Ann. Inst. Statist. Math.
Vol. 44, No. 4, 659–671 (1992)

TREATING BIAS AS VARIANCE FOR EXPERIMENTAL DESIGN PURPOSES

NORMAN R. DRAPER[1] AND IRWIN GUTTMAN[2]

[1]*Statistics Department, University of Wisconsin, 1210 West Dayton Street, Madison, WI 53706, U.S.A.*
[2]*Statistics Department, University of Toronto, Toronto, Ontario, Canada M5S 1A1*

(Received July 1, 1991; revised October 21, 1991)

Abstract. When an empirical model is fitted to data, bias can arise from terms that have not been incorporated, and this can have an important effect on the choice of an experimental design. Here, the biases are treated as random, and the consequences of this action are explored for the fitting of models of first and second order.

Key words and phrases: Bias error, experimental design, response surface, variance error.

1. Introduction

Suppose we wish to fit, by least squares, a linear model of the form

$$y = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{1.1}$$

where y is an $n \times 1$ vector of response observations, $\boldsymbol{X}$ is an $n \times p$ matrix whose p columns consist of n sets of observations on p predictor variables (one per column), where $\boldsymbol{\beta}$ is a $p \times 1$ vector of parameters to be estimated, and $\boldsymbol{\epsilon}$ is a vector of errors in the y's. Typically it is assumed $\boldsymbol{\epsilon} \sim N(\boldsymbol{0}, \boldsymbol{I}\sigma^2)$. The problem of experimental design is then the choice of n "design points" at which to observe the y-values; essentially we have to select $\boldsymbol{X}$ (or some part of it in cases where parts of $\boldsymbol{X}$ depend on other parts). For example, if (1.1) represents a polynomial in $x_1, x_2, \ldots, x_k$, we need to select the values $\boldsymbol{x}_u' = (x_{1u}, x_{2u}, \ldots, x_{ku})$, $u = 1, 2, \ldots, n$, from which $\boldsymbol{X}$ is generated. The $n \times k$ matrix whose u-th row is $\boldsymbol{x}_u'$ is usually called the design matrix. After the design has been run, $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$ provides the least squares vector estimate of $\boldsymbol{\beta}$.

If we are sure that our choice of (1.1) is correct, the design can be selected according to a "variance" criterion. Popular criteria are those of A-, D-, E- and G-optimality. A-optimality requires minimization of trace $(\boldsymbol{X}'\boldsymbol{X})^{-1}$; for D-optimality one maximizes $D = |\boldsymbol{X}'\boldsymbol{X}|/n^p$; the largest eigenvalue of $(\boldsymbol{X}'\boldsymbol{X})^{-1}$ is

minimized for E-optimality; while for G-optimality, one minimizes, over design choice ξ, the maximum over a selected design region, of the variance of the predicted value of $\hat{y}$; we can write this as

$$\min_{\xi} \max_{x} \boldsymbol{u}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{u} \tag{1.2}$$

where $\boldsymbol{u}'$ denotes a general vector whose form is that of a row of the $\boldsymbol{X}$ matrix.

We shall be concerned with designs with n distinct points. A more general "measure design" approach allows the proof of several powerful results including the famous equivalence theorem which connects D-optimality to other types of variance optimality. Much of the work in this area was initiated and carried out by Kiefer and coauthors. See Kiefer (1985) and also Bunke, H. and Bunke, O. ((1986), Section 8.1.3). A particularly wide set of references will also be found in Bandemer *et al.* (1987).

Another approach is to fit (1.1) but to fear that it may not be fully adequate. The fear is only a modest one. (If we felt sure that (1.1) were completely out of the question, it would not be fitted.) We can decide to choose our design to guard against the possibility that the model

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\epsilon} \tag{1.3}$$

is the correct one, where $\boldsymbol{X}_2\boldsymbol{\beta}_2$ represents extra terms not in (1.1). For example, (1.1) might be a plane, (1.3) a quadratic equation. The extra terms in (1.3) would, in general, be expected to produce bias error in the least squares estimates, in addition to the variance error considered in the variance-only approach above. See Box and Draper (1959, 1963) for the basic philosophy; this work has been continued and extended in numerous papers some of which are listed in the references, viz., Draper and Guttman (1986), Draper and Lawrence (1965*a*, 1965*b*, 1966, 1967).

In both of the above approaches, assumptions are usually made about a region of operability, O say, of the x-space within which experiments can be performed, and a region of interest, R say, within which predictive goodness is desired. Often O will include R. When only variance is considered, it is usually assumed that $R = O$ and then the best design is the one that places most or all points on the outer boundary of this joint region. When the model is feared to be inadequate, the appropriate best design depends on the relative size of the bias error B compared to the variance error V, where these quantities are defined in Box and Draper (1959). Although B is unknown, it can be shown that even if the variance error V (which *is* known apart from σ^2) is several times as large as the bias error B, the best design is much closer to the all-bias design than to the all-variance design. The all-bias design, which is the design theoretically appropriate when there is no variance error, and whose moments are the same as those of the region R, "crouches" well within the region R. As the size of V increases, the best design gradually expands. In a typical case, when V and B are of roughly the same size, the design points may lie inside and close to the boundary of R (see Draper and Guttman (1986)). For $V \gg B$, the design generally passes outside R and into the part of O outside R. When we reach the all-variance, no bias case, the

design points lie on the boundary of O, in theory. In practice, experimenters show reluctance to experiment outside R.

One can, of course, argue about the relative merits of the two approaches. Instead, we explore an approach that is, to a limited extent, a compromise between the two.

2. Bias as variance

We now assume that, in (1.3), the elements of $\boldsymbol{\beta}_2$ are not fixed effects, but random effects such that

$$\boldsymbol{\beta}_2 \sim N(\boldsymbol{\beta}_2^*, \sigma_\beta^2 \boldsymbol{I}). \tag{2.1}$$

(Two previous but different approaches have points of contact with this. Welch (1983) suggests that it might be possible to place an upper bound on the maximum absolute value of all the elements of $\boldsymbol{X}_2\boldsymbol{\beta}_2$ in (1.3). He offers a computer algorithm that can be used to obtain measure-type and integer type designs, and he illustrates his method by a discussion of designs in two factors based at the sites of a 3^2 factorial design and used for fitting a first order model. Steinberg (1985) applies a Bayesian approach to the entire model (1.3) assuming that the $\boldsymbol{X}_2\boldsymbol{\beta}_2$ terms consist of orthogonal polynomials. His examples also relate to first order models and designs, and he discusses 2^{k-p} type designs with 4, 8 and 16 runs. Our development is not Bayesian and uses a variance function as a criterion, incorporating the bias errors as random rather than fixed effects.) We now apply the result

$$V(\boldsymbol{u}) = E_v\{V(\boldsymbol{u} \mid \boldsymbol{v})\} + V_v\{E(\boldsymbol{u} \mid \boldsymbol{v})\} \tag{2.2}$$

to the usual least squares estimator $\hat{\boldsymbol{\beta}}$, regarded as $\boldsymbol{u}$, fitted on the (wrong) assumption that (1.1) is appropriate. Here $\boldsymbol{v}$ is $\boldsymbol{\beta}_2$. We find that

$$\boldsymbol{V}(\hat{\boldsymbol{\beta}}) = \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1} + \sigma_\beta^2 \boldsymbol{A}\boldsymbol{A}' \tag{2.3}$$

where $\boldsymbol{A} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X}_2$ is the so-called alias matrix (also sometimes called the bias matrix). Note that normalizing (2.3) by n/σ^2, and writing $\gamma = \sigma_\beta^2/\sigma^2$ provides

$$\boldsymbol{V} = (n/\sigma^2)\,\boldsymbol{V}(\hat{\boldsymbol{\beta}}) = n(\boldsymbol{X}'\boldsymbol{X})^{-1} + \gamma n \boldsymbol{A}\boldsymbol{A}'. \tag{2.4}$$

This enables us to consider a variance-type approach while incorporating "bias" through the presence of γ.

We now consider introduction of the regions R and O. The quantity $\boldsymbol{u}'\boldsymbol{V}\boldsymbol{u}$ where $\boldsymbol{V}$ is from (2.4), and where $\boldsymbol{u}$ is a row vector whose elements are of the same form as a row of $\boldsymbol{X}$, is n/σ^2 times the variance of $\hat{y}$ at a general location $(x_1, x_2, \ldots, x_k)$. The quantity

$$L = \int_O W(\boldsymbol{x})\boldsymbol{u}'\boldsymbol{V}\boldsymbol{u}d\boldsymbol{x} = \text{trace } \boldsymbol{V}\int_O W(\boldsymbol{x})\boldsymbol{u}\boldsymbol{u}'d\boldsymbol{x} = \text{trace } \boldsymbol{V}\boldsymbol{\mu}, \tag{2.5}$$

is an averaged variance value, over O, with respect to a weight function $W(\boldsymbol{x})$. A simple choice of $W(\boldsymbol{x})$, for example, would be to set $W(\boldsymbol{x}) = 0$ for x outside R and within O and, within R, to set

$$W^{-1}(\boldsymbol{x}) = \int_R d\boldsymbol{x}. \tag{2.6}$$

Then

$$\boldsymbol{\mu} = \int_R \boldsymbol{u}\boldsymbol{u}' d\boldsymbol{x} \Big/ \int_R d\boldsymbol{x}. \tag{2.7}$$

Such a choice corresponds to uniform interest within R, none outside, and then $\boldsymbol{\mu}$ is a matrix of moments of a uniform distribution over R. Whatever choice is made for the weight function we can use the criterion "minimize L" to select a suitable design under any specific set of assumptions. We illustrate use of this criterion via two examples.

3. Examples

Example 1. Suppose we wish to fit a first order model to k factors $x_1, x_2, \ldots, x_k$, fearing some second order bias. Assume R is the unit sphere of radius 1. Restrict the design class to *orthogonal* designs with third order moments zero and pure second moments all equal to c, so that for design points $(x_{1u}, x_{2u}, \ldots, x_{ku})$, $u = 1, 2, \ldots, n$,

$$c = n^{-1} \sum_{u=1}^{n} x_{iu}^2. \tag{3.1}$$

Take $W(\boldsymbol{x})$ uniform over R, zero outside R. This implies that $\boldsymbol{\mu} = \text{diagonal}(1, \mu_2, \ldots, \mu_2)$, with $\mu_2 = 1/(k+2)$. It is easily shown that (2.5) reduces to

$$L = 1 + \gamma k n c^2 + k\mu_2/c \tag{3.2}$$

which is minimized when

$$c = \{2\gamma n(k+2)\}^{-1/3}. \tag{3.3}$$

If we consider using a 2^{k-p} fractional factorial of resolution IV or higher with points at $(\pm\theta, \pm\theta, \ldots, \pm\theta)$ where θ is a scale factor, plus n_0 center points, making $n = 2^{k-p} + n_0$ in all, then $nc = 2^{k-p}\theta^2$ and (3.3) becomes

$$\theta = n^{1/3}\{2^{3(k-p)+1}\gamma(k+2)\}^{-1/6}. \tag{3.4}$$

The axial distance of the points is then $r = \theta k^{1/2}$ which can be compared with the unit radius of R to see if the points are inside or outside R for a particular solution.

What values of γ would be "typical"? We offer the argument that the variance contribution from the bias terms might, not unreasonably, be about the same size as $V(\epsilon) = \sigma^2$ over R. Now

$$V\{\beta_{11}x_1^2 + \cdots + \beta_{kk}x_k^2 + \beta_{12}x_1x_2 + \cdots + \beta_{k-1,k}x_{k-1}x_k\}$$
$$= \sigma_\beta^2\{x_1^4 + \cdots + x_k^4 + x_1^2x_2^2 + \cdots + x_{k-1}^2x_k^2\}$$

with a maximum value within R at $x_i = k^{-1/2}$, of $\sigma_\beta^2 k^{-2}\{k + k(k-1)/2\} = \sigma_\beta^2(k+1)/(2k)$. Setting this maximum value equal to σ^2 provides $\gamma = 2k/(k+1)$. For various k, this is

k =	2	3	4	5	6	7	8
γ =	1.33	1.50	1.60	1.67	1.71	1.75	1.77

so that, under this argument, comparatively large values of γ still lead only to a maximum doubling of the error variance over R. Table 1 shows some values of $r = \theta k^{1/2}$ for $0.1 \le \gamma \le 2$ for the values of k, p and n_0 shown. We see that in "typical" (as defined above) cases, the design points are drawn somewhat within the region R. Higher γ values (more bias) would shrink the designs more, lower γ values (less bias) would lead to an expanded design, as shown. We see that designs with points on or outside the boundary of R are appropriate only for quite low bias cases.

We argue that the above is an appropriate way to implement a variance type criterion, by defining it in a way to allow the possibility of an inadequate model to enter.

(An aside: What would happen if we ignored the region of interest and attempted to apply the A, D, E and G criteria directly to (2.3) rather than to $(\boldsymbol{X}'\boldsymbol{X})^{-1}$? For the A criterion, the solution is like (3.3) but with factor $(k+2)$ omitted so that θ is bigger by a factor of $(k+2)^{1/6}$; for the G criterion $(k+2)$ in (3.3) is replaced by $k/r_{\max}^2$ where $r_{\max}^2$ is the distance of the largest x-site desired. If $r_{\max}^2 = 1$, θ is smaller by a factor of $[k/(k+2)]^{1/6}$. The D criterion gives $c = (\gamma n)^{-1/2}$ if $k = 1$, and $c = \infty$ if $k \ge 2$. For the E criterion, the only solution possible occurs when the eigenvalues are equal namely when $1 + \gamma knc^2 = c^{-1}$. For a 2^{k-p} design this spreads the points somewhat more widely than does the criterion (3.4). The numbers shown in Table 1 would, typically, be about 15–20% larger.)

We can make a partial rough comparison of our Table 1 results with those of Steinberg ((1985), p. 523) for the 2^k designs, $k = 2$, 3 and 4, and the 2^{5-1} design. A difficulty is the following: our region R is a unit circle ($k = 2$) or sphere ($k = 3, 4$). This is equivalent to a weight function uniform over the unit circle or sphere. Steinberg's equivalent of R is a multivariate unit normal weight function whose highest point is at the origin and which falls off to zero in all directions. As a rough conversion, we divide Steinberg's d-ranges (Steinberg's d is our θ) by the radii of the 95% central region of his weight density, and multiply by $k^{1/2}$ to produce figures comparable to those of our Table 1. This provides Table 2, in which we also show the corresponding ranges of our values from Table 1. Steinberg's and

Table 1. Values of $r = \theta k^{1/2}$ for the factorial point distance of a 2^{k-p} design via (3.4), for $0.2 \leq \gamma \leq 2$ and for the values of k, p and n_0 shown. When $r < 1$, the points lie within the region of interest R.

$k = 2$, $p = 0$

	n_0				
γ	0	1	2	3	4
0.20	1.038	1.118	1.188	1.251	1.308
0.40	0.925	0.996	1.058	1.114	1.165
0.60	0.864	0.931	0.989	1.041	1.089
0.80	0.824	0.887	0.943	0.993	1.038
1.00	0.794	0.855	0.909	0.956	1.000
1.20	0.770	0.829	0.881	0.928	0.970
1.40	0.750	0.808	0.859	0.904	0.945
1.60	0.734	0.791	0.840	0.884	0.925
1.80	0.720	0.775	0.824	0.867	0.907
2.00	0.707	0.762	0.809	0.852	0.891

$k = 3$, $p = 0$

γ	0	1	2	3	4
0.20	1.091	1.135	1.175	1.213	1.249
0.40	0.972	1.011	1.047	1.081	1.113
0.60	0.909	0.945	0.979	1.010	1.040
0.80	0.866	0.901	0.933	0.963	0.991
1.00	0.834	0.868	0.899	0.928	0.955
1.20	0.809	0.842	0.872	0.900	0.927
1.40	0.789	0.820	0.850	0.877	0.903
1.60	0.772	0.802	0.831	0.858	0.883
1.80	0.757	0.787	0.815	0.841	0.866
2.00	0.743	0.773	0.801	0.827	0.851

$k = 4$, $p = 0$

γ	0	1	2	3	4
0.20	1.089	1.111	1.132	1.153	1.173
0.40	0.970	0.990	1.009	1.027	1.045
0.60	0.907	0.925	0.943	0.960	0.977
0.80	0.864	0.882	0.899	0.915	0.931
1.00	0.833	0.850	0.866	0.882	0.897
1.20	0.808	0.824	0.840	0.855	0.870
1.40	0.787	0.803	0.819	0.834	0.848
1.60	0.770	0.786	0.801	0.815	0.829
1.80	0.755	0.770	0.785	0.799	0.813
2.00	0.742	0.757	0.772	0.786	0.799

Table 1. (continued).

$k = 5, p = 0$

γ	0	1	2	3	4
0.20	1.057	1.068	1.079	1.089	1.099
0.40	0.942	0.951	0.961	0.970	0.979
0.60	0.880	0.889	0.898	0.907	0.915
0.80	0.839	0.848	0.856	0.864	0.873
1.00	0.808	0.817	0.825	0.833	0.841
1.20	0.784	0.792	0.800	0.808	0.816
1.40	0.764	0.772	0.780	0.787	0.795
1.60	0.747	0.755	0.763	0.770	0.777
1.80	0.733	0.740	0.748	0.755	0.762
2.00	0.720	0.728	0.735	0.742	0.749

$k = 5, p = 1$

γ	0	1	2	3	4
0.20	1.187	1.211	1.234	1.256	1.278
0.40	1.057	1.079	1.099	1.119	1.139
0.60	0.988	1.008	1.028	1.046	1.064
0.80	0.942	0.961	0.979	0.997	1.014
1.00	0.907	0.926	0.944	0.961	0.977
1.20	0.880	0.898	0.915	0.932	0.948
1.40	0.858	0.875	0.892	0.908	0.924
1.60	0.839	0.856	0.873	0.888	0.904
1.80	0.823	0.839	0.856	0.871	0.886
2.00	0.808	0.825	0.841	0.856	0.871

$k = 6, p = 1$

γ	0	1	2	3	4
0.20	1.132	1.144	1.156	1.167	1.178
0.40	1.009	1.019	1.030	1.040	1.049
0.60	0.943	0.953	0.962	0.972	0.981
0.80	0.899	0.908	0.917	0.926	0.935
1.00	0.866	0.875	0.884	0.892	0.901
1.20	0.840	0.849	0.857	0.866	0.874
1.40	0.819	0.827	0.836	0.844	0.852
1.60	0.801	0.809	0.817	0.825	0.833
1.80	0.785	0.793	0.801	0.809	0.817
2.00	0.772	0.779	0.787	0.795	0.802

Table 1. (continued).

$k = 7$, $p = 1$

γ	0	1	2	3	4
0.20	1.069	1.074	1.080	1.085	1.090
0.40	0.952	0.957	0.962	0.967	0.971
0.60	0.890	0.894	0.899	0.903	0.908
0.80	0.848	0.853	0.857	0.861	0.865
1.00	0.817	0.821	0.826	0.830	0.834
1.20	0.793	0.797	0.801	0.805	0.809
1.40	0.773	0.777	0.781	0.784	0.788
1.60	0.756	0.760	0.763	0.767	0.771
1.80	0.741	0.745	0.749	0.752	0.756
2.00	0.728	0.732	0.736	0.739	0.743

$k = 8$, $p = 1$

γ	0	1	2	3	4
0.20	1.000	1.003	1.005	1.008	1.010
0.40	0.891	0.893	0.896	0.898	0.900
0.60	0.833	0.835	0.837	0.839	0.841
0.80	0.794	0.796	0.798	0.800	0.802
1.00	0.765	0.767	0.769	0.771	0.773
1.20	0.742	0.744	0.746	0.748	0.749
1.40	0.723	0.725	0.727	0.729	0.730
1.60	0.707	0.709	0.711	0.713	0.714
1.80	0.693	0.695	0.697	0.699	0.701
2.00	0.681	0.683	0.685	0.687	0.688

Table 2. Rough conversion of Steinberg's ((1985), p. 523) d-ranges.

k	Steinberg's d	95% radius, r	$k^{1/2}d/r$ values	Range from Table 1
2	1.19–1.50	2.45	0.69–0.87	0.71–1.47
3	1.09–1.49	2.80	0.67–0.92	0.74–1.40
4	0.96–1.49	3.08	0.62–0.97	0.74–1.32
5	0.86–1.41	3.33	0.58–0.95	0.81–1.28

our results are broadly consistent in spite of the very different assumptions made in the two cases and the approximations made in the conversions.

Welch's (1983) example is for a 3^2-site design for fitting a first order model and he is concerned with design point weighting, rather than with design re-scaling.

Example 2. Suppose we wish to fit a second order model to k factors x_1, $x_2, \ldots, x_k$, fearing some third order bias. Assume R is the unit sphere of radius

1. Restrict the design class to second order rotatable designs with pure second moments c, pure fourth moments $3f$ and mixed even fourth order moments f (and all odd moments of order five or less are zero), so that for design points $(x_{1u}, x_{2u}, \ldots, x_{ku})$, $u = 1, 2, \ldots, n$

$$c = n^{-1}\sum_{u=1}^{n} x_{iu}^2, \qquad 3f = n^{-1}\sum_{u=1}^{n} x_{iu}^4 = 3n^{-1}\sum_{u=1}^{n} x_{iu}^2 x_{ju}^2. \tag{3.5}$$

Take $W(\boldsymbol{x})$ uniform over R, zero outside R. This implies that, when the second order polynomial terms are in the order

$$1; x_1, x_2, \ldots, x_k; x_1^2, x_2^2, \ldots, x_k^2; x_1x_2, \ldots, x_{k-1}x_k, \tag{3.6}$$

$$\boldsymbol{\mu} = \begin{bmatrix} 1 & \mathbf{0}' & \mu_2\mathbf{1}' & \mathbf{0} \\ \mathbf{0} & \mu_2\boldsymbol{I} & \mathbf{0} & \mathbf{0} \\ \mu_2\mathbf{1} & \mathbf{0} & (2\boldsymbol{I}+\boldsymbol{J})\mu_{22} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mu_{22}\boldsymbol{I} \end{bmatrix} \tag{3.7}$$

where $\mathbf{1}' = (1, 1, \ldots, 1)$ of length k, where $\boldsymbol{J} = \mathbf{1}\mathbf{1}'$, where the dimensions correspond to the semicolons in (3.6), with $\mu_2 = 1/(k+2)$ and $\mu_{22} = 1/\{(k+2)(k+4)\}$. It can be shown that (2.5) becomes

$$L = D + k\mu_2(2E + c^{-1} + n\gamma Q) + k\mu_{22}\{3F + (k-1)(P + (2f)^{-1})\} \tag{3.8}$$

where

$$\begin{aligned} D &= 2(k+2)f^2/H, \\ E &= -2fc/H, \\ F &= \{(k+1)f - (k-1)c^2\}/H, \\ P &= (c^2 - f)/H, \\ H &= 2f\{(k+2)f - kc^2\}, \\ Q &= (k+8)f^2/c^2. \end{aligned} \tag{3.9}$$

If we specialize further to a central composite design consisting of (i) a 2^{k-p} fractional factorial design, with points at $(\pm 1, \pm 1, \ldots, \pm 1)\theta$, (ii) $2k$ axial points $(\pm\alpha, 0, 0, \ldots, 0), (0, \pm\alpha, 0, \ldots, 0), \ldots, (0, 0, 0, \ldots, \pm\alpha)$ where $\alpha = 2^{(k-p)/4}\theta$ to ensure rotatability and (iii) n_0 center points at $(0, 0, \ldots, 0)$, then

$$\begin{aligned} nc &= (2^{k-p} + 2^{(k-p)/2})\theta^2, \\ nf &= 2^{k-p}\theta^4, \\ n &= 2^{k-p} + 2k + n_0 \end{aligned} \tag{3.10}$$

and we can determine the value of θ that minimizes L for given k, p, n_0 and γ. Again, the value of $\theta k^{1/2}$ enables us to see whether the factorial points lie inside, on or outside the unit radius region of interest.

Table 3. Values of $r = \theta k^{1/2}$, for the factorial point distance of the 2^{k-p} portion of a rotatable central composite design, which minimize (3.8) for $0.50 \leq \gamma \leq 4$ and for the values of k, p and n_0 shown. When $r < 1$ the 2^{k-p} portion points lie within the region of interest R.

$k = 2$, $p = 0$

	n_0				
γ	1	2	3	4	5
0.50	0.925	0.965	0.982	0.992	0.998
1.00	0.884	0.903	0.912	0.917	0.920
1.50	0.857	0.868	0.872	0.875	0.876
2.00	0.838	0.842	0.845	0.846	0.847
2.50	0.822	0.823	0.824	0.824	0.824
3.00	0.809	0.807	0.807	0.806	0.806
3.50	0.798	0.794	0.793	0.792	0.791
4.00	0.788	0.783	0.781	0.779	0.779

$k = 3$, $p = 0$

γ	1	2	3	4	5
0.50	0.985	1.032	1.052	1.064	1.071
1.00	0.942	0.967	0.978	0.984	0.987
1.50	0.914	0.929	0.935	0.939	0.941
2.00	0.894	0.902	0.906	0.908	0.910
2.50	0.877	0.882	0.884	0.885	0.886
3.00	0.864	0.865	0.866	0.866	0.867
3.50	0.852	0.851	0.851	0.851	0.851
4.00	0.842	0.839	0.838	0.838	0.837

$k = 4$, $p = 0$

γ	1	2	3	4	5
0.50	0.993	1.043	1.066	1.079	1.087
1.00	0.954	0.981	0.993	1.000	1.004
1.50	0.928	0.944	0.952	0.956	0.958
2.00	0.908	0.918	0.923	0.925	0.927
2.50	0.892	0.898	0.900	0.902	0.903
3.00	0.879	0.881	0.882	0.883	0.884
3.50	0.867	0.867	0.868	0.868	0.868
4.00	0.857	0.855	0.855	0.854	0.854

$k = 5$, $p = 0$

γ	1	2	3	4	5
0.50	0.986	1.031	1.053	1.066	1.074
1.00	0.948	0.972	0.984	0.991	0.995
1.50	0.923	0.937	0.944	0.948	0.951
2.00	0.904	0.912	0.916	0.919	0.920
2.50	0.888	0.893	0.895	0.896	0.897
3.00	0.875	0.877	0.877	0.878	0.878
3.50	0.864	0.863	0.863	0.863	0.863
4.00	0.854	0.852	0.851	0.850	0.849

Table 3. (continued).

$k = 5,\ p = 1$

γ	1	2	3	4	5
0.50	1.124	1.172	1.194	1.206	1.214
1.00	1.066	1.094	1.107	1.114	1.119
1.50	1.031	1.049	1.058	1.063	1.066
2.00	1.005	1.018	1.024	1.028	1.030
2.50	0.984	0.994	0.999	1.001	1.003
3.00	0.968	0.975	0.978	0.980	0.981
3.50	0.953	0.959	0.961	0.962	0.963
4.00	0.941	0.945	0.946	0.947	0.948

$k = 6,\ p = 1$

γ	1	2	3	4	5
0.50	1.059	1.119	1.146	1.162	1.171
1.00	1.018	1.053	1.069	1.078	1.083
1.50	0.991	1.014	1.025	1.030	1.034
2.00	0.970	0.986	0.994	0.998	1.000
2.50	0.953	0.965	0.970	0.973	0.975
3.00	0.939	0.947	0.951	0.953	0.954
3.50	0.927	0.933	0.935	0.936	0.937
4.00	0.917	0.920	0.921	0.922	0.922

$k = 7,\ p = 1$

γ	1	2	3	4	5
0.50	1.026	1.073	1.098	1.113	1.123
1.00	0.988	1.015	1.029	1.037	1.042
1.50	0.962	0.979	0.988	0.993	0.996
2.00	0.942	0.954	0.959	0.963	0.965
2.50	0.926	0.934	0.937	0.939	0.941
3.00	0.913	0.917	0.919	0.921	0.921
3.50	0.902	0.904	0.905	0.905	0.905
4.00	0.892	0.892	0.892	0.892	0.892

$k = 8,\ p = 1$

γ	1	2	3	4	5
0.50	1.043	1.058	1.068	1.076	1.081
1.00	0.984	0.992	0.997	1.000	1.003
1.50	0.949	0.953	0.956	0.958	0.959
2.00	0.923	0.925	0.927	0.928	0.929
2.50	0.903	0.904	0.905	0.905	0.906
3.00	0.887	0.887	0.887	0.887	0.887
3.50	0.873	0.873	0.872	0.872	0.872
4.00	0.862	0.860	0.860	0.859	0.859

We again need to ask what a "typical" value of γ might be, and we again argue that it is one such that the variance contribution from the bias terms is about as big as σ^2. We can write

$$V\{\beta_{111}x_1^3 + \cdots + \beta_{k-2,k-1,k}x_{k-2}x_{k-1}x_k\}$$
$$= \sigma_\beta^2\{x_1^6 + \cdots + x_1^2x_2^4 + \cdots + x_1^2x_2^2x_3^2 + \cdots\}.$$

If we set the maximum value of this over R equal to σ^2 by substituting $x_i = k^{-1/2}$, we find

$$\gamma^{-1} = k^{-3}\{k + k(k-1) + k(k-1)(k-2)/6\}$$
$$= (k+1)(k+2)/(6k^2)$$

so that $\gamma = 6k^2/\{(k+1)(k+2)\}$. Some values of this are as follows

k	=	2	3	4	5	6	7	8
γ	=	2	2.70	3.33	3.57	3.86	4.08	4.27

so that, under this argument, quite large values of γ still lead only to a maximum doubling of the error variance over R. Table 3 shows some calculations of $\theta k^{1/2}$ for γ's of this order and smaller. We see that, in a majority of cases, the factorial design points are slightly inside the region R of unit radius and that the results are insensitive to variation in the number of center points. As in Example 1, the addition of more center points causes the design to expand. An increase in γ causes the designs to contract. Note that, for the series of k and p values shown in Table 3, the ratios of the axial point distances to the cube point distances (from the origin) are given by $2^{(k-p)/4}/k^{1/2} = 1,\ 0.97,\ 1,\ 1.06,\ 0.89,\ 0.97,\ 1.07,\ 1.19$, respectively. Thus for (k,p) values of $(2,0)$, $(3,0)$, $(4,0)$, $(5,1)$, $(6,1)$, the axial points are either on the same sphere as the cube points, or slightly inside. For $(5,0)$, $(7,1)$ the axial points are slightly outside, and for $(8,1)$ rather more outside.

As we mentioned earlier, neither Welch (1983) nor Steinberg (1985) gave examples for second order model fitting, so that no comparisons are possible.

For both first and second order examples, the conclusion is that, in (what we have defined as) "typical situations", the design points should lie slightly inside the region R.

Acknowledgements

N. R. Draper gratefully acknowledges partial support from the National Science Foundation under Grants DMS-8701027 and DMS-8900426, and from the Wisconsin Alumni Research Foundation through the University of Wisconsin Graduate School. I. Guttman is grateful for partial support from the Canadian National Science and Engineering Research Council (NSERC) under Grant No. A8743. We thank the referees.

REFERENCES

Bandemer, H., Näther, W. and Pilz, J. (1987). Once more: optimal experimental design for regression models (with discussion), *Statistics*, **18**, 171–217.

Box, G. E. P. and Draper, N. R. (1959). A basis for the selection of a response surface design, *J. Amer. Statist. Assoc.*, **54**, 622–654.

Box, G. E. P. and Draper, N. R. (1963). The choice of a second order rotatable design, *Biometrika*, **50**, 335–352.

Bunke, H. and Bunke, O. (1986). *Statistical Inference in Linear Models*, Wiley, New York.

Draper, N. R. and Guttman, I. (1986). Response surface designs in flexible regions, *J. Amer. Statist. Assoc.*, **81**, 1089–1094.

Draper, N. R. and Lawrence, W. E. (1965*a*). Designs which minimize model inadequacies; cuboidal regions of interest, *Biometrika*, **52**, 111–118.

Draper, N. R. and Lawrence, W. E. (1965*b*). Mixtures designs for three factors, *J. Roy. Statist. Soc. Ser. B*, **27**, 450–465.

Draper, N. R. and Lawrence, W. E. (1966). The use of second-order "spherical" and "cuboidal" designs in the wrong regions, *Biometrika*, **53**, 596–599.

Draper, N. R. and Lawrence, W. E. (1967). Sequential designs for spherical weight functions, *Technometrics*, **9**, 517–529.

Kiefer, J. (1985). *Collected Papers III, Design of Experiments*, Springer, New York.

Steinberg, D. M. (1985). Model robust response surface designs: scaling two-level factorials, *Biometrika*, **72**, 513–526.

Welch, W. J. (1983). A mean squared error criterion for the design of experiments, *Biometrika*, **70**, 205–213.