x_j)A(p)f(x_j) = a(x_j, p), \quad (13)$$

where $A(p)$ is the $k \times k$ matrix

$$M^{-1}(p)\{r^{-1} \Sigma\, f(x_i)f^T(x_i)\}M^{-1}(p).$$

Hence, if the quadratic forms $a(x_j, p)$ are maintained, an addition or subtraction step is straightforward. Efficient updating of $a(x_j, p)$ $(j = 1, \ldots, r)$ as p changes also requires maintenance of $A(p)$. Because two sets of quadratic forms—$d(x_j, p)$ and $a(x_j, p)$—are updated, with similar computational complexities, a V-optimality step is about twice as expensive as D optimality.

2.3 Random Initial Designs

Like DETMAX, the generalized algorithm does not guarantee a design minimizing $c(p)$ as it may lead to a local optimum. Following Mitchell (1974), repeated attempts are made from a number of random initial designs, abandoning an attempt when the excursions are forced to $n + 6$ or $n - 6$ points. To generate the starting design, Galil and Kiefer (1980) advocated sampling n_{ran} random candidates, with n_{ran} itself randomly distributed between 1 and $k/2$. By sequentially adding $n - n_{\text{ran}}$ further points according to (4), the initial n-point design is produced. This strategy is adopted for all of the examples in Section 3.

Galil and Kiefer also suggest that experimentation with larger n_{ran} can be profitable in some (probably uncommon) cases. We have an option to allow n_{ran} between user-defined limits a and b. Selecting these limits to allow $n_{\text{ran}} > n$ and then subtracting $n_{\text{ran}} - n$ points does not appear to have been investigated. For instance, consider the first-order model with six explanatory variables,

$$E(Y_{(i)}) = \beta_0 + \beta_1 x_{(i)1} + \cdots + \beta_6 x_{(i)6},$$

where $x_{(i)} = (x_{(i)1}, \ldots, x_{(i)6})$. Twelve observations are to be taken from the design region consisting of the 64 candidate vectors $x = (\pm 1, \ldots, \pm 1)$. Galil and Kiefer reported difficulty in obtaining the D-optimal design with DETMAX for this and related cases, only succeeding in 14 out of 200 attempts when $(a, b) = (1, 6)$. With (a, b) taking the values (1, 3), (4, 12), (13, 64), and (65, 256), we achieved D-optimality success frequencies from 50 attempts of 5, 9, 11, and 16, respectively, clearly suggestive of a trend with n_{ran}.

Since the D-, G-, and V-optimal designs coincide for this example, sharing the orthogonal Hadamard structure, experimentation with the criterion may also be helpful. The G and V criteria attained optimality in 8 out of 100 and 85 out of 100 attempts, respectively. Both of these rates are aggregated over the four pairs of (a, b), as varying (a, b) has little effect. It is not clear why the V algorithm is so successful in locating a Hadamard structure here.

3. EXAMPLES

Experiments with mixtures have received much attention in the literature; Cornell (1981) provides a review. They involve explanatory variables with nonnegative levels summing to one to represent the proportions of components in a mixture. Vuchkov, Damgaliev, and Yontchev (1981) (VDY) describe the sequential generation of D-optimal designs for mixture experiments that also include independent process variables.

The first series of examples applies the generalized excursion algorithm to the case of three mixture variables and one process variable. A design point $x_{(i)} = (x_{(i)1}, \ldots, x_{(i)4})$ is therefore constrained such that

$$x_{(i)1} + x_{(i)2} + x_{(i)3} = 1;$$
$$x_{(i)s} \geq 0, \qquad s = 1, \ldots, 3. \quad (14)$$

The second-order model canonical parameterization given by VDY is

$$E(Y_{(i)}) = \sum_{s=1}^{3} \beta_s x_{(i)s} + \sum_{s=1}^{3} \sum_{t=s+1}^{4} \beta_{st} x_{(i)s} x_{(i)t} + \beta_{44} x_{(i)4}^2.$$

This model has $k = 10$ parameters. As a design region

suppose that the mixture proportions are continuous between zero and one, and approximate their ranges by the seven levels $x_{(i)s} = 0(1/6)1$ $(s = 1, \ldots, 3)$. To ascertain whether seven levels are sufficient, 13-level grids spaced at intervals of 1/12 are also used. In contrast, assume that the process variable is confined to only three values—coded -1, 0, and 1—for any of the reasons described in Section 1. Clearly if this variable could be continuously set with ease, then again, a finer grid of values would be appropriate. With three seven-level mixture components and a three-level process variable, there are 84 combinations of levels satisfying the constraints (14) to comprise the design region. The number of candidates increases to 273 if the mixtures are at 13 levels.

The excursion algorithm was executed 10 times for $n = 10, \ldots, 15$ observations, with the D, G, and V criteria and both the 84-point and 273-point grids. During a G or V excursion, $d_{\max}$ or d_{ave} is evaluated over the chosen search grid. To facilitate comparison, however, the $d_{\max}$ and d_{ave} values presented for the final designs are always computed over the finer 273-point grid. We also report $\{\det(M^{-1})\}^{1/k}$, the quantity minimized by D optimality; the exponent $1/k$ reduces the determinant to units of variance for consistency with the other criteria. Another property of possible interest is $\lambda_{\max}$, the maximum eigenvalue of $M^{-1}(p)$, also interpretable as the maximum variance of any linear combination of the parameter estimators $a^T\hat{\beta}$ subject to $a^Ta = 1$. Minimizing $\lambda_{\max}$ is commonly called E optimality.

A set of 10 attempts often yields a variety of designs. Table 1 lists the properties of selected designs performing well for one, or preferably several, of the tabulated criteria. When no design simultaneously minimizes all criteria, the choice is subjective. Thus when $n = 13$, two designs found by employing G optimality on a 273-point grid are listed. The first has the smaller $d_{\max}$ and $\lambda_{\max}$, whereas the second has better d_{ave} and $\{\det(M^{-1})\}^{1/k}$.

The designs in Table 1 demonstrate relatively little variability in $\{\det(M^{-1})\}^{1/k}$. Hence for these examples, D optimality appears to exploit negligible improvements in $\{\det(M^{-1})\}^{1/k}$. Conversely, the remaining criteria show greater diversity. Most striking is the case $n = 11$, where $d_{\max}$ is 18.2 for D optimality and 12.8 for G and V optimality, with an even larger proportional improvement for $\lambda_{\max}$. Except for $n = 10$, G optimality produces designs with values of $d_{\max}$ and d_{ave} substantially lower than those of D optimality. Therefore, the G optimal designs with asterisks have been subjectively selected as a compromise among the various criteria. For four out of the six values of n, these designs exploit the finer 273-point grid, but there would be little practical disadvantage in restricting the choice to the 84-point design region.

Table 1. *Design Properties for the Example With Three Mixture Variables and One Process Variable*

	Criterion					
n	$r = 84$	$r = 273$	d_{max}	d_{ave}	$\{det\ (M^{-1})\}^{1/k}$	λ_{max}
10		G	17.8	9.9	14.6	390*
	D, V, G	D, V	18.3	10.2	14.3	389
		V	23.3	9.6	15.4	391
11	V, G	V, G	12.8	8.2	14.3	213*
	D	D	18.2	9.1	14.2	340
12		G	13.5	8.2	14.4	234*
		G	13.8	8.1	14.3	227
	G	V	14.0	8.2	14.4	229
	V		15.3	8.1	14.3	250
	D	D	15.5	8.7	13.8	344
13		G	13.0	7.9	14.3	219*
	V		13.1	7.7	14.2	225
		G	13.1	7.7	14.1	225
	G		13.3	7.6	13.8	237
		V	13.9	7.7	13.9	244
	D	D	15.2	8.4	13.8	236
14	G		13.1	7.6	13.9	209*
		G	13.1	7.7	14.2	251
		V	13.6	7.4	14.2	203
	V		14.4	7.5	13.9	247
		D	16.0	8.0	13.7	231
	D		16.4	8.6	13.7	206
15		G	13.4	7.4	13.9	210*
	G		13.4	7.5	14.0	205
		V	14.7	7.3	14.2	225
		D	14.7	7.6	13.6	220
	V	V	15.0	7.2	13.6	258
	D		15.0	8.9	13.6	241

* Compromise among the various criteria.
NOTE: Search grids of size $r = 84$ and $r = 273$ are employed, but d_{max} and d_{ave} are always computed over the 273-point grid.

The design points of the selected designs with asterisks form the first part of Table 2—the seven points common to all of the designs plus those that relate to a particular n. The first two elements of the tabulated points are the levels $0, 1, \ldots, 12$ of two of the mixture variables; the third element represents the levels 0, 1, or 2 of the process variable. The remaining mixture component is fixed by the mixture constraints (14).

For $r = 84$, mean execution times per attempt on an IBM 3081 computer were .68, 1.32, and 1.73 seconds for D, G, and V optimality. For $r = 273$, the mean execution times increased to 1.88, 4.26, and 4.87 seconds, respectively.The averages are taken across the six values of n and are based on 60 attempts.

A fair comparison of the designs produced here with those of VDY is not possible: their corresponding designs perform poorly but are disadvantaged by being sequentially constructed over a less flexible grid.

The second series of examples typifies the program of leaching experiments that motivated this research. In leaching uranium from an ore to recover the metal in solution, four operating conditions might be ex-

Table 2. *Design-Point Levels for Designs Marked * in Tables 1 and 3*

n	Levels							
		Example With Three Mixture Variables and One Process Variable[a]						
10	0, 1, 2	4, 4, 0	6, 0, 1					
11	0, 0, 2	4, 4, 0	4, 4, 2	6, 0, 1				
12	0, 0, 2	3, 3, 0	3, 3, 2	5, 0, 1	6, 6, 2			
13	0, 0, 2	0, 6, 2	4, 3, 0	5, 3, 2	6, 0, 1	6, 6, 2		
14	0, 0, 1	0, 0, 2	0, 6, 2	4, 4, 0	6, 0, 1	6, 0, 2	6, 6, 2	
15	0, 0, 2	0, 5, 2	0, 12, 1	3, 6 0	6, 0, 0	6, 0, 2	6, 6, 2	7, 0, 1
		Four-Variable Leaching Experiment Example[b]						
15	1, 1, 1, 0	1, 1, 2, 1	1, 2, 1, 1					
16	0, 1, 1, 1	1, 0, 1, 1	1, 1, 1, 2	1, 1, 2, 1				
17	0, 2, 1, 1	1, 0, 1, 1	1, 1, 0, 1	1, 1, 1, 0	2, 1, 1, 1			
18	0, 1, 1, 2	1, 1, 1, 0	1, 1, 2, 1	1, 2, 1, 1	2, 0, 1, 1	2, 1, 0, 1		
19	0, 1, 1, 2	0, 1, 2, 1	0, 2, 1, 1	1, 0, 1, 1	1, 1, 0, 1	1, 1, 1, 0	2, 1, 1, 1	
20	0, 0, 1, 1 2, 1, 2, 1	0, 1, 1, 0	0, 1, 2, 1	1, 1, 0, 1	1, 1, 1, 2	1, 2, 1, 1	2, 0, 1, 1,	

[a]Common: 0, 0, 0; 0, 6, 1; 0, 12, 0; 0, 12, 2; 6, 6, 1; 12, 0, 0; 12, 0, 2.
[b]Fixed: 0, 0, 0, 0; 0, 0, 2, 2; 0, 2, 0, 2; 0, 2, 2, 0; 2, 0, 0, 0; 2, 0, 2, 2; 2, 2, 0, 2; 2, 2, 2, 0. Common: 1, 0, 0, 2; 1, 0, 2, 0; 1, 2, 0, 0; 1, 2, 2, 2.

plored: ore grind size, reaction time, sulphuric acid concentration, and temperature. All of these explanatory variables are continuous, but it is convenient in practice to restrict the experiment to three levels for each factor, the minimum for estimation of a second-order model. The levels are

1. Grind size (mm)—.5, 1.0, 1.5
2. Reaction time (hours)—8, 16, 24
3. Acid concentration (g/litre)—30, 65, 100
4. Temperature (°C)—50, 70, 90

with the lower and upper levels coinciding with the range of operating conditions of interest.

Although it is convenient to experiment at only these levels, one usually wishes to estimate the yield at any set of conditions within the factor ranges. Therefore, $d_{\max}$ and d_{ave} are considered over a finer grid with each explanatory variable at seven recoded levels $-1(1/3)1$. Thus the region of interest is approximated by $r = 7^4 = 2{,}401$ points. To restrict the design to three levels, we set the upper-replication constraint in (3) as $u_j = 1$ if x_j has all four variables at levels -1, 0, or 1, and $u_j = 0$ otherwise ($j = 1, \ldots, 2{,}401$).

Similarly, the lower constraints in (3) may be employed for augmentation. Suppose that observations have already been taken for a 2^{4-1} experiment: the eight fixed points in the second part of Table 2. At these points in the design region we must have at least one observation in the augmented design, and l_j in (3) adopts the value $1/n$ accordingly. Elsewhere we set $l_j = 0$.

The second-order model in the four explanatory variables $x_{(i)} = (x_{(i)1}, \ldots, x_{(i)4})$,

$$E(Y_{(i)}) = \beta_0 + \sum_{s=1}^{4} \beta_s x_{(i)s} + \sum_{s=1}^{4} \sum_{t=s}^{4} \beta_{st} x_{(i)s} x_{(i)t}, \qquad i = 1, \ldots, n,$$

had 15 parameters. We consider augmented designs of $n = 15, \ldots, 20$ observations. Ten attempts were made with the D and V excursion criteria for each n; G optimality was also applied 10 times for $n = 15$ and 16 but only five times for $n = 17, \ldots, 20$.

The properties of selected designs are presented in Table 3. As in the first example, the cataloged designs show little variation in $\{\det(M^{-1})\}^{1/k}$, but this time there are only modest differences in $d_{\max}$. The G excursion criterion also fails to achieve the smallest $d_{\max}$ for $n = 16$ and $n = 18$, though the difference is marginal in the latter instance.

Table 3. *Design Properties for the Four-Variable Leaching Experiment Example*

n	Criterion	d_{max}	d_{ave}	$\{det\ (M^{-1})\}^{1/k}$	λ_{max}
15	*V*	30.0	15.8	2.36	20.3*
	D, G	30.0	19.0	2.36	34.4
16	*V*	27.2	13.7	2.36	17.7*
	G	28.0	14.3	2.42	18.1
	D	30.8	17.3	2.33	30.8
17	*G*	27.2	14.4	2.37	19.3
	V	28.0	13.1	2.35	18.8*
	D	29.5	13.5	2.30	21.1
	V	66.0	13.0	2.40	12.1
18	*D*	24.7	12.9	2.28	19.6
	G	25.2	13.0	2.29	19.9
	D	26.0	14.1	2.26	22.1
	V	26.1	11.6	2.28	11.7*
19	*G*	24.3	12.7	2.29	20.6
	V	25.3	11.2	2.31	10.2*
	D	25.5	13.6	2.27	24.0
	D	26.7	12.1	2.24	12.9
	D	27.9	14.3	2.23	23.1
20	*G*	25.1	12.8	2.29	18.1
	V	27.1	11.0	2.29	12.1*
	D	27.8	14.1	2.23	27.1
	D	28.0	12.3	2.23	13.5
	V	30.4	11.3	2.41	10.6
	D	31.9	14.6	2.20	22.4

* Compromise among the various criteria.

Turning to d_{ave} and λ_{max}, the V criterion is superior in all cases, with particularly large improvements relative to D optimality for $n = 15$, 16, and 18. The saturated case $n = 15$ is of interest in that the three excursion criteria produce designs with the same d_{max} and $(\det(M^{-1})\}^{1/k}$, yet D and G happen to choose a design with 20% larger d_{ave} and 70% larger λ_{max} than the V-optimal design.

For all six values of n, the V-optimal designs with asterisks in Table 3 offer a reasonable compromise among the various criteria. Their design points may be found in Table 2. The 15-point design, for instance, is made up of the eight points fixed by prior experimentation, four points commonly chosen, and the three further points particular to $n = 15$. The levels 0, 1, and 2 correspond to recoded values -1, 0, and 1.

Mean execution times per attempt on an IBM 3081 computer were 5.6, 7.3, and 31.1 seconds for D, V, and G optimality. The D and V averages are based on 60 attempts and the G average on 40 tries. G optimality is much more expensive in this example partly because it considers and, therefore, updates the response-estimator variances throughout the 2,401-point grid. D and V optimality only update their working arrays over the restricted 81-point grid. For V optimality, all of the 2,401 points are involved in the summation that initializes $A(p)$ in (13), but the quadratic forms $a(x_j, p)$ are only required for potential design points x_j in the restricted set.

4. CONCLUSIONS

D optimality has received more attention in the design literature than any other optimality criterion. Yet the foregoing examples (and others not reported here) show that D optimality may exploit improvements in $\{\det(M^{-1})\}^{1/k}$ that are of little practical significance. These improvements may be at appreciable expense to other criteria, at least for small, near-saturated designs. Ironically DETMAX can be too successful in finding D-optimal designs: If it produced a comprehensive range of suboptimal designs, we might find one with better all-round performance.

Nevertheless, the excursion in DETMAX is a powerful heuristic. Where the objective is response estimation, the G and V variants may produce improvements in d_{max} and d_{ave} large enough to be practically worthwhile. It has been shown that these criteria sometimes generate designs that also achieve substantially smaller values of λ_{max}. Whatever the objective—in practice most experiments have multiple objectives—criterion-robust designs are preferred. By trying a number of excursion criteria, one is more likely to find such a robust design.

Part of D optimality's popularity stems from computational convenience. Attention to efficient updating and other computational details, however, makes the time penalty of the G and V criteria more acceptable. An analysis of the computer operations involved suggests that V optimality will be about twice as expensive as D optimality. The corresponding ratio for G optimality is less clear in general, but the noted examples required from two to six times the computational effort for D optimality. The implemented algorithms have been applied to unrestricted design regions of up to about 1,000 points.

The approach adopted here is to specify the design region by a finite set of points. This affords flexibility to restrict the design to convenient levels of the explanatory variables or to approximate continuous variables by finer grids. Such an approach will become computationally cumbersome, though, for a large number of continuous variables.

A FORTRAN listing of the Algorithms for the Construction of Experimental Designs (ACED) computer package is available from the author. Version 1.5.1 implements all of the facilities described here.

ACKNOWLEDGMENTS

This research was funded by Warren Spring Laboratory in Stevenage, England, while the author was a member of the Department of Mathematics of Imperial College. The author also thanks an associate editor and a referee for their comments and A. C. Atkinson for his guidance and encouragement.

[*Received July 1981. Revised September 1983.*]

REFERENCES

ATWOOD, C. L. (1973), "Sequences Converging to *D*-Optimal Designs of Experiments," *Annals of Statistics*, 1, 342–352.

BOX, G. E. P., and DRAPER, N. R. (1959), "A Basis for the Selection of a Response Surface Design," *Journal of the American Statistical Association*, 54, 622–654.

CORNELL, J. A. (1981), *Experiments With Mixtures*, New York: John Wiley.

FEDOROV, V. V. (1972), *Theory of Optimal Experiments*, eds. and trans. W. J. Studden and E. M. Klimko, New York: Academic Press.

GALIL, Z., and KIEFER, J. (1977), "Comparison of Box–Draper and *D*-Optimum Designs for Experiments With Mixtures," *Technometrics*, 19, 441–444.

——— (1980), "Time- and Space-Saving Computer Methods, Related to Mitchell's DETMAX, for Finding *D*-Optimum Designs," *Technometrics*, 22, 301–313.

KIEFER, J. (1975), "Optimal Design: Variation in Structure and Performance Under Change of Criterion," *Biometrika*, 62, 277–288.

KIEFER, J., and WOLFOWITZ, J. (1960), "The Equivalence of Two Extremum Problems," *Canadian Journal of Mathematics*, 12, 363–366.

KNUTH, D. E. (1973), *The Art of Computer Programming, Vol. 3, Sorting and Searching*, Reading, Mass.: Addison-Wesley.

MITCHELL, T. J. (1974), "An Algorithm for the Construction of *D*-Optimal Experimental Designs," *Technometrics*, 16, 203–210.

VUCHKOV, I. N., DAMGALIEV, D. L., and YONTCHEV, C. A.

(1981), "Sequentially Generated Second Order Quasi *D*-Optimal Designs for Experiments With Mixture and Process Variables," *Technometrics*, 23, 233–238.

WELCH, W. J. (1982), "Branch-and-Bound Search for Experimental Designs Based on *D* Optimality and Other Criteria," *Technometrics*, 24, 41–48.

——— (1983), "A Mean Squared Error Criterion for the Design of Experiments," *Biometrika*, 70, 205–213.