TWO ROBUST DESIGN APPROACHES
FOR LINEAR MODELS WITH CORRELATED ERRORS

Julie Zhou

University of Victoria

Abstract: In this paper, infinitesimal and minimax approaches are used to construct robust regression designs for linear models with correlated errors. We consider IMSE (Integrated Mean Squared Error) as the loss function. Using an infinitesimal approach, we minimize IMSE at the ideal model subject to two robust constraints to derive M-robust designs. We also minimize the maximum of the IMSE to obtain minimax designs. In particular, M-robust and minimax designs are constructed for an approximately linear model with MA(1) errors. These designs are robust against small departures from the assumed regression response and small departures from the assumption of uncorrelated errors. It is interesting that M-robust and minimax designs have the same form of density function, while M-robust designs require less restrictive ordering of design points. Implementation is discussed and examples are given.

Key words and phrases: Approximately linear regression response, infinitesimal approach, M-robust design, minimax design, moving average errors, robust regression design.

1. Introduction

In this paper we study two robust design approaches for the model:

$$y_i = \theta_0 + \boldsymbol{\theta}_1^T \mathbf{x}_i + f(\mathbf{x}_i) + \epsilon_i, \quad i = 1, \ldots, n, \tag{1.1}$$

where design points $\mathbf{x}_i$'s belong to a q-dimensional space $\mathcal{S}$, and f is from a class of functions $\mathcal{F}$. Model departure from linear regression is reflected in the term f. If $f \equiv 0$, the regression response degenerates to the linear case.

The assumption of uncorrelated errors is sometimes unrealistic. Successive observations are likely to be correlated. The presence of serial correlation may affect our design strategy. Throughout the paper we assume that errors follow a first-order moving average (MA(1)) process. Specifically, $\epsilon_i = e_i + ae_{i-1}, \quad i = 1, \cdots, n$, where the e_i are white noise variables with mean 0 and variance σ_0^2, and a is the parameter of MA(1) process, $-1 < a < 1$. For this error model,

$$COV(\boldsymbol{\epsilon}) = \sigma_\epsilon^2 P, \tag{1.2}$$

where $\sigma_\epsilon^2 = 2\sigma_0^2/(1 + \sqrt{1 - 4\rho^2})$ with $\rho = a/(1 + a^2)$, and the autocorrelation matrix $P = (p_{i,j})$ has $p_{i,i} = 1$, $p_{i,i+1} = p_{i+1,i} = \rho$, and $p_{i,j} = 0$ otherwise. For $-1 < a < 1$, ρ^2 is less than $1/4$.

Let $\hat{\boldsymbol{\theta}}$ be the least squares estimate of $\boldsymbol{\theta} = (\theta_0, \boldsymbol{\theta}_1^T)^T$, i.e., $\hat{\boldsymbol{\theta}} = \left(X^T X\right)^{-1} X^T \mathbf{y}$, where the ith row of matrix X is $(1, \mathbf{x}_i^T)$ and $\mathbf{y} = (y_1, \ldots, y_n)^T$. The loss function we consider is the Integrated Mean Squared Error (IMSE) of the estimated response, i.e.,

$$IMSE_n(f, \rho, \xi) = n \int_{\mathcal{S}} E[(\hat{\theta}_0 + \hat{\boldsymbol{\theta}}_1^T \mathbf{x} - E[y|\mathbf{x}])^2] d\mathbf{x}, \qquad (1.3)$$

where ξ is the design measure of $\mathbf{x}_1, \ldots, \mathbf{x}_n$ on $\mathcal{S}$.

A design ξ^* is called a *minimax design* if it minimizes the maximum (over ρ and f) of the $IMSE_n(f, \rho, \xi)$ asymptotically:

$$\lim_{n\to\infty} \sup_{\rho\in\mathcal{P}} \sup_{f\in\mathcal{F}} IMSE_n(f, \rho, \xi^*) \le \lim_{n\to\infty} \sup_{\rho\in\mathcal{P}} \sup_{f\in\mathcal{F}} IMSE_n(f, \rho, \xi) \ for \ all \ \xi, \qquad (1.4)$$

for a given class $\mathcal{P}$ of ρ.

A second approach to robust designs is to minimize $IMSE_n(f, \rho, \xi)$ at the ideal model ($f = 0$, $\rho = 0$, i.e., the regression response is linear with uncorrelated errors) subject to two constraints. One constraint guarantees a small change in $IMSE_n$ if there is a small departure in the assumed regression response. Another guarantees a small change in $IMSE_n$ if errors are correlated. Specifically, we call ξ^M an M-robust design if it is a solution to

$$\min_{\xi} IMSE_n(0, 0, \xi) \qquad (1.5)$$

$$s.t. \quad \lim_{n\to\infty} \sup_{f\in\mathcal{F}} \left\{ \frac{\partial^2}{\partial t^2} IMSE_n(tf, 0, \xi)|_{t=0} \right\} \le \alpha, \qquad (1.6)$$

$$\lim_{n\to\infty} \sup_{\rho\in\mathcal{P}} \left\{ \frac{\partial}{\partial s} IMSE_n(0, s\rho, \xi)|_{s=0} \right\} \le \beta, \qquad (1.7)$$

where $\alpha > 0$ and $\beta > 0$ are two constants. The second derivative is used in (1.6), since $IMSE_n(tf, 0, \xi)$ is a quadratic function of t.

Minimax designs have received much attention. Examples of work in the area are Huber (1975, 1981), Kiefer and Wynn (1984), Li and Notz (1982), Wiens (1992), Sitter (1992), Wong (1992), and Wiens and Zhou (1996, 1998). In Wiens and Zhou (1996), minimax designs are studied for (1.1) and (1.2) without assuming any specific error structure, and a minimax design for uncorrelated errors retains its optimality under autocorrelation if the design points are a random sample, or a random permutation, of points from this distribution. The result concludes that randomization is robust against autocorrelation of errors.

However, if additional information is available about the error structure, one could use the information to improve the designs. In Wiens and Zhou (1998) a first-order autoregressive AR(1) process is considered, and the result shows that the design points should be selected purposively and not randomly.

The infinitesimal approach to robust design was introduced by Wiens and Zhou (1997). Their loss function is the determinant of the mean squared error matrix, and various robust designs (V-robust, B-robust, and M-robust) are defined and studied. These designs minimize the loss function at the ideal model subject to some robustness constraints formulated in terms of boundedness of the Gateaux derivatives of the loss function in the directions of a contaminating response function and/or autocorrelation structure. The theory of the infinitesimal approach to robust statistics in general can be found in Hampel, Ronchetti, Rousseeuw and Stahel (1986).

In this paper we derive minimax and M-robust designs for MA(1) error processes comparing to designs for the AR(1) error processes in Wiens and Zhou (1998), and to the designs for a general stationary error process in Wiens and Zhou (1996). We also provide a detailed comparison of minimax and M-robust designs. Our results suggest that the two design approaches yield similar distributions of design points but that M-robust designs require less restriction to the order in which they are taken.

In Section 2, we derive minimax (ξ^*) and M-robust (ξ^M) designs. In Section 3, we compare M-robust designs with minimax designs, and discuss how to implement them in practice. Guidelines are given to select finite design points from ξ^* and ξ^M. The proofs of some Theorems are presented in the Appendix.

2. Minimax Design ξ^* and M-robust Design ξ^M

The following class of disturbance functions is considered in this paper: $\mathcal{F} \doteq \{f \mid \int_{\mathcal{S}} f(\mathbf{x})d\mathbf{x} = 0, \int_{\mathcal{S}} \mathbf{x}f(\mathbf{x})d\mathbf{x} = \mathbf{0}, \int_{\mathcal{S}} f^2(\mathbf{x})d\mathbf{x} \leq \eta_n^2\}$. This ensures the identifibility of regression parameter in model (1.1). For the design space $\mathcal{S}$, we use the sphere with volume 1 and radius $(\Gamma(0.5q + 1))^{1/q}/\sqrt{\pi}$, denoted r. Define $A = \int_{\mathcal{S}} \begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix} (\,1\ \mathbf{x}^T\,)d\mathbf{x} = 1 \oplus \gamma_0 I_q, \ \gamma_0 = r^2/(q+2), \ B(\xi) = \int_{\mathcal{S}} \begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix} (\,1\ \mathbf{x}^T\,)d\xi(\mathbf{x}), \ D(\xi, \rho) = X^T P X/n, \mathbf{b}_{f,\xi} = \int_{\mathcal{S}} f(\mathbf{x}) \begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix} d\xi(\mathbf{x})$. Then the integrated mean squared error can be decomposed into three terms

$$IMSE_n(f, \rho, \xi) = n\mathbf{b}_{f,\xi}^T B^{-1}(\xi) A B^{-1}(\xi) \mathbf{b}_{f,\xi}$$
$$+ \sigma_\epsilon^2 trace(B^{-1}(\xi)D(\xi,\rho)B^{-1}(\xi)A) + n\int_{\mathcal{S}} f^2(\mathbf{x})d\mathbf{x}. \quad (2.1)$$

The first term arises from the bias of regression estimates, the second from the covariance matrix, and the third from model misspecification.

To derive minimax and M-robust designs, we assume (A1) ξ has a density function $m(\mathbf{x})$; (A2) ξ is spherically symmetric; (A3) $\lim_{n\to\infty} n\eta_n^2$ exists and equals $\tau^2 > 0$. Assumption A1 is necessary for $\sup_{f\in\mathcal{F}} IMSE_n(f,\rho,\xi) < \infty$, see Lemma 1 in Wiens (1992). Spherically symmetric designs have nice properties such as orthogonality and rotatability. Under A2, $B(\xi)$ is diagonal, and the term $trace(B^{-1}(\xi)D(\xi,\rho)B^{-1}(\xi)A)$ in (2.1) can be simplified. Assumption A3 is needed in order that bias and variance are of the same magnitude in $IMSE_n(f,\rho,\xi)$.

Under A2, the density function $m(\mathbf{x})$ can be written as $g(\| \mathbf{x} \|)$, where $g(u)$ satisfies

$$\int_0^r \frac{qu^{q-1}}{r^q} g(u)du = 1. \tag{2.2}$$

The density function for $U = \| \mathbf{x} \|$ is $h(u) = (qu^{q-1}/r^q)g(u)$. Let $\gamma = E[X_j^2] = \int_0^r \frac{u^{q+1}}{r^q}g(u)du$ and $J_0(g,\gamma) = \int_0^r \frac{qu^{q-1}}{r^q}(g(u)-1)^2 du$. Applying Theorem 1 in Wiens (1992), we get

$$\sup_{f\in\mathcal{F}} IMSE_n(f,\rho,\xi) = n\eta_n^2 J_0(g,\gamma) + \sigma_\epsilon^2 trace(B^{-1}(\xi)D(\xi,\rho)B^{-1}(\xi)A) + n\eta_n^2. \tag{2.3}$$

To analyze the second term in (2.3), we introduce lag-1 autocorrelation for each regressor x_j: $r_{j,n} = \sum_{i=2}^n x_{ij}x_{(i-1),j}/\sum_{i=1}^n x_{ij}^2$, $j = 1,\ldots,q$, where x_{ij} is the value of x_j from ith design point. Let $r_j = \lim_{n\to\infty} r_{j,n}$ for all $j = 1,\ldots,q$, then

$$trace(B^{-1}(\xi)D(\xi,\rho)B^{-1}(\xi)A) = 1 + q\frac{\gamma_0}{\gamma} + 2\rho\Big(\frac{n-1}{n} + \frac{\gamma_0}{\gamma}\sum_{i=1}^q r_{i,n}\Big). \tag{2.4}$$

Therefore from (2.3),

$$\sup_{f\in\mathcal{F}} IMSE_n(f,\rho,\xi)$$
$$= n\eta_n^2 + n\eta_n^2 J_0(g,\gamma) + \sigma_\epsilon^2(1 + q\frac{\gamma_0}{\gamma}) + 2\rho\sigma_\epsilon^2(\frac{n-1}{n} + \frac{\gamma_0}{\gamma}\sum_{i=1}^q r_{i,n}). \tag{2.5}$$

We derive minimax designs and M-robust designs for MA(1) processes when $\mathcal{P}_1 = \{ \rho \mid 0 < a_0 \leq \rho \leq a_1 < 0.5 \}$ and $\mathcal{P}_2 = \{ \rho \mid -0.5 < b_0 \leq \rho \leq b_1 < 0 \}$.

Theorem 2.1. *Suppose the density g_ν^* minimizes $L(g,\nu) = J_0(g,\gamma) + q\nu\frac{\gamma_0}{\gamma}$ for fixed ν. Then the minimax design ξ^* has density function $g_{\nu_0}^*$ and requires $r_1 = \cdots = r_q = -sign(\rho)$, where ν_0 is either $\sigma_0^2(2-4a_0)/(\tau^2(1+\sqrt{1-4a_0^2}))$*

or $\sigma_0^2(2 - 4a_1)/(\tau^2(1 + \sqrt{1 - 4a_1^2}))$ for $\mathcal{P}_1$, and ν_0 is either $\sigma_0^2(2 + 4b_0)/(\tau^2(1 + \sqrt{1 - 4b_0^2}))$ or $\sigma_0^2(2 + 4b_1)/(\tau^2(1 + \sqrt{1 - 4b_1^2}))$ for $\mathcal{P}_2$.

The requirement $r_1 = \cdots = r_q = -sign(\rho)$ in Theorem 2.1 imposes a particular ordering for design points for minimax designs: consecutive design points are as far away as possible for $\mathcal{P}_1$ and as close as possible for $\mathcal{P}_2$. Detailed strategies of implementing minimax designs are discussed in Section 3. From Wiens (1992), the density function g_ν^* has the form $g_\nu^*(u) = a^*(u^2 - b^*r^2)^+$, $a^* > 0$, $b^* \leq 1$, $0 \leq u \leq r$, where constants a^* and b^* depend on the value of ν. For small ν, we solve γ/γ_0 from the equation $((q + 4)/2)(\gamma/\gamma_0 - 1)(\gamma/\gamma_0)^2 = \nu$. If $1 \leq \gamma/\gamma_0 \leq (q + 2)^2/(q(q + 4))$, then $g_\nu^*(u) = 1 + (\gamma/\gamma_0 - 1)((q + 4)/4)(u^2/\gamma_0 - q)$, $0 \leq u \leq r$. For large ν, we solve b from equation $\nu = 2K_{q+2}^2(b)/((q + 2)K_q^3(b))$, where $K_q(b) = q\int_{\sqrt{b}}^1 v^{q-1}(v^2 - b)dv$, and then $g_\nu^*(u) = [(u/r)^2 - b]/K_q(b)$, $r\sqrt{b} \leq u \leq r$.

Theorem 2.2. *For $\mathcal{P}_1$, the M-robust design ξ^M has density function $g^M = a^*(u^2 - b^*r^2)^+$ (Lemma 4.1 in the Appendix) with $\alpha' = (\alpha - \tau^2)/\tau^2$, $\beta' = 1/q - \beta/(2\sigma_0^2 a_1 q)$, and ξ^M requires that $r_1 + \cdots + r_q \leq (\beta/(2\sigma_0^2 a_1) - 1)\gamma(g^M)/\gamma_0$.*

The constants a^* and b^* are determined by α' and β' in Lemma 4.1 in the Appendix. It is true that $(\beta/(2\sigma_0^2 a_1) - 1)\gamma(g^M)/\gamma_0 \geq -q$, therefore $r_1 = -1, \ldots, r_q = -1$ always satisfies the requirement on the r_i. A similar result is obtained for $\mathcal{P}_2$ with $\beta' = \beta/(2\sigma_0^2 b_1 q) - 1/q$, and the design ξ^M requires that $r_1 + \cdots + r_q \geq (\beta/(2\sigma_0^2 b_1) - 1)\gamma(g^M)/\gamma_0$. The condition on the r_i can be satisfied with $r_1 > 0, \ldots, r_q > 0$.

3. Implementation and Comparison

From Section 2, both minimax and M-robust designs have density functions of the form $a^*(u^2 - b^*r^2)^+$. For ξ^*, a^* and b^* depend on σ_0^2/τ^2 and on the range of autocorrelation parameter ρ. For ξ^M, a^* and b^* depend mainly on the values of the constraint bounds α and β.

In the limiting case of $\tau^2 = \lim_{n\to\infty} n\eta_n^2 = 0$, i.e. $\sigma_0^2/\tau^2 \to \infty$, the bias effect in $IMSE$ is 0, so we have a pure variance problem. The optimal distribution of $||\mathbf{x}||$ from the minimax design ξ^* is the pointmass at $||\mathbf{x}|| = r$. For the M-robust design ξ^M, the optimal distribution is determined by the constraint parameter β in (1.7), because constraint (1.6) becomes trivial ($\tau^2 + \tau^2 J_0(g, \gamma) \leq \alpha$ for all $\alpha \geq 0$). For example, if $q = 1$ and $\beta' = 5/9$, the optimal distribution for $||\mathbf{x}||$ is $H(u) = 8u^3$, $0 \leq u \leq 0.5$; if $q = 1$ and $\beta' = 1$, the optimal distribution for $||\mathbf{x}||$ is the uniform distribution on $[0, r]$, and $H(u) = 2u, 0 \leq u \leq 0.5$; if $q = 1$ and $\beta' < 1/3$, then constraint (1.7) always holds and the pointmass at $||\mathbf{x}|| = r$ is optimal.

In the case of $\tau^2 = \lim_{n\to\infty} n\eta_n^2 = \infty$, i.e., $\sigma_0^2/\tau^2 \to 0$, the variance term in $IMSE$ vanishes. The optimal distribution of $\|\mathbf{x}\|$ from the minimax design ξ^* is uniform on $[0, r]$. For the M-robust design ξ^M, constraint (1.6) does not hold because of τ^2. However we can have $J_0(g, \gamma) = 0$ if only $\tau^2 J_0(g, \gamma) \leq \alpha$ is considered, which also leads to the uniform distribution on $[0, r]$.

Both ξ^* and ξ^M have requirements on $r_1, \ldots, r_q$. As a result, ξ^* and ξ^M impose orders on the design points $\mathbf{x}_1, \ldots, \mathbf{x}_n$. From Theorem 2.1, the requirement on $r_1, \ldots, r_q$ by ξ^* is very restrictive: all r_i equal -1 for $\mathcal{P}_1$ and 1 for $\mathcal{P}_2$. On the other hand the requirement by ξ^M in Theorem 2.2 is less restrictive, since $r_1 > 0, \ldots, r_q > 0$ satisfies the requirement for M-robust designs for $\mathcal{P}_2$, and $r_1 < 0, \ldots, r_q < 0$ usually satisfies the requirement for $\mathcal{P}_1$. Furthermore the requirements on $r_1, \ldots, r_q$ of ξ^* always satisfy those of ξ^M. So the minimax designs are special cases of M-robust designs if their distribution functions are the same.

To implement these designs, we can select design points $\mathbf{x}_1, \ldots, \mathbf{x}_n$ in two steps. Suppose H is the distribution function of $U = \| \mathbf{x} \|$.

Step 1. Choose n spherically symmetric points $\mathbf{w}_1, \ldots, \mathbf{w}_n$ in the design space $\mathcal{S}$ such that the empirical distribution function of $\|\mathbf{w}_1\|, \ldots, \|\mathbf{w}_n\|$ converges to H.

Step 2. Select $\mathbf{x}_1, \ldots, \mathbf{x}_n$ as one particular permutation of $\mathbf{w}_1, \ldots, \mathbf{w}_n$ which is determined by $r_1, \ldots, r_q$.

Example 1. We implement the minimax design ξ^* for $q = 1$. The design space is $\mathcal{S} = [-0.5, 0.5]$. In step 1, choose n symmetric points $w_1, \ldots, w_n$ as follows. Let $l = [n/2]$, $w_i = -H^{-1}((l - i + 0.5)/l)$ for $i = 1, \ldots, l$, and $w_i = -w_{n-i+1}$ for $i = n - l + 1, \ldots, n$. If n is odd, then there is a middle point $w_{(n+1)/2} = 0$. These points are in an increasing order, $w_1 \leq w_2 \leq \ldots \leq w_n$. In step 2, select $x_1, \ldots, x_n$ by the requirement on r_1. For $\mathcal{P}_1$, ξ^* requires $r_1 = -1$, i.e., $\sum x_i x_{i+1} / \sum x_i^2 \to -1$, as $n \to \infty$. The following selection satisfies this requirement: $x_1 = w_1, x_2 = w_n, x_3 = w_2, x_4 = w_{n-1}, \ldots, x_{2i-1} = w_i, x_{2i} = w_{n-i+1}, \ldots$. For $\mathcal{P}_2$, ξ^* requires $r_1 = 1$, so we select $x_1 = w_1, x_2 = w_2, \ldots, x_n = w_n$.

Example 2. Implement ξ^M for $q = 1$. In step 1, choose $w_1, \ldots, w_n$ as in Example 1. In step 2, the requirement on r_1 imposed by ξ^M is not so restrictive. For $\mathcal{P}_1$ we only need $r_1 < 0$. Therefore there are many ways to select $\mathbf{x}_1, \ldots, \mathbf{x}_n$ including the one in Example 1. Let $w^I = \{ w_i \,|\, w_i < 0 \}$ and $w^{II} = \{ w_i \,|\, w_i \geq 0 \}$, then $x_1, \ldots, x_n$ can be any permutation of $w_1, \ldots, w_n$ such that design points alternate between w^I and w^{II}. For $\mathcal{P}_2$, ξ^M requires $r_1 > 0$. So $x_1, \ldots, x_n$ can be any permutation of $w_1, \ldots, w_n$ such that the first half of design points are from w^I and the second half are from w^{II}.

In general, step 1 is the same, and step 2 is different in implementing ξ^* and ξ^M. Stricter requirements on $r_1, \ldots, r_q$ imposed by ξ^* allow little flexibility

in choosing $\mathbf{x}_1, \ldots, \mathbf{x}_n$ at step 2. The requirement $r_1 = \ldots = r_q = -sign(\rho)$ is equivalent to $\lim_{n \to \infty} E[||\mathbf{x}_{i+1} + sign(\rho)\mathbf{x}_i||^2] = 0$, which is the requirement imposed by the minimax designs for approximately linear models with AR(1) errors in Wiens and Zhou (1998). Therefore the strategies in Wiens and Zhou (1998) can be used in Step 2 to implement ξ^*. For $\mathcal{P}_2$, we can select $\mathbf{x}_1, \ldots, \mathbf{x}_n$ by the *nearest neighbour method*:

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

The design points for $\mathcal{P}_1$ can then be taken as $(-1)^1\mathbf{x}_1, (-1)^2\mathbf{x}_2, \ldots, (-1)^n\mathbf{x}_n$.

For ξ^M, it is relatively easier to implement step 2. For $q = 2$, points $\mathbf{w}_1, \ldots, \mathbf{w}_n$ selected in step 1 fall in four quadrants, say $\mathbf{w}^1, \mathbf{w}^2, \mathbf{w}^3$ and $\mathbf{w}^4$. For $\mathcal{P}_1$, ξ^M requires designs points $\mathbf{x}_1, \ldots, \mathbf{x}_n$ alternating between $\mathbf{w}^1$ and $\mathbf{w}^3$ first, then alternating between $\mathbf{w}^2$ and $\mathbf{w}^4$. For $\mathcal{P}_2$, ξ^M requires that consecutive design points stay within the same quadrant as long as possible. It is important to emphasize that this strategy can be easily generalized for $q > 2$. Denote the sign change for two consecutive design points $\mathbf{x}_i$ and $\mathbf{x}_{i+1}$ as

$$s_i = \sum_{j=1}^{q} \{1 - sign(x_{ij}) \cdot sign(x_{(i+1),j})\}/2.$$

Set initial value $\mathbf{x}_1 = \mathbf{w}_1$, and choose the next design point $\mathbf{x}_i$ to keep s_{i-1} as small as possible for $\mathcal{P}_2$, and to keep s_{i-1} as large as possible for $\mathcal{P}_1$ (*sign change method*).

Example 3. For $q = 2$, we construct $n = 16$ points from both the minimax and the M-robust designs for $\mathcal{P}_1$ and $\mathcal{P}_2$. The optimal distribution for $||\mathbf{x}||$ is $H(u) = (\pi/2)u^2 + (\pi^2/2)u^4$, $0 \le u \le \pi^{-1/2}$, which corresponds to $\nu = 49/72, \alpha' = 1/3$ and $\beta' = 6/7$. Since $\mathbf{x}/||\mathbf{x}||$ is uniformly distributed over the unit circle in $\mathbb{R}^2$, and $||\mathbf{x}||$ and $\mathbf{x}/||\mathbf{x}||$ are independent, one way to select the design points is to take

$$\mathbf{w}_i = t_i \begin{pmatrix} \sin \phi_i \\ \cos \phi_i \end{pmatrix}, \quad i = 1, \ldots, n,$$

where $t_i = H^{-1}((i - 0.5)/n)$, and ϕ_i is randomly selected from the uniform distribution on $[0, 2\pi)$, or ϕ_i can be randomly selected from $\{0, 2\pi/n, \ldots, 2(n-1)\pi/n\}$ without replacement. The latter is used in this example. In this case, the empirical distribution of $||\mathbf{w}_i||$ converges to H. Using the nearest neighbour method, we can arrange the order of design points $\mathbf{w}_i$ to get minimax designs. For M-robust designs, the sign change method is applied. Figure 1 shows the

distribution of design points of the minimax designs and the M-robust designs respectively.
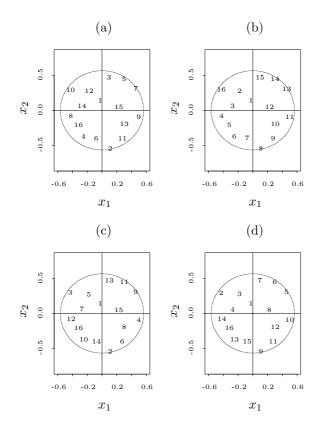

Figure 1. Minimax and M-robust design points in two dimensional space: (a) minimax design for $\mathcal{P}_1$, (b) minimax design for $\mathcal{P}_2$, (c) M-robust design for $\mathcal{P}_1$, (d) M-robust design for $\mathcal{P}_2$. Order of implementation is numerical.

Example 4. For $q = 4$, we construct $n = 16$ points from both the minimax and the M-robust designs for $\mathcal{P}_2$. Suppose the optimal distribution for $\|\mathbf{x}\|$ is $H(u) = 8^{-1/2}\pi^3 u^6$, $0 \leq u \leq 2^{1/4}\pi^{-1/2}$ corresponding to $\nu = 81/128, \alpha' = 1/5$ and $\beta' = 8/9$. Let $\mathbf{w}_i = t_i \mathbf{u}_i$, $i = 1, \ldots, n$, where $t_i = H^{-1}((i - 0.5)/n)$, and the $\mathbf{u}_i$ are randomly selected from the uniform distribution over the surface of the unit sphere. One way to select $\mathbf{u}_i$ is

$$\mathbf{v}_i = \begin{pmatrix} \sin\phi_{1,i}\sin\phi_{2,i}\sin\phi_{3,i} \\ \cos\phi_{1,i}\sin\phi_{2,i}\sin\phi_{3,i} \\ \cos\phi_{2,i}\sin\phi_{3,i} \\ \cos\phi_{3,i} \end{pmatrix}, \quad \mathbf{u}_i = T_i\mathbf{v}_i,$$

where $\phi_{1,i}, \phi_{2,i}, \phi_{3,i}$ are randomly and independently generated from the uniform distribution on $[0, 2\pi)$ for each $\mathbf{v}_i$, and T_i is a permutation matrix randomly generated for each $\mathbf{u}_i$. It is easy to verify that $||\mathbf{u}_i|| = ||\mathbf{v}_i|| = 1$ for $i = 1, \ldots, n$, and each point on the surface of the unit sphere can be represented in the form of $\mathbf{v}_i$. The permutation matrix T_i makes $\mathbf{u}_i$ uniformly distributed on the surface. Splus is used to generate $\phi_{1,i}, \phi_{2,i}, \phi_{3,i}$ T_i, and to carry out both sign change method and nearest neighbour method.

Based on the $\mathbf{w}_i$'s, we determine the sequence $\mathbf{x}_1, \ldots, \mathbf{x}_{16}$ for the M-robust design by minimizing sign changes for consecutive design points. One sequence of the design points $\mathbf{x}_1, \ldots, \mathbf{x}_{16}$ is displayed in Figure 2. Since design points are in 4-dimensional space, we use two plots (x_1 versus x_2 and x_3 versus x_4) to show the design points graphically. The minimax design is obtained by using the nearest neighbour method, and has exactly the same points as the M-robust design but a different ordering: $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_4, \mathbf{x}_{13}, \mathbf{x}_{11}, \mathbf{x}_{12}, \mathbf{x}_{10}, \mathbf{x}_8, \mathbf{x}_9, \mathbf{x}_7, \mathbf{x}_{14}, \mathbf{x}_{15}, \mathbf{x}_{16}$.
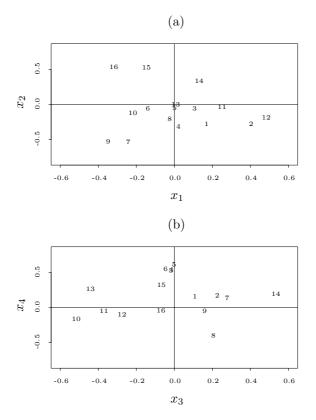

Figure 2. M-robust design points in 4-dimensional space: (a) x_1 vs x_2, (b) x_3 vs x_4. Order of implementation is numerical.

We conclude that minimax and infinitesimal approaches produce consistent robust designs for linear models with correlated errors. M-robust designs have the advantage of a less restrictive ordering of design points. In practice it is not necessary to choose η_n^2. One only needs to choose ν (related to the ratio σ_0^2/τ^2) to determine optimal density function g_ν^* in Theorems 2.1, which leads to the optimal density $h(u) = (qu^{q-1}/r^q)g_\nu^*(u)$ for $||\mathbf{x}||$. Since $0 \leq \nu < \infty$, and two limiting cases $\nu = 0$ and $\nu = \infty$ correspond to the pure bias problem and pure variance problem respectively, ν can be viewed as a measure of efficiency versus robustness against model departure in the response. When ν increases, we gain more efficiency but less robustness. Furthermore, the optimal designs are not sensitive to the value of ν if ν is not too close to 0 or ∞.

Acknowledgement

Appendix

Proof of Theorem 2.1. First $\sup_{\rho \in \mathcal{P}_1} \sup_{f \in \mathcal{F}} IMSE_n(f, \rho, \xi)$ is obtained at either $\rho = a_0$ or a_1. Without loss of generality we assume that $\rho = a_0$ gives the maximum value of $\lim_{n \to \infty} \sup_{f \in \mathcal{F}} IMSE_n(f, \rho, \xi^*)$, then $\lim_{n \to \infty} \sup_{\rho \in \mathcal{P}_1}$ $\sup_{f \in \mathcal{F}} IMSE_n(f, \rho, \xi^*) = \tau^2(1 + L(g_{\nu_1}^*, \nu_1)) + \sigma_0^2(2 + 4a_0)/(\tau^2(1 + \sqrt{1 - 4a_0^2}))$ with $\nu_1 = \sigma_0^2(2 - 4a_0)/(\tau^2(1 + \sqrt{1 - 4a_0^2}))$. Using the fact that $g_{\nu_1}^*$ minimizes $L(g, \nu_1)$ and $r_i = -1$, we have, for any ξ,

$$\lim_{n \to \infty} \sup_{\rho \in \mathcal{P}_1} \sup_{f \in \mathcal{F}} IMSE_n(f, \rho, \xi) \geq \lim_{n \to \infty} \sup_{\rho \in \mathcal{P}_1} \sup_{f \in \mathcal{F}} IMSE_n(f, \rho, \xi^*).$$

The result for $\mathcal{P}_2$ can be proved similarly.

The following lemma is used in Theorem 2.2.

Lemma 4.1. *The density* $g^M = a^*(u^2 - b^*r^2)^+$ *is the solution to*

$$\min_{\xi} \frac{\gamma_0}{\gamma}, \qquad s.t. \ \ J_0(g, \gamma) \leq \alpha' \quad and \quad \frac{\gamma_0}{\gamma} \geq \beta', \tag{A.1}$$

where constants a^* *and* b^* *depend on* α' *and* β'.

The proof is based on the idea that the density $g = a^*(u^2 - b^*r^2)^+$ minimizes $J_0(g, \gamma)$. We need to find the minimum of γ_0/γ by choosing a^* and b^* properly. Constants a^* and b^* are determined as follows. Formulas for γ_0/γ and $J_0(g, \gamma)$ can be found in Wiens (1992, p.364).

Case 1. For $0 \leq \alpha' \leq 4/(q(q+4))$,

$$g^M(u) = 1 + (\frac{1}{c_M} - 1)(\frac{q+4}{4})(\frac{u^2}{\gamma_0} - q), \quad 0 \leq u \leq r, \tag{A.2}$$

where $c_M = \max\{\beta', \sqrt{q(q+4)}/(\sqrt{4\alpha'} + \sqrt{q(q+4)})\}$.

Case 2. For $\alpha' > 4/(q(q+4))$ and $q(q+4)/(q+2)^2 \leq \beta' \leq 1$, $g^M(u)$ has the form of (A.2) with $c_M = \beta'$.

Case 3. For $\alpha' > 4/(q(q+4))$ and $q/(q+2) \leq \beta' \leq q(q+4)/(q+2)^2$, first we solve for b in

$$\begin{cases} \dfrac{\gamma}{\gamma_0} = \dfrac{K_{q+2}(b)}{K_q(b)} \\ J_0(g, \gamma) = \dfrac{q\gamma - br^2}{r^2 K_q(b)} - 1 = \alpha', \end{cases}$$

and write the solution as $\hat{b}$. Then we solve for b in $\gamma/\gamma_0 = K_{q+2}(b) / K_q(b) = \beta'$ and denote the solution as $\tilde{b}$. Set $b^1 = min\{\hat{b}, \tilde{b}\}$, then the density is $g^M(u) = ((u/r)^2 - b^1)/K_q(b^1), \quad r\sqrt{b^1} \leq u \leq r$.

Proof of Theorem 2.2. The optimization problem defined by (1.5) - (1.7) is equivalent to (A.1) in Lemma 4.1 based on the following:

$$IMSE_n(0, 0, \xi) = \sigma_0^2(1 + q\frac{\gamma_0}{\gamma}),$$

$$\lim_{n \to \infty} \sup_{f \in \mathcal{F}} \left\{ \frac{\partial^2}{\partial t^2} IMSE_n(tf, 0, \xi)|_{t=0} \right\} = \tau^2(1 + J_0(g, \gamma)),$$

$$\lim_{n \to \infty} \sup_{\rho \in \mathcal{P}_1} \left\{ \frac{\partial}{\partial s} IMSE_n(0, s\rho, \xi)|_{s=0} \right\} = 2\sigma_0^2 a_1(1 + \frac{\gamma_0}{\gamma} \sum_{i=1}^q r_i).$$

The condition on r_i is obtained from constraint (1.7).

References

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. (1986). *Robust Statistics: The Approach Based on Influence Functions*. Wiley, New York.

Huber, P. J. (1975). Robustness and designs. In *A Survey of Statistical Design and Linear Models*. (Edited by J. N. Srivastava), 287-303. North Holland, Amsterdam.

Huber, P. J. (1981). *Robust Statistics*. Wiley, New York.

Kiefer, J. and Wynn, H. P. (1984). Optimum and minimax exact treatment designs for one-dimentional autoregressive error process. *Ann. Statist.* **12**, 414-450.

Li, K. C. and Notz, W. (1982). Robust designs for nearly linear regression. *J. Statist. Plann. Inference* **6**, 135-151.

Sitter, R. R. (1992). Robust designs for binary data. *Biometrics* **48**, 1145-1156.

Wiens, D. P. (1992). Minimax designs for approximately linear regression. *J. Statist. Plann. Inference* **31**, 353-371.

Wiens, D. P. and Zhou, J. (1996). Minimax regression designs for approximately linear models with autocorrelated errors. *J. Statist. Plann. Inference* **55**, 95-106.

Wiens, D. P. and Zhou, J. (1997). Robust designs based on the infinitesimal approach. *J. Amer. Statist. Assoc.* **92**, 1503-1511.

Wiens, D. P. and Zhou, J. (1998). Minimax designs for approximately linear models with AR(1) errors. *Canad. J. Statist.* to appear.

Wong, W. K. (1992). A unified approach to the construction of mini-max designs. *Biometrika* **79**, 611-620.

Department of Mathematics and Statistics, University of Victoria, Victoria, BC, Canada V8W 3P4.

E-mail: jzhou@math.uvic.ca