

Minimax designs for approximately linear models with AR(1) errors

Douglas P. WIENS and Julie ZHOU

Key words and phrases: Approximately linear regression response; autoregressive errors; integrated mean squared error; minimax; nearest neighbour; regression design.

AMS 1991 subject classification: Primary 62K05, 62F35; secondary 62J05, 62M10.

ABSTRACT

We obtain designs for linear regression models under two main departures from the classical assumptions: (1) the response is taken to be only *approximately* linear, and (2) the errors are not assumed to be independent, but to instead follow a first-order autoregressive process. These designs have the property that they minimize (a modification of) the maximum integrated mean squared error of the estimated response, with the maximum taken over a class of departures from strict linearity and over all autoregression parameters ρ , $|\rho| < 1$, of fixed sign. Specific methods of implementation are discussed. We find that an asymptotically optimal procedure for AR(1) models consists of choosing points from that design measure which is optimal for uncorrelated errors, and then implementing them in an appropriate order.

RÉSUMÉ

Les auteurs abordent le problème de la planification de la cueillette de données dans le cadre de modèles de régression où la variable endogène s'exprime comme une fonction *approximativement* linéaire des variables exogènes et où les erreurs ne sont pas supposées indépendantes mais obéissent plutôt à un processus autorégressif du premier ordre. Les plans proposés minimisent une version de l'erreur quadratique moyenne intégrée maximale de la variable réponse estimée, le maximum étant pris sur une classe d'écarts au postulat de linéarité stricte et sur l'ensemble des valeurs possibles du paramètre autorégressif ρ , $|\rho| < 1$, de signe donné. Les mécanismes d'implantation de ces plans minimax sont présentés. Il appert qu'en choisissant les valeurs des variables hétérogènes selon un schéma qui est optimal dans le cas où les erreurs sont non corrélées, on obtient une procédure asymptotiquement optimale pour les modèles AR(1), dans la mesure où l'ordre dans lequel les observations sont alors recueillies est judicieusement choisi.

1. INTRODUCTION

In this paper we study optimal designs for linear models under two main departures from the classical assumptions:

- (1) the response is taken to be only *approximately* linear, and
- (2) the errors are not assumed to be independent, but to instead follow a first-order autoregressive (AR(1)) process.

We consider the following approximately linear model with AR(1) errors. We suppose that an experimenter is to observe a random variable Y at locations $\mathbf{x}_1, \dots, \mathbf{x}_n$ in a q -dimensional design space \mathcal{S} . The index i of \mathbf{x}_i may be taken to represent time. The response $E(Y|\mathbf{x}) \approx \theta_0 + \theta'_1 \mathbf{x}$ is thought to be approximately linear and is observed subject to additive and possibly autoregressive errors, with mean zero. With $f(\mathbf{x}) := \sqrt{n} \{E(Y|\mathbf{x}) - \theta_0 - \theta'_1 \mathbf{x}\}$ the observations then satisfy

$$Y_i = \theta_0 + \theta'_1 \mathbf{x}_i + n^{-\frac{1}{2}} f(\mathbf{x}_i) + \varepsilon_i, \quad i = 1, \dots, n. \quad (1)$$

The covariance matrix of the errors is that of an AR(1) process:

$$\text{cov}(\varepsilon) = \frac{\sigma^2}{1 - \rho^2} P, \quad P_{i,j} = \rho^{|i-j|}, \quad 0 \leq |\rho| < 1.$$

The estimate $\hat{\theta}$ is taken to be the *best linear unbiased estimate* (BLUE), computed using a consistent estimate of ρ . Here, both *best* and *unbiased* refer only to properties of the estimator in an exactly linear model with a correctly specified autocorrelation process. The experimenter intends to fit a linear surface, and anticipates AR(1) errors. He desires a design that will afford a measure of protection against the consequent increases in bias should the assumption of exact linearity not be realized, as well as protection against errors due to variation and to autocorrelation.

In the classical linear model (with $f \equiv 0$ and a fixed autocorrelation structure in (1)), the general experimental design problem is to choose n design points $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ in such a way that a certain loss function is minimized. This loss is a scalar-valued function of the covariance matrix of $\hat{\theta}$, such as the trace (A-optimality), the determinant (D-optimality) or the largest eigenvalue (E-optimality). If the errors are uncorrelated, $\hat{\theta}$ is typically the ordinary least-squares estimator. If the errors are correlated with a known correlation matrix, then generalized least squares yields the BLUE.

The study of experimental designs for finite sample sizes often leads to intractable integer optimization problem. For a discussion see Pukelsheim (1993). A remedy is to seek *asymptotically optimal* designs, characterized by their limiting design measures. To implement such designs, one selects points $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ from the optimal distribution in such a way that the empirical distributions converge to the optimal distribution as $n \rightarrow \infty$.

In Wiens & Zhou (1996), minimax designs were studied for the approximately linear model with errors obeying a very general model of dependence. A main result of that paper is that a design distribution which is asymptotically (minimax) optimal for uncorrelated errors retains its optimality under autocorrelation if the design points are *randomly sampled*, or obtained by *randomly permuting* a fixed design, from this distribution. By random sampling in this context we mean that each design point is randomly and independently selected according to the optimal distribution. A randomly permuted fixed design is one in which design points are first selected in a systematic manner from the optimal distribution. For instance, in the one-dimensional case one can choose the quantiles $x_i = G^{-1}\{(i-1)/(n-1)\}$. The points so chosen are then randomly permuted to determine the order in which they are implemented, i.e., the i^{th} point to be implemented is $x_{\pi(i)}$, where π is a random permutation of $(1, 2, \dots, n)$.

In the aforementioned paper the error process was allowed to vary over a neighbourhood of the uncorrelated error process and was otherwise assumed only to be weakly stationary. If additional information is available, one can improve on the designs described above.

Example 1. Suppose that the design space is a sphere centered at the origin, and that an AR(1) model holds. We show in this paper that if consecutive observations are *negatively* correlated, then an asymptotically optimal procedure is to first choose a collection of design points from a design optimal for uncorrelated errors, and to then order the points in such a way that each point is the nearest neighbour, among the points not yet used, of the preceding design point. If the first, third, fifth, \dots points so obtained are then multiplied by -1 , the resulting sequence forms an asymptotically optimal design for *positively* correlated errors. This is made precise in Section 3, where other methods of implementation are also discussed.

We define the “true” values of the regression parameters to be those which minimize the $L^2(S)$ norm of f . This implies that f is $L^2(S)$ -orthogonal to the regressors $(1, \mathbf{x})$. Our asymptotic approach requires further regularity of f , hence we take $f \in \mathcal{F}$, where

$$\mathcal{F} = \left\{ f \mid f \text{ is bounded and a.e. continuous, } \int_S f^2(\mathbf{x}) d\mathbf{x} \leq \eta^2, \int_S \begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix} f(\mathbf{x}) d\mathbf{x} = 0 \right\}. \quad (2)$$

The bound η^2 on $\int_S f^2(\mathbf{x}) d\mathbf{x}$ is not required to be known to the experimenter, although *some* bound is necessary in order that errors due to variance and to bias remain of the same order of

magnitude in finite samples and asymptotically, thus keeping the model discrimination problem from degenerating. An implication (see (1)) is that the effect of f on the regression response decreases at the rate $n^{-1/2}$. This is required for the asymptotics, and is a situation quite analogous to that in the theory of hypothesis testing, where one typically considers contiguous alternatives tending to the null hypothesis at the rate $n^{-1/2}$. The reason is the same in both situations – bias and variance must decrease at the same rate if the limit is not to be overwhelmed by the contribution of one of them.

Our results depend on η^2 and on the noise variance σ^2 only through the ratio $\nu := \sigma^2/\eta^2$. Thus for any particular value of n the experimenter may choose ν according to his judgement of the relative importance of variance versus bias. The limiting case $\nu = 0$ corresponds to a “pure bias” problem, in which only the effect of f is of concern. In this case optimal designs are uniform over the design space. As $\nu \rightarrow \infty$ variance becomes completely dominant, so that in the limit one has the classical design problem, whose solution places all design points at the extremes of the design space.

The assumptions of boundedness and continuity of f are required in order that the loss function converge as the design converges to its limiting value. We note that the orthogonality of f to the regressors also ensures the identifiability of the parameter vector $\theta = (\theta_0, \theta'_1)'$.

As design space we shall take a q -dimensional sphere of unit volume centered at the origin, hence with radius $r := \{\Gamma(1+q/2)\}^{1/q}/\sqrt{\pi}$. Other choices of \mathcal{S} are possible, but *some* structure must be assumed. The assumptions that \mathcal{S} is centered and normalized are made without loss of generality – each corresponds merely to a linear transformation of \mathbf{x} . If the design space is a sphere of radius s , centered at \mathbf{a} , then design points obtained from this paper are to be subjected to the transformation $\mathbf{x} \rightarrow \mathbf{a} + (s/r) \cdot \mathbf{x}$. Of course an analogous transformation holds for an elliptical design space.

The purpose of this paper is to derive minimax designs for the model described by (1) and (2) with loss taken to be (a modification of) the integrated mean squared error (IMSE). This is analogous to the classical Q -optimality criterion (Fedorov 1972, p. 142), also termed I -optimality by Studden (1977). The designs obtained here minimize the supremum of the loss asymptotically. The supremum is taken over \mathcal{F} and over all autoregression parameters ρ , $|\rho| < 1$, of fixed sign. The latter condition allows for optimal experimentation when there is no knowledge of ρ apart from the sign. It should be possible to obtain minimax designs in which ρ is restricted to lie in a particular sub-interval of $(-1, 1)$, but this is something we have not yet investigated.

The parameter ρ may be estimated from the data through, e.g., the Cochran-Orcutt procedure (Montgomery & Peck 1982, p. 355). We note however that in order to implement the designs constructed here, only the *sign* of ρ need be known. This sign is often easily determined by the nature of the experiment. For instance, in an experiment to investigate crop yield per plant (y) relating to density of planting (x) on several plots, suppose that the i^{th} plot is to be assigned density x_i . The effect of neighbouring plots will induce a positive correlation ($\rho > 0$) due to common factors such as temperature and soil characteristics, not all of which can be controlled. If the sign of ρ (or the specific autocorrelation model) cannot be determined, one might adopt a more conservative approach as in Wiens & Zhou (1996).

Example 2. Box, Hunter & Hunter (1978, Section 18.4) describe a feedback control experiment in which polymer is coloured by adding dye at the inlet of a continuous reactor. The purpose is to adjust the rate x_i at which dye is added at time i , in order to maintain the colour index y_i at a constant target value. An AR(1) model is posited and justified on empirical grounds. It is assumed that the increase in x required to eventually increase the colour index by 1 unit is *known*; the manner in which it might become known is not explained. Suppose then that this parameter is not known, and consider the related problem of its estimation, i.e., the estimation of θ_1 in $E(y_i | x_i) \simeq \theta_0 + \theta_1 x_i$. It seems reasonable to assume an approximately linear model for which we desire an estimate of the slope parameter which is robust against response nonlinearity and sensitive to the effects of autocorrelation. The designs detailed in this paper are appropriate

for such a situation. The carry-over effects of previous additions of dye result in a positive lag-1 autocorrelation. Then if the desired number n of observations is odd, if $\nu = 6.48$ (see Remark 2(b) of Section 2), if a and b are respectively the midrange and the range of x and if $z_i = (x_i - a)/b$, the design resulting from the considerations in Section 3.1 below is given as follows. With

$$t_i = \left\{ \frac{i-1}{4(n-1)} - \frac{1}{8} \right\}^{\frac{1}{3}}, \quad i = 1, \dots, n,$$

the ordered design points are

$$\langle z_1, z_2, \dots, z_n \rangle = \langle t_n, t_2, t_{n-2}, t_4, t_{n-4}, t_6, \dots, t_{(n+1)/2} = 0, \dots, t_{n-3}, t_3, t_{n-1}, t_1 \rangle.$$

If n is even, the point at 0 is omitted.

The present paper represents part of an ongoing attempt to combine the theory of designs for models which are exactly linear but which have correlated errors with the theory of designs for models which have uncorrelated errors but also have possible departures as at (1) and (2). References to related work, and other approaches to this synthesis, may be found in Wiens & Zhou (1996, 1997). See in particular Bickel & Herzberg (1979) and Herzberg & Huda (1981).

In Section 2 we obtain the asymptotic value of the loss function, define the notion of asymptotic optimality, and give properties required of an asymptotically optimal sequence of designs. In Section 3 we construct explicit examples of asymptotically optimal sequences for $q = 1$, $q = 2$ and $q = 6$, and outline an algorithm by which such sequences may be constructed for general values of q . The results of a Monte Carlo study are presented to show the gains, over some common competitors, which may be enjoyed by the use of these designs.

2. ASYMPTOTIC OPTIMALITY

Let \mathbf{X} be the $n \times (q+1)$ model matrix with i^{th} row $(1, \mathbf{x}_i')$, and define $\mathbf{y} = (y_1, \dots, y_n)'$. Assume that ρ is consistently estimated so that, asymptotically, $\hat{\boldsymbol{\theta}} = (\mathbf{X}'\mathbf{P}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{P}^{-1}\mathbf{y}$, where

$$\mathbf{P}^{-1} = \frac{1}{1-\rho^2} \begin{pmatrix} 1 & -\rho & 0 & \cdots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \cdots & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{pmatrix}.$$

With $\mathbf{f} = (f(\mathbf{x}_1), \dots, f(\mathbf{x}_n))'$ the mean-squared-error matrix is

$$\begin{aligned} \mathbf{MSE}_n &= E[(\sqrt{n}\hat{\boldsymbol{\theta}} - \sqrt{n}\boldsymbol{\theta})(\sqrt{n}\hat{\boldsymbol{\theta}} - \sqrt{n}\boldsymbol{\theta})'] \\ &= n \left(\frac{\mathbf{X}'\mathbf{P}^{-1}\mathbf{X}}{n} \right)^{-1} \frac{\mathbf{X}'\mathbf{P}^{-1}\mathbf{f}}{n} \frac{\mathbf{f}'\mathbf{P}^{-1}\mathbf{X}}{n} \left(\frac{\mathbf{X}'\mathbf{P}^{-1}\mathbf{X}}{n} \right)^{-1} + \frac{\sigma^2}{1-\rho^2} \left(\frac{\mathbf{X}'\mathbf{P}^{-1}\mathbf{X}}{n} \right)^{-1}. \end{aligned} \quad (3)$$

Define $\gamma_0 = r^2/(q + 2)$ and $\mathbf{A} = \int_{\mathcal{S}} \begin{pmatrix} 1 \\ \mathbf{x}' \end{pmatrix} (1 \ \mathbf{x}') d\mathbf{x} = 1 \oplus \gamma_0 \mathbf{I}_q$. Then the normalized integrated mean squared error, of which we seek the limit as $n \rightarrow \infty$, is

$$\begin{aligned} \text{IMSE}_n &= n \int_{\mathcal{S}} E[\{\hat{\theta}_0 + \hat{\boldsymbol{\theta}}_1' \mathbf{x} - E(y \mid \mathbf{x})\}^2] d\mathbf{x} \\ &= \int_{\mathcal{S}} (1, \mathbf{x}') \mathbf{MSE}_n(1, \mathbf{x}')' d\mathbf{x} + \int_{\mathcal{S}} f^2(\mathbf{x}) d\mathbf{x} \\ &= \text{tr}(\mathbf{MSE}_n \cdot \mathbf{A}) + \int_{\mathcal{S}} f^2(\mathbf{x}) d\mathbf{x}. \end{aligned} \quad (4)$$

These quantities are more easily handled in terms of the empirical distribution function ξ_n of the consecutive design points $(\mathbf{x}_i, \mathbf{x}_{i+1})$, $i = 1, \dots, n - 1$, viz.

$$\xi_n(\mathbf{u}, \mathbf{v}) = \frac{1}{n-1} \sum_{i=1}^{n-1} \delta_{\{\mathbf{x}_i, \mathbf{x}_{i+1}\}}(\mathbf{u}, \mathbf{v}).$$

Here $\delta_{\{\mathbf{x}_i, \mathbf{x}_{i+1}\}}$ is point mass at $(\mathbf{x}_i, \mathbf{x}_{i+1})$. Denote the marginals of ξ_n by $M_{n,1}(\mathbf{u})$ and $M_{n,2}(\mathbf{v})$ respectively.

To obtain the asymptotic value of (4) we require three assumptions. Assume (A1) ξ_n converges weakly to a probability distribution ξ_0 on $\mathcal{S} \times \mathcal{S}$. Let $M_0(\cdot)$ be the corresponding weak limit of $M_{n,1}(\cdot)$ and $M_{n,2}(\cdot)$ and assume that (A2) all first moments of $M_0(\cdot)$ vanish, and that (A3) $M_0(\{\mathbf{0}\}) < 1$. Assumption (A1) is necessary for an asymptotic description of the designs, (A2) implies that the intercept and slope components of $\hat{\boldsymbol{\theta}}$ are asymptotically uncorrelated, and (A3) ensures that the covariance matrix of $\hat{\boldsymbol{\theta}}$ has a nonsingular limit as $n \rightarrow \infty$.

Define a $q + 1$ -vector by

$$\mathbf{b}(f, \xi_n, \rho) = \frac{1}{1 - \rho^2} \begin{pmatrix} b_1(f, \xi_n, \rho) \\ \mathbf{b}_2(f, \xi_n, \rho) \end{pmatrix},$$

where

$$\begin{aligned} b_1(f, \xi_n, \rho) &= (1 - \rho)^2 E_{\xi_n}\{f(\mathbf{U})\}, \\ \mathbf{b}_2(f, \xi_n, \rho) &= E_{\xi_n}\{(1 + \rho^2)\mathbf{U}f(\mathbf{U}) - \rho\mathbf{U}f(\mathbf{V}) - \rho\mathbf{V}f(\mathbf{U})\}. \end{aligned}$$

We find that

$$\frac{\mathbf{X}'\mathbf{P}^{-1}\mathbf{f}}{n} = \mathbf{b}(f, \xi_n, \rho) + O\left(\frac{r\|\mathbf{f}\|_{\infty}}{n}\right). \quad (5)$$

Similarly,

$$\frac{\mathbf{X}'\mathbf{P}^{-1}\mathbf{X}}{n} = \mathbf{W}(\xi_n, \rho) + O\left(\frac{r^2}{n}\right), \quad (6)$$

where

$$\begin{aligned} \mathbf{W}(\xi_n, \rho) &= \frac{1}{1 - \rho^2} \begin{pmatrix} (1 - \rho)^2 & (1 - \rho)^2 E_{\xi_n}(\mathbf{U}') \\ (1 - \rho)^2 E_{\xi_n}(\mathbf{U}) & (1 + \rho^2)\mathbf{B}(\xi_n) - \rho\mathbf{C}(\xi_n) \end{pmatrix}, \\ \mathbf{B}(\xi_n) &= E_{\xi_n}(\mathbf{U}\mathbf{U}'), \quad \mathbf{C}(\xi_n) = E_{\xi_n}(\mathbf{U}\mathbf{V}' + \mathbf{V}\mathbf{U}'). \end{aligned}$$

The remainders in (5) and (6) arise from the end effects $n^{-1}\mathbf{x}_n\mathbf{x}_n'$, $n^{-1}\mathbf{x}_nf(\mathbf{x}_n)$, etc.

Substituting (3) into (4) and using (5) and (6) yields $\text{IMSE}_n = \text{IMSE}(f, \xi_n, \rho) + O(n^{-1})$, where

$$\begin{aligned} \text{IMSE}(f, \xi_n, \rho) &= \frac{\sigma^2}{1 - \rho^2} \text{tr} \{ \mathbf{W}^{-1}(\xi_n, \rho) \mathbf{A} \} \\ &\quad + \mathbf{b}'(f, \xi_n, \rho) \mathbf{W}^{-1}(\xi_n, \rho) \mathbf{A} \mathbf{W}^{-1}(\xi_n, \rho) \mathbf{b}(f, \xi_n, \rho) + \int_{\mathcal{S}} f^2(\mathbf{x}) d\mathbf{x}. \end{aligned}$$

Assumptions (A1), (A3) and the defining properties of \mathcal{F} ensure that $\mathbf{W}(\xi_n, \rho)$ is nonsingular for sufficiently large n , and that $\mathbf{b}(f, \xi_n, \rho)$ and $\mathbf{W}(\xi_n, \rho)$ converge to $\mathbf{b}(f, \xi_0, \rho)$ and $\mathbf{W}(\xi_0, \rho)$ respectively. The asymptotic value of IMSE_n is then

$$\lim_{n \rightarrow \infty} \text{IMSE}_n = \text{IMSE}(f, \xi_0, \rho). \quad (7)$$

As our loss function we shall take

$$\text{IMSE}'(f, \xi, \rho) = \text{IMSE}(f, \xi, \rho) - \frac{\sigma^2}{(1 - \rho)^2}.$$

The subtraction of the asymptotic variance of the intercept estimate, which cannot be controlled by the design and is in any event typically of less interest than that of the slope estimates, ensures that the supremum of the loss will be finite in the case $\rho \geq 0$, given an appropriate design. If $\rho \leq 0$ then this step is without loss of generality, since in this case both $\text{IMSE}'(f, \xi, \rho)$, when evaluated at the optimal design, and $\sigma^2/(1 - \rho)^2$ are maximized at $\rho = 0$.

Evaluating (7) gives

$$\begin{aligned} \text{IMSE}'(f, \xi, \rho) &= \sigma^2 \gamma_0 \text{tr} \{ [(1 + \rho^2) \mathbf{B}(\xi) - \rho \mathbf{C}(\xi)]^{-1} \} + [\mathbf{E}_M \{ f(\mathbf{U}) \}]^2 + \int_{\mathcal{S}} f^2(\mathbf{x}) d\mathbf{x} \\ &\quad + \gamma_0 \mathbf{b}'_2(f, \xi, \rho) \{ (1 + \rho^2) \mathbf{B}(\xi) - \rho \mathbf{C}(\xi) \}^{-2} \mathbf{b}_2(f, \xi, \rho). \end{aligned} \quad (8)$$

Here we have used (A2), which implies that $\mathbf{W}(\xi_0, \rho)$ is block-diagonal. This simplifying assumption can be done away with, but at the cost of considerable additional complexity. We note that the first summand in (8) arises solely from the covariance matrix of the regression estimates, the second and third summands from their bias and the fourth from the errors in model specification.

We say that a sequence $(\xi_n)_{n=1}^\infty$ is *asymptotically optimal* if the weak limit ξ_0 satisfies

$$\sup_{f, \rho} \text{IMSE}'(f, \xi_0, \rho) = \inf_{\xi} \sup_{f, \rho} \text{IMSE}'(f, \xi, \rho).$$

The suprema are taken over $f \in \mathcal{F}$ and all ρ , $|\rho| < 1$, of fixed sign. The infimum is taken over all distributions satisfying assumptions (A1)–(A3) above.

Our main result is Theorem 1, which is stated below and proven in the appendix. Before stating it we describe the design which is minimax optimal for the model described by (1) and (2), with uncorrelated errors. It turns out that an asymptotically optimal procedure for AR(1) models consists of choosing points from this design and then implementing them in an appropriate order.

The minimax design was derived by Huber (1975) in the case $q = 1$, and by Wiens (1992) for general q . It has a density $m_0(\mathbf{u}) = g_0(\|\mathbf{u}\|)$ on \mathcal{S} , where $g_0(z)$ is defined as follows:

Case 1. If $0 \leq \nu \leq 2(q + 2)^4/q^3(q + 4)^2$ then

$$g_0(z) = 1 + \left(\frac{\gamma}{\gamma_0} - 1 \right) \left(\frac{q + 4}{4} \right) \left(\frac{z^2}{\gamma_0} - q \right),$$

where γ is determined from

$$\nu = \left(\frac{q+4}{2} \right) \left(\frac{\gamma}{\gamma_0} - 1 \right) \left(\frac{\gamma}{\gamma_0} \right)^2.$$

Case 2. If $\nu \geq 2(q+2)^4/(q^3(q+4)^2)$ then $g_0(z) = \{(z/r)^2 - b\}^+/K_q(b)$, where $x^+ := \max(0, x)$, $K_q(b) = (1-b) - 2\{1 - b^{(q/2)+1}\}/(q+2)$, and b, γ are determined by the equations $\gamma = \gamma_0 \cdot K_{q+2}(b)/K_q(b)$, $\nu = 2K_{q+2}^2(b)/((q+2)K_q^3(b))$.

If \mathbf{U} has density $m_0(\mathbf{u})$, then $Z = \|\mathbf{U}\|$ has distribution function $H_0(z)$ ($0 \leq z \leq r$), with density $h_0(z) = (qz^{q-1}/r^q)g_0(z)$. The parameter γ has the interpretation $q\gamma = E_{H_0}(Z^2)$.

THEOREM 1. Fix the sign of ρ . Suppose that the empirical distributions $\xi_n(\mathbf{u}, \mathbf{v})$ of the consecutive design points $(\mathbf{x}_i, \mathbf{x}_{i+1})$, $i = 1, \dots, n-1$, are such that

(B1) The marginal distribution $M_0(\mathbf{u})$ of the weak limit ξ_0 has density $m_0(\mathbf{u})$; and

(B2) $E_{\xi_n}\{\|\mathbf{U} + \text{sign}(\rho)\mathbf{V}\|^2\} \rightarrow 0$ as $n \rightarrow \infty$.

Then the sequence of designs $(\xi_n)_{n=1}^\infty$ is asymptotically optimal.

It is interesting to observe that:

(1) When $\rho < 0$, requirement (B2) of Theorem 1 states that the average squared distance between consecutive design points is to tend to 0. When $\rho > 0$, the average of the squared norms of the pairwise averages of consecutive design points is to tend to 0. These two situations can be linked as follows. First suppose that $\rho < 0$, that (ξ_n) satisfies (B1) and (B2) and that ξ_n has support points $(\mathbf{x}_{ni})_{i=1}^n$. Let (ξ'_n) be the design with support points $\{(-1)^i \mathbf{x}_{ni}\}_{i=1}^n$. Then (ξ'_n) satisfies (B1) by virtue of the symmetry of m_0 , and (B2) holds for $\rho > 0$. Thus if (ξ_n) is asymptotically optimal for $\rho < 0$ then (ξ'_n) is asymptotically optimal for $\rho > 0$.

(2) Some special cases of the density m_0 are:

- (a) When $\nu = 0$ we have $\gamma = \gamma_0$ and $m_0(\mathbf{u}) \equiv 1$, the continuous uniform density on \mathcal{S} . This is a “pure bias” problem, in that as $\nu \rightarrow 0$ the MSE matrix becomes a function of bias alone.
- (b) When $\nu = 2(q+2)^4/\{q^3(q+4)^2\}$, i.e., on the boundary between Cases 1 and 2, we have $b = 0$, $H_0(z) = (z/r)^{q+2}$ for $0 \leq z \leq r$, $\gamma = \gamma_0[1 + \{4/q(q+4)\}]$, and $m_0(\mathbf{u}) = \|\mathbf{u}\|^2/(q\gamma_0)$ for $0 \leq \|\mathbf{u}\| \leq r$. This is the value of ν used in Example 2 of Section 1.
- (c) As $\nu \rightarrow \infty$ – giving a “pure variance” problem – we have $b \rightarrow 1$, $\gamma \rightarrow r^2/q$ and $\|\mathbf{U}\| \rightarrow$ point mass at r .

(3) It is clear from the definition that asymptotically optimal sequences are not unique. In Section 3 below we give some guidelines for their construction.

3. SPECIAL CASES

3.1. $q=1$: Straight-Line Regression.

When $q = 1$, the density $m_0(x)$ of Theorem 1 is of the form

$$m_0(x) = \alpha(x^2 + \beta)^+, \quad |x| \leq \frac{1}{2}. \quad (9)$$

Table 1 gives some representative values of the constants in terms of ν .

For $\rho < 0$ we choose the design points $x_i = M_0^{-1}\{(i-1)/(n-1)\}$. Then (B1) of Theorem 1 is immediate, and (B2) holds as long as

$$\lim_{n \rightarrow \infty} \frac{1}{n-1} \sum_{i=1}^{n-1} |x_i - x_{i+1}|^2 = 0.$$

(10)

But $\sum |x_i - x_{i+1}|^2 \leq \sum |x_i - x_{i+1}| = x_n - x_1 = 1$, and (10) follows. Thus this sequence of designs is asymptotically optimal when $\rho < 0$.

If $\nu = 0$, the optimal density is $m_0(x) \equiv 1$ and $x_i = (i-1)/(n-1) - 1/2$. As $\nu \rightarrow \infty$ the optimal design tends to $M_0(\{\pm 1/2\}) = 1/2$ and the above prescription calls for the first half of the design points to be placed at $-1/2$ and the last half at $1/2$ if n is even. If n is odd, there is as well a design point $x_{(n+1)/2} = 0$.

TABLE 1: Constants for $m_0(x)$; $q = 1$.

$\nu = \sigma^2/\eta^2$	α	β	$\nu = \sigma^2/\eta^2$	α	β
0	0	∞			
0.0001	0.0006	1666	10	15.55	-0.0240
0.001	0.0060	166.6	100	90.23	-0.1487
0.01	0.0595	16.72	1000	737.0	-0.2136
0.1	0.5580	1.709	10000	6886	-0.2379
1	3.810	0.1780	∞	∞	-0.25

When $\rho > 0$ an asymptotically optimal sequence of designs may be obtained from Remark 1 of the previous section. In particular, as $\nu \rightarrow \infty$ designs obtained in this manner tend to that with $x_i = x_{n-i+1} = 0.5 \times (-1)^{i+1}$ for $i < (n+1)/2$ and n even, with as well $x_{(n+1)/2} = 0$ when n is odd. For $\nu = 0$ we have $x_i = (-1)^i\{(i-1)/(n-1) - 1/2\}$. If $\nu = 6.48$, so that $\beta = 0$ in (9), see Example 2 of Section 1.

In the case of even n these limiting designs as $\nu \rightarrow \infty$ were also given by Jenkins & Chamugam (1962). Constantine (1989) gave these designs as well for the BLUE and MA(1) errors.

3.1. The case $q=2$.

For $q \geq 2$ we use the fact that if the random vector \mathbf{X} has a spherically symmetric distribution M_0 , then $Z = \|\mathbf{X}\|$ is distributed independently of $\mathbf{X}/\|\mathbf{X}\|$, which is in turn uniformly distributed over the surface of the unit sphere in \mathbb{R}^q . When $q = 2$ we then have the representation

$$\mathbf{X} = Z \cdot (\cos \Phi, \sin \Phi)',$$

where Z has distribution function H_0 and Φ is uniformly distributed over $[0, 2\pi)$, independently of Z .

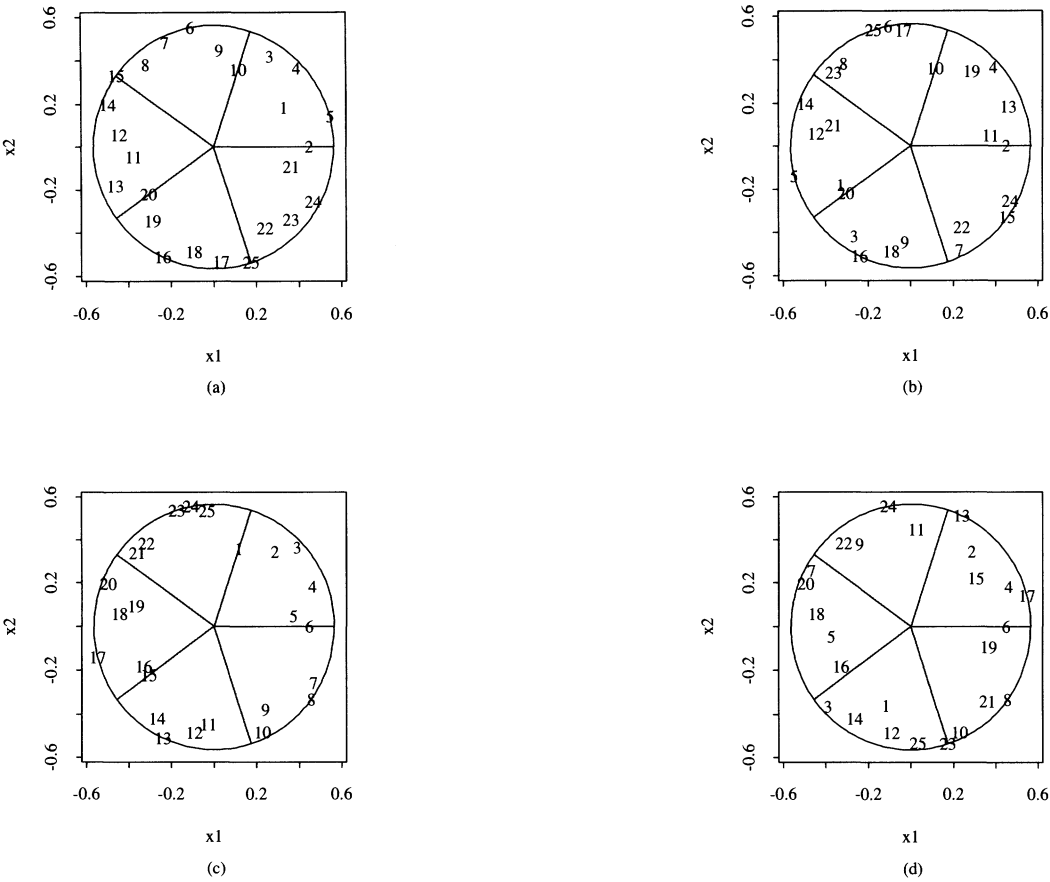


FIGURE 1: Minimax designs for $q = 2, n = 25$ and $\nu = 16/9$. Indices of design points are plotted for (a), (c): $\rho < 0$, (b), (d): $\rho > 0$. Designs (a) and (b) are as in Section 3.2; designs (c) and (d) use the nearest-neighbour algorithm of Section 3.3. Design space is a circle of unit area. Angles of rays are $2\pi i / [\sqrt{n}]$, $i = 1, \dots, [\sqrt{n}]$.

To satisfy the requirements of Theorem 1 when $\rho < 0$, first define $k = [\sqrt{n}]$, $l = n - k^2$. Of the k^2 angles $\{0, 2\pi/k^2, 2 \times 2\pi/k^2, \dots, (k^2 - 1) \times 2\pi/k^2\}$, exactly k are in the interval $I_i := [2\pi(i - 1)/k, 2\pi i/k)$. Assign these at random to the norms $z_j = H_0^{-1}(j/k)$, $j = 1, \dots, k$, to obtain the k points

$$\mathbf{t}_{ij} = z_j \cdot (\cos \phi_{ij}, \sin \phi_{ij})', \quad \phi_{ij} \in I_i, \quad j = 1, \dots, k. \tag{11}$$

For $i = 1, \dots, k$ define $2 \times k$ matrices

$$\mathbf{X}_i = \begin{cases} (\mathbf{t}_{i1}, \mathbf{t}_{i2}, \dots, \mathbf{t}_{ik}), & i \text{ odd,} \\ (\mathbf{t}_{ik}, \dots, \mathbf{t}_{i2}, \mathbf{t}_{i1}), & i \text{ even.} \end{cases}$$

Let ξ_n be the design measure whose points of support $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ are the columns of the $2 \times n$ matrix

$$(\mathbf{0}_{2 \times l} \ \mathbf{X}_1 \ \mathbf{X}_2 \ \cdots \ \mathbf{X}_k),$$

ordered in the indicated manner. Thus design points with angles in I_i are grouped together. Within such a group they are ordered by increasing or decreasing values of their norms.

That condition (B1) of Theorem 1 holds is clear. To check (B2) note that

$$\begin{aligned} \frac{1}{n-1} \sum_{i=1}^{n-1} \|\mathbf{x}_i - \mathbf{x}_{i+1}\|^2 &\leq \frac{1}{n-1} \left\{ 4kr^2 + \sum_{i=1}^k \sum_{j=1}^{k-1} \|\mathbf{t}_{ij} - \mathbf{t}_{i,j+1}\|^2 \right\} \\ &\leq \frac{4r^2}{k-1} + \frac{1}{k+1} \sum_{i=1}^k \left\{ \frac{1}{k-1} \sum_{j=1}^{k-1} \|\mathbf{t}_{ij} - \mathbf{t}_{i,j+1}\|^2 \right\} \end{aligned} \quad (12)$$

and that

$$\|\mathbf{t}_{ij} - \mathbf{t}_{i,j+1}\|^2 \leq (z_j - z_{j+1})^2 + 2z_j z_{j+1} \left(1 - \cos \frac{2\pi}{k} \right). \quad (13)$$

In exactly the same way that (10) was established we find that $\lim_{k \rightarrow \infty} \frac{1}{k-1} \sum_{i=1}^{k-1} (z_j - z_{j+1})^2 = 0$; this together with (12) and (13) gives

$$\frac{1}{n-1} \sum_{i=1}^{n-1} \|\mathbf{x}_i - \mathbf{x}_{i+1}\|^2 \rightarrow 0 \quad \text{as } n \rightarrow \infty. \quad (14)$$

Thus (ξ_n) is asymptotically optimal for $\rho < 0$. For $\rho > 0$ one may apply the method of Remark 1 of Section 2.

See Figure 1(a) and (b) for designs constructed in this manner for $n = 25$ and $\nu = 16/9$, for which $H_0(z) = (\pi z^2)^2$, $0 \leq z \leq \pi^{-\frac{1}{2}}$.

3.3. General q .

Assume that points $(\mathbf{t}_j)_{j=1}^n$ have been generated satisfying (B1) of Theorem 1, perhaps by exploiting the characterization of spherically symmetric distributions given at the beginning of Section 3.2. Consider the following algorithm by which these points may then be ordered:

1. $\mathbf{x}_0 := \mathbf{0}$.
2. For $i = 1, \dots, n$ define \mathbf{x}_i to be the nearest neighbour, among those \mathbf{t}_j not yet chosen, of \mathbf{x}_{i-1} .

A simple argument based on the compactness of \mathcal{S} shows that (14) holds, so that we have:

THEOREM 2. *The sequence (ξ_n) , where ξ_n has design points $(\mathbf{x}_i)_{i=1}^n$ chosen as above, is asymptotically optimal for $\rho < 0$.*

See Figure 1(c) and (d) for designs constructed in this manner, from the same design points as used in the designs of Figure 1(a) and (b).

Example 3. As a further application of Theorem 2 we construct $n = 10$ design points for $q = 6$. We choose $\nu = 256/675$, which again is the boundary between the cases in the definition of H_0 . Let \mathbf{T} have the corresponding density m_0 , so that $Z = \|\mathbf{T}\|$ has distribution function $H_0(z) = 6^{-4/3} \pi^4 z^8$, $0 \leq z \leq 6^{1/6} \pi^{-1/2}$. Then Z and $\mathbf{U} = \mathbf{T}/\|\mathbf{T}\|$ are independent and \mathbf{U} is uniformly distributed over the surface of the unit sphere in \mathbb{R}^6 , so that we can construct design points as follows. We set

$$\mathbf{t}_j = z_j \mathbf{u}_j, \quad j = 1, \dots, 10,$$

where $z_j = H_0^{-1}\{(j-1)/9\} = \{(j-1)/9\}^{1/8}6^{1/6}\pi^{-1/2}$ and the \mathbf{u}_j are independent realizations of \mathbf{U} . One way to select the $\mathbf{u}_j = (u_{j1}, \dots, u_{j6})'$ is to generalize the representation at (11) to

$$u_{ji} = \begin{cases} \sin \phi_{j,1} \sin \phi_{j,2} \cdots \sin \phi_{j,5}, & i = 1, \\ \cos \phi_{j,i-1} \sin \phi_{j,i} \sin \phi_{j,i+1} \cdots \sin \phi_{j,5}, & 1 < i < 6, \\ \cos \phi_{j,5}, & i = 6, \end{cases}$$

where the angles $\phi_{j,1}, \dots, \phi_{j,5}$ are independently and uniformly distributed over $[0, 2\pi)$ for each j . Then (B1) of Theorem 1 holds.

Now obtain the successive nearest neighbours \mathbf{x}_i from among $(\mathbf{t}_1, \dots, \mathbf{t}_{10})$, starting with $\mathbf{x}_0 = \mathbf{0}$, to obtain the design for $\rho < 0$ shown in Table 2. For $\rho > 0$ an optimal design is $(-\mathbf{x}_1, \mathbf{x}_2, -\mathbf{x}_3, \dots, \mathbf{x}_{10})$.

TABLE 2: Design for $\rho < 0$.

i	1	2	3	4	5	6
x_1	0	0	0	0	0	0
x_2	−.192	−.127	+.128	−.146	−.481	−.108
x_3	−.002	+.023	+.348	−.078	−.557	+.291
x_4	+.122	−.465	+.066	+.015	−.315	+.324
x_5	+.019	+.001	+.594	−.180	+.029	−.101
x_6	−.248	+.386	+.328	+.010	−.131	−.370
x_7	−.457	+.510	+.038	−.137	−.164	−.163
x_8	−.061	+.206	−.098	+.109	+.457	+.472
x_9	+.459	−.285	−.423	−.044	+.322	−.035
x_{10}	−.224	−.295	−.136	+.603	+.101	−.180

3.4. Simulation Study.

In this subsection the performance of some optimal designs obtained in this paper is compared, through numerical simulation, with those of some competitors. We consider the simple linear regression model $y_i = \theta_0 + \theta_1 x_i + n^{-1/2} f(x_i) + \varepsilon_i$, with true value $(\theta_0, \theta_1) = (2, 2)$, design space $\mathcal{S} = [-0.5, 0.5]$, $f(x) = \sqrt{180}(x^2 - 1/12)$ (so that $f \in \mathcal{F}$ and $\eta^2 = 1$), and AR(1) errors with parameter ρ and $\sigma^2 = 1$. Since $\nu = \sigma^2/\eta^2 = 1$, the optimal density from Table 1 is $m_0(x) = 3.81(x^2 + 0.178)$, $|x| \leq 0.5$, and the optimal distribution function is $M_0(x) = 1.276x^3 + 0.681x + 0.50$.

We selected $n = 25$ design points from three distributions: $M_0(\cdot)$, the uniform distribution on $[-0.5, 0.5]$ and the distribution with equal mass $1/2$ at each of ± 0.5 . We investigated two orderings for each of the uniform and point-mass distributions, obtaining the following five designs:

- (1) O (optimal): $x_i = \{-\text{sign}(\rho)\}^i M_0^{-1}((i-1)/(n-1))$, $i = 1, \dots, 25$;
- (2) U1 (uniform): $x_i = (i-1)/(n-1) - 0.5$, $i = 1, \dots, 25$;
- (3) U2 (uniform): $x_i = (-1)^i((i-1)/(n-1) - 0.5)$, $i = 1, \dots, 25$;

- (4) P1 (point mass): $(x_i)_{i=1}^n = -0.5, \dots, -0.5, 0, 0.5, \dots, 0.5$;
- (5) P2 (point mass): $(x_i)_{i=1}^n = 0.5, -0.5, \dots, 0.5, -0.5, 0, 0.5, -0.5, \dots, 0.5, -0.5$.

The ordering for U1 and P1 is optimal for $\rho < 0$, and the ordering for U2 and P2 is optimal for $\rho > 0$. For each of a selection of values of ρ we computed the BLUE. After 1000 simulations we calculated the mean squared error of the simulated estimates $\hat{\theta}^{(k)}$:

$$MSE = \frac{1}{1000} \sum_{k=1}^{1000} \{(\hat{\theta}_0^{(k)} - 2)^2 + (\hat{\theta}_1^{(k)} - 2)^2\}.$$

Representative results for $\rho = \pm 0.1, \pm 0.3$ are reported in Table 3. From these values one sees that the MSE is reduced significantly both by the use of the optimal distribution M_0 and by the use of a judicious ordering of the points so obtained.

TABLE 3: Mean squared errors for the designs of Section 3.4.

Design	$\rho = 0.1$	$\rho = 0.3$	$\rho = -0.1$	$\rho = -0.3$
O	.353	.304	.323	.259
U1	.604	.847	.385	.305
U2	.421	.350	.558	.869
P1	.429	.556	.343	.309
P2	.357	.360	.414	.531

APPENDIX. PROOF OF THEOREM 1

Let (ξ_n) converge to a distribution ξ which satisfies (A1)–(A3) but is otherwise arbitrary. By dropping the third summand in (8) and evaluating the remaining terms at $\rho = 0$ we obtain

$$\sup_{f,\rho} \text{IMSE}'(f, \xi, \rho) \geq \sigma^2 \gamma_0 \text{tr} \{ \mathbf{B}^{-1}(\xi) \} + \sup_f \left[\text{E}_M \{ f(\mathbf{U}) \}^2 + \int_{\mathcal{S}} f^2(\mathbf{x}) \, d\mathbf{x} \right]. \tag{A.1}$$

We may clearly eliminate from consideration any ξ for which the maximum loss is infinite. But as in Lemma 1 of Wiens (1992), the supremum on the right side of (A.1) is infinite unless M is absolutely continuous. Thus, let $m = M'$ be the density. Now define

$$f_m(\mathbf{u}) = \frac{\eta \{ m(\mathbf{u}) - 1 \}}{[\int_{\mathcal{S}} \{ m(\mathbf{u}) - 1 \}^2 \, d\mathbf{u}]^{\frac{1}{2}}}$$

and note that $f_m \in \mathcal{F}$ by virtue of assumption (A2) in Section 2. Evaluating the second term in (A.1) at $f = f_m$ gives

$$\sup_{f,\rho} \text{IMSE}'(f, \xi, \rho) \geq \sigma^2 \gamma_0 \text{tr} \{ \mathbf{B}^{-1}(\xi) \} + \eta^2 \int_{\mathcal{S}} m^2(\mathbf{u}) \, d\mathbf{u} =: \Phi(m). \tag{A.2}$$

We will show that

- (i) $\Phi(m)$ is minimized, over all densities $m(\mathbf{u})$, by $m_0(\mathbf{u}) = g_0(\|\mathbf{u}\|)$; and that
- (ii) $\text{E}_{M_0}(\|\mathbf{U}\|^2) = q\gamma$.

Before doing this we first show that (ii), (B1) and (B2) together imply that

$$\sup_{f, \rho} \text{IMSE}'(f, \xi_0, \rho) = \Phi(m_0). \quad (\text{A.3})$$

This, (i) and (A.2) yield the asymptotic optimality.

To see (A.3), note that (ii) and the spherical symmetry of M_0 imply that $B(\xi_0) = \gamma \mathbf{I}$. By (B2) and the weak convergence of ξ_n to ξ_0 we have

$$E_{\xi_0} \{ \|\mathbf{U} + \text{sign}(\rho) \mathbf{V}\|^2 \} = 0, \quad (\text{A.4})$$

so that

$$0 = E_{\xi_0} \{ \{\mathbf{U} + \text{sign}(\rho) \mathbf{V}\} \{\mathbf{U} + \text{sign}(\rho) \mathbf{V}\}' \} = 2\gamma \mathbf{I} + \text{sign}(\rho) \mathbf{C}(\xi_0),$$

whence

$$(1 + \rho^2) \mathbf{B}(\xi_0) - \rho \mathbf{C}(\xi_0) = (1 + |\rho|)^2 \gamma \mathbf{I}. \quad (\text{A.5})$$

Again using (A.4) (together with the Cauchy-Schwarz inequality) we have that

$$\begin{aligned} 0 &= E_{\xi_0} \{ \{\mathbf{U} + \text{sign}(\rho) \mathbf{V}\} \{f(\mathbf{U}) + \text{sign}(\rho) f(\mathbf{V})\} \} \\ &= 2E_{M_0} \{ \mathbf{U} f(\mathbf{U}) \} + \text{sign}(\rho) E_{\xi_0} \{ \mathbf{U} f(\mathbf{V}) + \mathbf{V} f(\mathbf{U}) \}, \end{aligned}$$

whence

$$\mathbf{b}_2(f, \xi_0, \rho) = (1 + |\rho|)^2 E_{M_0} \{ \mathbf{U} f(\mathbf{U}) \}. \quad (\text{A.6})$$

Now (8), (A.5) and (A.6) give

$$\text{IMSE}'(f, \xi_0, \rho) = \frac{\sigma^2}{(1 + |\rho|)^2} \frac{q\gamma_0}{\gamma} + [E_{M_0} \{ f(\mathbf{U}) \}]^2 + \frac{\gamma_0}{\gamma^2} \|E_{M_0} \{ \mathbf{U} f(\mathbf{U}) \}\|^2 + \int_{\mathcal{S}} f^2(\mathbf{x}) d\mathbf{x}. \quad (\text{A.7})$$

The supremum over ρ of the right side of (A.7) is clearly attained at $\rho = 0$. That over f may be obtained by applying Theorem 1 of Wiens (1992), where it is shown that the least favourable function f , for m_0 , is f_{m_0} . Substitution of $\rho = 0$ and $f = f_{m_0}$ in (A.7) gives (A.3).

It remains to establish (i) and (ii) above. For any orthogonal matrix \mathbf{Q} let m_Q be the density on \mathcal{S} given by $m_Q(\mathbf{u}) = m(\mathbf{Q}\mathbf{u})$. Then $\Phi(m_Q) = \Phi(m)$. Since Φ is a convex function of m , we have $\Phi(\frac{1}{2}m + \frac{1}{2}m_Q) \leq \Phi(m)$, and it follows that to minimize Φ we need only consider densities satisfying $m = m_Q$ for all \mathbf{Q} . Such densities are spherically symmetric: $m(\mathbf{u}) = g(\|\mathbf{u}\|)$ for some nonnegative function g satisfying

$$\int_0^r \frac{qz^{q-1}}{r^q} g(z) dz = 1. \quad (\text{A.8})$$

The integrand in (A.8) is the density $h(z)$ of $Z = \|\mathbf{U}\|$. Define $\gamma = E(U_1^2)$ to satisfy (ii), i.e.,

$$\gamma = \int_0^r \frac{z^{q+1}}{r^q} g(z) dz. \quad (\text{A.9})$$

In terms of g we have

$$\Phi(m) = \frac{\sigma^2 \gamma_0}{q\gamma} + \eta^2 \int_0^r \frac{qz^{q-1}}{r^q} g^2(z) dz. \quad (\text{A.10})$$

We first minimize this for fixed γ , i.e., for g satisfying (A.8) and (A.9). Equivalently we minimize

$$\begin{aligned} \int_0^r \frac{qz^{q-1}}{r^q} g^2(z) dz + 2abr^2 \int_0^r \frac{qz^{q-1}}{r^q} g(z) dz - 2aq \int_0^r \frac{z^{q+1}}{r^q} g(z) dz \\ = \int_0^r \frac{qz^{q-1}}{r^q} [g^2(z) - \{2a(z^2 - br^2)\}g(z)] dz \end{aligned} \quad (\text{A.11})$$

for some Lagrange multipliers $a > 0$ and b . This can be done by minimizing the integrand in (A.11) pointwise; the minimizer is $g(z) = a(z^2 - br^2)^+$.

Now determine $a = a(\gamma)$ and $b = b(\gamma)$ to satisfy (A.8) and (A.9). In terms of these constants the right side of (A.10) is

$$\eta^2 \left[\frac{\nu\gamma_0}{q\gamma} + a(\gamma)\{q\gamma - b(\gamma)r^2\} \right]. \quad (\text{A.12})$$

The optimal g_0 is now obtained by minimizing (A.12) over γ for fixed ν . This minimization results in the design described by Cases 1 and 2 preceding the statement of Theorem 1, and completes the proof.

ACKNOWLEDGEMENT

The research of both authors is supported by the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

- P. J. Bickel & A. M. Herzberg (1979). Robustness of design against autocorrelation in time I: Asymptotic theory, optimality for location and linear regression. *The Annals of Statistics*, 7, 77–95.
- G. E. P. Box, W. G. Hunter & J. S. Hunter (1978). *Statistics for Experimenters*. Wiley, Toronto.
- G. M. Constantine (1989). Robust designs for serially correlated observations. *Biometrika*, 76, 245–251.
- V. V. Fedorov (1972). *Theory of Optimal Experiments*. Academic Press, New York.
- A. M. Herzberg & S. Huda (1981). A comparison of equally spaced designs with different correlation structures in one and more dimensions. *The Canadian Journal of Statistics*, 9, 203–208.
- P. J. Huber (1975). Robustness and Designs. In *A Survey of Statistical Design and Linear Models* (J. N. Srivastava, ed.), Amsterdam: North Holland, pp. 287–303.
- G. M. Jenkins & J. Chanmugam (1962). The estimation of slope when the errors are autocorrelated. *Journal of the Royal Statistical Society Series B*, 24, 199–214.
- D. C. Montgomery & E. A. Peck (1982). *Introduction to Linear Regression Analysis*. Wiley, Toronto.
- F. Pukelsheim (1993). *Optimal Design of Experiments*. Wiley, Toronto.
- W. J. Studden (1977). Optimal designs for integrated variance in polynomial regression. In *Statistical Decision Theory and Related Topics II* (S. S. Gupta & D. S. Moore, eds.), New York: Academic Press, pp. 411–420.
- D. P. Wiens (1992). Minimax designs for approximately linear regression. *Journal of Statistical Planning and Inference*, 31, 353–371.
- D. P. Wiens & J. Zhou (1996). Minimax regression designs for approximately linear models with autocorrelated errors. *Journal of Statistical Planning and Inference*, 55, 95–106.
- D. P. Wiens & J. Zhou (1997). Robust designs based on the infinitesimal approach. *Journal of the American Statistical Association*, 92, 1503–1511.

Received 11 February 1998

Accepted 4 November 1998

Douglas P. WIENS: doug.wiens@ualberta.ca

Statistics Centre, Dept. of Mathematical Sciences
University of Alberta, Edmonton, Alberta, Canada T6G 2G1

Julie ZHOU: jzhou@math.uvic.ca

Dept. of Mathematics and Statistics, University of Victoria
Victoria, British Columbia, Canada V8W 3P4