Robust Minimax Designs for Multiple Linear Regression

Douglas P. Wiens*
Department of Statistics and Applied Probability
University of Alberta
Edmonton, Alberta T6G 2G1, Canada

Submitted by George P. H. Styan

ABSTRACT

We establish an extension, to the case of multiple regression, of a result on minimax simple regression designs due to P. Huber. Designs are found which are minimax with respect to integrated mean squared error as the true response function varies over an $\mathscr{L}_2$-neighbourhood of (1) a p-dimensional plane or (2) a bivariate surface with possible interactions between the regressors.

1. INTRODUCTION AND SUMMARY

In this paper, we address some problems in the theory of minimax designs for multiple linear regression. Box and Draper (1959) made apparent the problems which can arise from too strict an adherence to the assumed form of a regression function, by analyzing the relative importance of errors due to bias and to variance. They found that very small deviations from the model form can eliminate any supposed gains arising from the use of a design which minimizes variance alone.

In the work of Box and Draper, it is assumed that the true regression function is a polynomial, of possibly higher degree than that of the assumed regression function. Huber (1975) and Marcus and Sacks (1976) formulated and solved some less restricted minimax problems, in which designs are found which minimize the maximum mean squared error as the true regression

*Research supported by the Natural Sciences and Engineering Research Council of Canada.

655 Avenue of the Americas, New York, NY 10010 0024-3795/90/$3.50

function varies over full neighbourhoods, albeit of linear functions of a single variable only. From a robustness point of view the neighbourhoods used by Marcus and Sacks are evidently rather thin, leading as they do to "optimal" designs which concentrate all mass at just two points. Such designs allow no possibility to check the assumed linearity of the regression function.

Somewhat surprisingly, the approach of Huber (1975) has not been generalized to multiple regression. This paper represents an attempt to fill this gap in the literature.

We suppose that an experimenter is to observe values of a regression function $f(\mathbf{x})$ subject to error, at n (not necessarily distinct) design points $\mathbf{x}_i$. The goal of the statistician is to construct a design which minimizes the maximum integrated mean squared error when f is only approximately known. This indeterminateness is formalized by allowing f to vary freely over an $\mathscr{L}_2$-neighbourhood of a fixed class $\mathscr{G}$ of regression functions, linear in several parameters. Specifically,

$$\mathscr{G} = \left\{ g(\mathbf{x}) = \sum_{j=1}^{N} \alpha_j g_j(\mathbf{x}) \,|\, g_1, \ldots, g_N \text{ known, linearly independent} \right.$$
$$\left. \text{functions, } \alpha_1, \ldots, \alpha_N \text{ unknown parameters} \right\}.$$

Assume that the regressors $\mathbf{x} = (x_1, \ldots, x_p)^T$ have been transformed to lie in a region R with unit volume. Specific choices of R will be dictated, in the examples, by the form of $\mathscr{G}$. All integrals are over R.

The approximating class is, for a fixed number $\eta \geqslant 0$,

$$\mathscr{F} = \left\{ f(\mathbf{x}) \,\middle|\, \inf_{g \in \mathscr{G}} \|f - g\| \leqslant \eta \right\}. \tag{1.1}$$

Here, $\|f\| = (\int f^2(\mathbf{x})\, d\mathbf{x})^{1/2}$, the $\mathscr{L}_2$-norm. The experimenter behaves as if $f \in \mathscr{G}$, and calculates estimates $\hat{\alpha}_j$ and $\hat{f}(\mathbf{x}) = \sum_{j=1}^{N} \hat{\alpha}_j g_j(\mathbf{x})$ accordingly. The design, however, is required to give protection against deviations from $\mathscr{G}$ into $\mathscr{F}$.

We remark that this describes a situation commonly faced in practice. When we fit a linear response to a set of data we are typically well aware that the true response is not exactly linear. However, we would be hard pressed to define a nonlinear family of responses appropriate to the given data.

We assume here that $\hat{f}$ is obtained by least squares, and that the design matrix is of full rank. Similar problems, in which $\hat{f}$ is obtained as an M-estimate, are currently being investigated.

The observations are

$$y_i = f(\mathbf{x}_i) + \varepsilon_i, \qquad i = 1, 2, \ldots, n,$$

where the ε_i are uncorrelated, zero mean errors with finite variance σ^2. Let $\xi(\mathbf{x})$ be the design measure, placing mass n^{-1} at each $\mathbf{x}_i$. The loss function is integrated mean squared error:

$$Q(f, \xi) = \int E\left\{\left[f(\mathbf{x}) - \hat{f}(\mathbf{x})\right]^2\right\} d\mathbf{x} \qquad \left(d\mathbf{x} = \prod_{i=1}^{p} dx_i\right). \tag{1.2}$$

The problem is then to choose a ξ_0 which satisfies

$$\sup_f Q(f, \xi_0) = \inf_\xi \sup_f Q(f, \xi). \tag{1.3}$$

The first step, carried out in Section 2, is to construct the least favourable $f_0 \in \mathscr{F}$, for a fixed design ξ. The derivation is an extension of that of Huber (1975).

For any particular class $\mathscr{G}$, the minimax problem can be solved completely. This does not seem possible in general, however, since f_0 depends upon the maximum eigenvalue, and corresponding eigenvector, of a matrix whose elements are functionals of ξ, and whose form varies with $\mathscr{G}$.

In Section 3, these eigenvalue problems are solved in two specific cases. In the first, the response surface to be investigated is thought to be, approximately, a p-dimensional plane. In the second it is a bivariate surface with possible interactions between the regressors. We then find designs ξ_0 minimizing $Q(f_0, \xi)$. These give saddlepoint solutions to (1.3):

$$Q(f, \xi_0) \leqslant Q(f_0, \xi_0) \leqslant Q(f_0, \xi) \tag{1.4}$$

for all $f \in \mathscr{F}$ and all designs ξ.

2. DETERMINATION OF THE LEAST FAVOURABLE f_0

Recall (1.1), (1.2). If g_f is the $\mathscr{L}_2$-closest member of $\mathscr{G}$ to f, then $f - g_f$ is orthogonal to each $g \in \mathscr{G}$, and in particular to $g_f - \hat{f}$. It follows that $Q(f, \xi)$ decomposes into an intrinsic error term, a bias term, and a variance term:

$$Q(f, \xi) = \int [f(\mathbf{x}) - g_f(\mathbf{x})]^2 \, d\mathbf{x} + \int [g_f(\mathbf{x}) - E\hat{f}(\mathbf{x})]^2 \, d\mathbf{x} + \int \mathrm{var}[\hat{f}(\mathbf{x})] \, d\mathbf{x}$$

$$=: Q_f + B(f, \xi) + V(\xi). \tag{2.1}$$

In terms of

$$\mathbf{z} = \mathbf{z}(\mathbf{x}) = (g_1(\mathbf{x}), \ldots, g_N(\mathbf{x}))^T,$$

$$\mathbf{a} = \int \mathbf{z} f(\mathbf{x}) \, d\mathbf{x}, \qquad \mathbf{A} = \int \mathbf{z}\mathbf{z}^T \, d\mathbf{x}, \qquad \boldsymbol{\alpha}^0 = \mathbf{A}^{-1}\mathbf{a},$$

we find by least squares that

$$g_f(\mathbf{x}) = \mathbf{z}^T \boldsymbol{\alpha}^0. \tag{2.2}$$

Similarly, the least squares estimate $\hat{f}(\mathbf{x})$ is obtained in terms of

$$\hat{\mathbf{b}} = \int \mathbf{z} y \, d\xi(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{z}(\mathbf{x}_i) y_i,$$

$$\mathbf{B} = \int \mathbf{z}\mathbf{z}^T \, d\xi(\mathbf{x}), \qquad \hat{\boldsymbol{\alpha}} = \mathbf{B}^{-1}\hat{\mathbf{b}},$$

and is given by

$$\hat{f}(\mathbf{x}) = \mathbf{z}^T \hat{\boldsymbol{\alpha}}. \tag{2.3}$$

With

$$\mathbf{b}(f) = \mathbf{b} := E\hat{\mathbf{b}} = \int \mathbf{z} f(\mathbf{x}) \, d\xi(\mathbf{x}), \qquad \boldsymbol{\alpha}^1 = E\hat{\boldsymbol{\alpha}} = \mathbf{B}^{-1}\mathbf{b}, \tag{2.4}$$

we then have

$$E\hat{f}(\mathbf{x}) = \mathbf{z}^T \boldsymbol{\alpha}^1. \tag{2.5}$$

Now (2.2)–(2.5) in (2.1) give

$$B(f,\xi) = (\boldsymbol{\alpha}^0 - \boldsymbol{\alpha}^1)^T \mathbf{A}(\boldsymbol{\alpha}^0 - \boldsymbol{\alpha}^1),$$

$$V(\xi) = \frac{\sigma^2}{n} \operatorname{tr} \mathbf{A}\mathbf{B}^{-1}.$$

Note that our assumptions imply that $\mathbf{A}$ and $\mathbf{B}$ are positive definite.

To maximize $Q(f,\xi)$ for fixed ξ, we first note that if $h := f - g_f$, then $g_h(\mathbf{x}) \equiv 0$, $Q_f = Q_h = \|h\|^2$, and, using (2.2), $B(f,\xi) = B(h,\xi)$. Clearly $V(\xi)$ is unaffected by the change. Thus $Q(f,\xi) = Q(h,\xi)$ and so we may assume, without loss of generality, that $g_f = 0$. With

$$\mathbf{H} = \mathbf{B}\mathbf{A}^{-1}\mathbf{B}$$

the problem is then:

$$\begin{aligned}&\text{Maximize } B(f,\xi) = \mathbf{b}^T\mathbf{H}^{-1}\mathbf{b} \text{ subject to}\\ &\text{(i) } \int \mathbf{z} f(\mathbf{x})\, d\mathbf{x} = \mathbf{0}, \text{ (ii) } \int f^2(\mathbf{x})\, d\mathbf{x} = \eta^2.\end{aligned} \tag{2.6}$$

We have equality in (ii) because if f satisfies (i), but $\|f\| = c\eta < \eta$, then $c^{-1}f$ satisfies (i) and (ii), and has $B(c^{-1}f,\xi) = c^{-2}B(f,\xi) > B(f,\xi)$.

We now restrict to absolutely continuous ξ, with density $\xi' = m$. In practice, such ξ would be approximated by discrete measures. Let $\mathscr{F}_1$ be the convex set $\{f \mid \|f\| \leqslant \eta\}$, and for $f_0, f_1 \in \mathscr{F}_1$ and $\lambda \in [0,1]$ put $f_\lambda = (1-\lambda) f_0 + \lambda f_1$. Let $\mathbf{d} \in \mathbb{R}^N$, $d_0 \in \mathbb{R}$ be Lagrange multipliers, and put

$$\phi(\lambda) = B(f_\lambda, \xi) + \mathbf{d}^T \int \mathbf{z} f_\lambda(\mathbf{x})\, d\mathbf{x} + d_0 \int f_\lambda^2(\mathbf{x})\, d\mathbf{x}.$$

A maximizing f_0 must then satisfy, for each f_1,

$$0 \geqslant \phi'(0) = \int (f_1 - f_0)(\mathbf{x}) \left[2\mathbf{b}^T\mathbf{H}^{-1}\mathbf{z} m(\mathbf{x}) + \mathbf{d}^T\mathbf{z} + 2 d_0 f_0(\mathbf{x}) \right] d\mathbf{x}.$$

Here, $\mathbf{b} = \mathbf{b}(f_0)$. The above inequality suggests that

$$f_0(\mathbf{x}) = m(\mathbf{x}) \mathbf{z}^T \mathbf{H}^{-1} \boldsymbol{\beta} + \mathbf{z}^T \mathbf{c}$$

for some N-vectors $\boldsymbol{\beta}$ [proportional to $\mathbf{b}(f_0)$] and $\mathbf{c}$. Adjusting $\mathbf{c}$ to satisfy (2.6)(i) gives

$$f_0(\mathbf{x}) = \mathbf{z}^T\left[m(\mathbf{x})\mathbf{H}^{-1} - \mathbf{B}^{-1}\right]\boldsymbol{\beta}. \tag{2.7}$$

To see that there in fact exists a bias-maximizing f_0, and that it has the form (2.7) for *some* $\boldsymbol{\beta}$, first define functions

$$h(\mathbf{x}, \boldsymbol{\beta}) = s\mathbf{z}^T\left[m(\mathbf{x})\mathbf{H}^{-1} - \mathbf{B}^{-1}\right]\boldsymbol{\beta},$$

where $s > 0$ is chosen so that

$$\|h(\cdot, \boldsymbol{\beta})\| = \eta. \tag{2.8}$$

Then $h(\mathbf{x}, \boldsymbol{\beta})$ satisfies (2.6)(i), (ii) for any $\boldsymbol{\beta}$.

For a fixed but arbitrary f satisfying (2.6)(i), (ii), consider $h(\mathbf{x}, \mathbf{b}(f))$. We claim that

$$B(f, \xi) \leqslant B(h(\cdot, \mathbf{b}(f)), \xi), \tag{2.9}$$

so that we may restrict the search for a maximizing f_0 to those functions of the form (2.7). Furthermore, we will show that equality holds in (2.9) if and only if

$$f(\mathbf{x}) = h(\mathbf{x}, \mathbf{b}(f)) \qquad \text{a.e. } \mathbf{x}. \tag{2.10}$$

From (2.8), and with

$$\mathbf{K} := \int \mathbf{z}\mathbf{z}^T m^2(\mathbf{x})\, d\mathbf{x},$$

we have

$$s^2 = \frac{\eta^2}{\mathbf{b}^T(f)[\mathbf{H}^{-1}\mathbf{K}\mathbf{H}^{-1} - \mathbf{H}^{-1}]\mathbf{b}(f)}. \tag{2.11}$$

Upon replacing f by $h(\cdot, \mathbf{b}(f))$ in (2.4) we find

$$\mathbf{b}(h(\cdot, \mathbf{b}(f))) = s[\mathbf{K}\mathbf{H}^{-1} - \mathbf{I}]\mathbf{b}(f), \tag{2.12}$$

whence

$$B(h(\cdot,\mathbf{b}(f)),\xi)=s^2\mathbf{b}^T(f)[\mathbf{H}^{-1}\mathbf{K}-\mathbf{I}]\mathbf{H}^{-1}[\mathbf{K}\mathbf{H}^{-1}-\mathbf{I}]\mathbf{b}(f). \quad (2.13)$$

Note that, since f satisfies (2.6)(i) and (ii),

$$s^2\left[\mathbf{b}^T(f)\mathbf{H}^{-1}\mathbf{b}(f)\right]^2=\left(\int f(\mathbf{x})h(\mathbf{x},\mathbf{b}(f))\,d\mathbf{x}\right)^2$$

$$\leqslant \int f^2(x)\,d\mathbf{x}\int h^2(\mathbf{x},\mathbf{b}(f))\,d\mathbf{x}=\eta^4. \quad (2.14)$$

Let $\mathbf{H}^{-1/2}$ be a symmetric, p.d. root of $\mathbf{H}^{-1}$, and put

$$\mathbf{l}=\mathbf{H}^{-1/2}\mathbf{b}(f), \qquad \mathbf{J}=\mathbf{H}^{-1/2}\mathbf{K}\mathbf{H}^{-1/2}-\mathbf{I}.$$

Represent $\mathbf{J}$ as $\mathbf{J}=\mathbf{Q}\Lambda\mathbf{Q}^T$, where $\mathbf{Q}$ is orthogonal and Λ is diagonal. Define

$$\mathbf{u}=\mathbf{Q}^T\mathbf{l}/(\mathbf{l}^T\mathbf{l})^{1/2}.$$

Note that $\mathbf{u}^T\mathbf{u}=1$, and put $\bar{\lambda}=\Sigma_{i=1}^N\lambda_i u_i^2$, where $\lambda_i=\Lambda_{ii}$. In this notation, we have

$$B(h(\cdot,\mathbf{b}(f)),\xi)=s^2\mathbf{l}^T\mathbf{J}^2\mathbf{l}, \qquad B(f,\xi)=\mathbf{l}^T\mathbf{l}. \quad (2.15)$$

From (2.11) and (2.14),

$$s^2=\eta^2/\mathbf{l}^T\mathbf{J}\mathbf{l} \quad \text{and} \quad \eta^2\geqslant s\mathbf{l}^T\mathbf{l}, \quad (2.16)$$

so that

$$\eta^2\geqslant\frac{(\mathbf{l}^T\mathbf{l})^2}{\mathbf{l}^T\mathbf{J}\mathbf{l}}=\frac{\mathbf{l}^T\mathbf{l}}{\mathbf{u}^T\Lambda\mathbf{u}}. \quad (2.17)$$

Now (2.15)–(2.17) give

$$\frac{B(h(\cdot,\mathbf{b}(f)),\xi)}{B(f,\xi)}-1=\frac{\eta^2}{\mathbf{l}^T\mathbf{l}}\cdot\frac{\mathbf{u}^T\Lambda^2\mathbf{u}}{\mathbf{u}^T\Lambda\mathbf{u}}-1$$

$$\geqslant\frac{\mathbf{u}^T\Lambda^2\mathbf{u}}{(\mathbf{u}^T\Lambda\mathbf{u})^2}-1=\frac{\Sigma_{i=1}^N(\lambda_i-\bar{\lambda})^2u_i^2}{\bar{\lambda}^2}\geqslant 0. \quad (2.18)$$

This proves (2.9). Equality in (2.18) requires equality in (2.14), and hence that (2.10) hold. Clearly, (2.10) implies equality in (2.18). Thus, equality in (2.9) is equivalent to (2.10).

Rather than investigate solutions to (2.10), it is now simpler to determine an f_0 of the form (2.7), with $\boldsymbol{\beta}$ chosen to maximize

$$B(f_0, \xi) = \boldsymbol{\beta}^T(\mathbf{K}\mathbf{H}^{-1} - \mathbf{I})^T \mathbf{H}^{-1}(\mathbf{K}\mathbf{H}^{-1} - \mathbf{I})\boldsymbol{\beta},$$

subject to

$$\int f_0^2(\mathbf{x})\, d\mathbf{x} = \boldsymbol{\beta}^T(\mathbf{H}^{-1}\mathbf{K}\mathbf{H}^{-1} - \mathbf{H}^{-1})\boldsymbol{\beta} = \eta^2. \tag{2.19}$$

Equivalently, with

$$\boldsymbol{\delta} = \mathbf{K}^{-1}\boldsymbol{\beta}, \qquad \mathbf{G} = (\mathbf{K}\mathbf{H}^{-1} - \mathbf{I})\mathbf{K}\mathbf{H}^{-1}\mathbf{K}, \qquad \mathbf{F} = \mathbf{G}(\mathbf{H}^{-1}\mathbf{K} - \mathbf{I}),$$

we maximize

$$B(f_0, \xi) = \boldsymbol{\delta}^T \mathbf{F} \boldsymbol{\delta} \tag{2.20}$$

subject to

$$\boldsymbol{\delta}^T \mathbf{G} \boldsymbol{\delta} = \eta^2. \tag{2.21}$$

This is a standard eigenvalue problem. The solution is obtained by finding the largest solution μ_ξ to $|\mathbf{F} - \mu\mathbf{G}| = 0$, and then choosing $\boldsymbol{\delta}$ to satisfy

$$(\mathbf{F} - \mu_\xi \mathbf{G})\boldsymbol{\delta} = \mathbf{0},$$

normalized to satisfy (2.21). In terms of $\mathbf{H}$ and $\mathbf{K}$, $\nu_\xi = 1 + \mu_\xi$ is the largest solution to

$$|\mathbf{K} - \nu\mathbf{H}| = 0, \tag{2.22}$$

and

$$(\mathbf{K} - \nu_\xi \mathbf{H})\boldsymbol{\delta} = \mathbf{0}. \tag{2.23}$$

From (2.7),

$$f_0(\mathbf{x}) = \nu_\xi \mathbf{z}^T \left[m(\mathbf{x})\mathbf{I} - \mathbf{A}^{-1}\mathbf{B} \right] \boldsymbol{\delta}. \tag{2.24}$$

The maximum bias is $B(f_0, \xi) = \mu_\xi \eta^2$, so that

$$Q(f_0, \xi) = \nu_\xi \eta^2 + \frac{\sigma^2}{n} \operatorname{tr} \mathbf{A}\mathbf{B}^{-1}. \tag{2.25}$$

3. SPECIAL CASES

In this section we consider two particular types of response surfaces, as described in Section 1. We find, by variational methods, design densities $m_0(\mathbf{x})$ minimizing (2.25), hence satisfying the saddlepoint property (1.4).

3.1. *Response Surface a Plane*

Take $\mathcal{G}$ to be the class of functions of the form

$$g(\mathbf{x}) = \alpha_0 + \sum_{j=1}^{p} \alpha_j x_j,$$

so that $\mathbf{x} = (x_1, x_2, \ldots, x_p)^T$, $\mathbf{z} = (1, \mathbf{x}^T)^T$. We restrict to densities $m(\mathbf{x})$ which are symmetric in each variable and in which the variables are exchangeable. In view of (3.3) below, the most appropriate choice of R is that of a sphere of unit volume:

$$R = \left\{ \mathbf{x} \middle| \|\mathbf{x}\| \leqslant r_p := \frac{[\Gamma(p/2+1)]^{1/p}}{\sqrt{\pi}} \right\}.$$

See Section 4 of Box and Draper (1959) for a discussion of this point.

We find

$$\mathbf{A} = 1 \oplus \gamma_0 \mathbf{I}_p, \qquad \mathbf{B} = 1 \oplus \gamma \mathbf{I}_p,$$

where

$$\gamma_0 = \int x_1^2 \, d\mathbf{x} = \frac{r_p^2}{p+2}, \qquad \gamma = \int x_1^2 m(\mathbf{x}) \, d\mathbf{x},$$

$$\mathbf{K} = \int m^2(\mathbf{x}) \, d\mathbf{x} \oplus \left(\int x_1^2 m^2(\mathbf{x}) \, d\mathbf{x} \right) \mathbf{I}_p,$$

$$\nu_\xi = \max\left[\int m^2(\mathbf{x}) \, d\mathbf{x}, \frac{\gamma_0}{\gamma^2} \int x_1^2 m^2(\mathbf{x}) \, d\mathbf{x} \right].$$

For the solutions below we have

$$\nu_\xi = \int m^2(\mathbf{x}) \, d\mathbf{x}, \tag{3.1}$$

whence

$$Q(f_0, \xi) = \eta^2 \int m^2(\mathbf{x}) \, d\mathbf{x} + \frac{\sigma^2}{n} \left(1 + \frac{p\gamma_0}{\gamma} \right). \tag{3.2}$$

From (2.24),

$$f_0(\mathbf{x}) \propto m_0(\mathbf{x}) - 1.$$

We first minimize $Q(f_0, \xi)$ for fixed γ, then minimize over γ. At the first stage, for some multipliers a, b we minimize

$$\int \left[m^2(\mathbf{x}) - 2a|\mathbf{x}|^2 m(\mathbf{x}) - 2bm(\mathbf{x}) \right] d\mathbf{x}$$

by minimizing the integrand pointwise. We find that

$$m(\mathbf{x}) = \left(a|\mathbf{x}|^2 + b \right)^+ \tag{3.3}$$

with a, b determined by

$$\int m(\mathbf{x}) \, d\mathbf{x} = 1, \qquad \int |\mathbf{x}|^2 m(\mathbf{x}) \, d\mathbf{x} = p\gamma. \tag{3.4}$$

For $\gamma \geqslant \gamma_0$ the minimax design assumes two forms, depending upon the sign of b. In each case, we leave it to the reader to verify (3.1).

Case 1: $a, b > 0$. This form is valid for

$$1 \leqslant \frac{\gamma}{\gamma_0} \leqslant \frac{(p+2)^2}{p(p+4)},$$

corresponding to small values of σ^2/n. The design is most conveniently described in terms of γ. Put

$$\kappa^2 = \sigma^2 \left\{ 2n \left(\frac{\gamma}{\gamma_0} \right)^2 \left(\frac{\gamma}{\gamma_0} - 1 \right) \right\}^{-1}.$$

Upon solving for a, b from (3.4) and then minimizing (3.2) over γ, we find that ξ_0 has the density

$$m_0(\mathbf{x}) = 1 + \left(\frac{p+4}{4} \right) \left(\frac{\gamma}{\gamma_0} - 1 \right) \left(\frac{|\mathbf{x}|^2}{\gamma_0} - p \right), \qquad |\mathbf{x}| \leqslant r_p,$$

and

$$f_0(\mathbf{x}) = \frac{\kappa}{\sqrt{p}} \left(\frac{|\mathbf{x}|^2}{\gamma_0} - p \right),$$

$$\eta = \frac{2\kappa}{\sqrt{p+4}}.$$

Note that $\gamma = \gamma_0$ corresponds to the uniform design $m(\mathbf{x}) \equiv 1$, which becomes minimax as $n \to \infty$.

Case 2: $a > 0$, $b < 0$. Define $c \in [0, 1)$ by $c^2 r_p^2 = -b/a$. Set

$$J_p(c) = \frac{p(1-c^2) - 2c^2(1-c^p)}{p+2}.$$

Proceeding as in case 1, we find that for

$$\frac{(p+2)^2}{p(p+4)} \leqslant \frac{\gamma}{\gamma_0} \leqslant \frac{p+2}{p}$$

the solution is given by

$$m_0(\mathbf{x}) = \frac{1}{J_p(c)}\left(\left(\frac{|\mathbf{x}|}{r_p}\right)^2 - c^2\right)^+,$$

$$\frac{\gamma}{\gamma_0} = \frac{J_{p+2}(c)}{J_p(c)},$$

$$f_0(\mathbf{x}) = \kappa\left[m_0(\mathbf{x}) - 1\right],$$

where

$$\kappa^2 = \frac{(p+2)J_p^3(c)}{2SJ_{p+2}^2(c)} \cdot \frac{\sigma^2}{n}, \qquad \eta^2 = S\kappa^2$$

and

$$S = \frac{pJ_{p+2}(c)}{(p+2)J_p^2(c)} - \frac{c^2}{J_p(c)} - 1.$$

The limiting case $\gamma = \gamma_0(p+2)/p = r_p^2/p$, $c = 1$ corresponds to the design with all mass at $|\mathbf{x}| = r_p$.

3.2. *Two Interacting Regressors*

We here take $\mathscr{G}$ to be the class of functions of the form

$$g(x_1, x_2) = \alpha_1 + \alpha_2 x_1 + \alpha_3 x_2 + \alpha_4 x_1 x_2.$$

on $R = [-\frac{1}{2}, \frac{1}{2}] \times [-\frac{1}{2}, \frac{1}{2}]$. Thus $\mathbf{x} = (x_1, x_2)^T$, $\mathbf{z} = (1, x_1, x_2, x_1 x_2)^T$. We again restrict to symmetric, exchangeable m and find

$$\mathbf{A} = \operatorname{diag}\left(1, \tfrac{1}{12}, \tfrac{1}{12}, \tfrac{1}{144}\right), \qquad \mathbf{B} = \operatorname{diag}(1, \gamma, \gamma, \gamma_{12}),$$

where

$$\gamma = \int x_1^2 m(\mathbf{x})\, d\mathbf{x}, \qquad \gamma_{12} = \int x_1^2 x_2^2 m(\mathbf{x})\, d\mathbf{x};$$

$$\mathbf{K} = \operatorname{diag}\left(\int m^2(\mathbf{x})\, d\mathbf{x}, \int x_1^2 m^2(\mathbf{x})\, d\mathbf{x}, \int x_1^2 m^2(\mathbf{x})\, d\mathbf{x}, \int x_1^2 x_2^2 m^2(\mathbf{x})\, d\mathbf{x} \right),$$

$$\nu_\xi = \max\left[\int m^2(\mathbf{x})\, d\mathbf{x}, \frac{\int x_1^2 m^2(\mathbf{x})\, d\mathbf{x}}{12\gamma^2}, \frac{\int x_1^2 x_2^2 m^2(\mathbf{x})\, d\mathbf{x}}{144\gamma_{12}^2} \right].$$

For the range in (3.8) below, ν_ξ is again as at (3.1), whence

$$Q(f_0, \xi) = \eta^2 \int m^2(\mathbf{x})\, d\mathbf{x} + \frac{\sigma^2}{n}\left(1 + \frac{1}{6\gamma} + \frac{1}{144\gamma_{12}}\right). \tag{3.5}$$

Minimizing Q for fixed γ, γ_{12} gives

$$m_0(x_1, x_2) = \left(\lambda + \mu\left(x_1^2 + x_2^2\right) + \delta x_1^2 x_2^2\right)^+ \tag{3.6}$$

with λ, μ, δ determined by

$$\int m_0(\mathbf{x}) = 1, \qquad \int x_1^2 m_0(\mathbf{x})\, d\mathbf{x} = \gamma, \qquad \int x_1^2 x_2^2 m_0(\mathbf{x})\, d\mathbf{x} = \gamma_{12}. \tag{3.7}$$

We give here the details of that form of the solution which is valid for small values of $\sigma^2/n\eta^2$, corresponding to

$$\tfrac{1}{12} \leqslant \gamma \leqslant \tfrac{7}{60}. \tag{3.8}$$

In this case γ, μ, δ are nonnegative; solving for them from (3.7) gives

$$\lambda = \tfrac{9}{16}(400\gamma_{12} - 120\gamma + 9),$$

$$\mu = \tfrac{45}{4}(-240\gamma_{12} + 56\gamma - 3),$$

$$\delta = 225(144\gamma_{12} - 24\gamma + 1).$$

Upon minimizing (3.5) over γ, γ_{12} we find that we can express first γ_{12}, then $\sigma^2/n\eta^2$ in terms of γ. Specifically, determine γ_{12} from

$$34{,}560\gamma_{12}^3 + 240(1 - 24\gamma)\gamma_{12}^2 + 240\gamma^2\gamma_{12} + \gamma^2(3 - 56\gamma) = 0,$$

TABLE 1

VALUES OF THE CONSTANTS FOR THE DESIGN (3.6)

γ	γ_{12}	λ	μ	δ	$\sigma^2/n\eta^2$
$\frac{15}{180}$	$\frac{1}{144}$	1	0	0	0
$\frac{16}{180}$	.0080	.8716	0.5409	5.5091	0.1026
$\frac{17}{180}$	.0091	.7379	1.1455	10.2538	0.2452
$\frac{18}{180}$	.0102	.6004	1.7947	14.4636	0.4307
$\frac{19}{180}$	.0112	.4603	2.4761	18.2869	0.6621
$\frac{20}{180}$	.0122	.3182	3.1816	21.8213	0.9427
$\frac{21}{180}$	.0133	.1745	3.9055	25.134	1.2758

using the root which satisfies

$$\tfrac{1}{6}\gamma - \tfrac{1}{144} \leqslant \gamma_{12} \leqslant \tfrac{7}{30}\gamma - \tfrac{1}{80}$$

to ensure $\lambda, \mu, \delta \geqslant 0$. Now put

$$\sigma^2/n\eta^2 = 270\gamma^2(56\gamma - 240\gamma_{12} - 3).$$

Then $m_0(\mathbf{x})$ is given by (3.6), and

$$f_0(\mathbf{x}) = \eta \frac{m_0(\mathbf{x}) - 1}{(\lambda + 2\mu\gamma + \delta\gamma_{12} - 1)^{1/2}}.$$

Note that $\gamma = \frac{1}{12}$ corresponds to the uniform design, minimax as $n \to \infty$.
Some relevant numbers are given in Table 1.

REFERENCES

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.

Huber, P. 1975. Robustness and designs, in *A Survey of Statistical Design and Linear Models* (Srivastava, J. N., Ed.), North Holland, Amsterdam, pp. 287–303.

Marcus, M. B. and Sacks, J. 1976. Robust designs for regression problems, in *Statistical Decision Theory and Related Topics II* (S. S. Gupta and D. S. Moore, Eds.), Academic, New York, pp. 245–268.

Received 23 September 1988; final manuscript accepted 29 June 1989