The Korean Communications
in Statistics Vol. 3, No. 3, 1996
pp. 93-101

Minimax Average MSE Designs for Estimating Mean Responses

Joong-Yang Park[1)]

Abstract

The unknown response function is usually approximated by a low order polynomial model. Such an approximation always accompanies bias due to model departure. The minimax Average MSE (AMSE) designs are suggested for estimating mean responses. A class of first order minimax AMSE designs is derived and a specific first order minimax AMSE design is selected from the class by optimizing the secondary criterion related to the power of the lack of fit test.

1. Introduction

A response function $\eta(\boldsymbol{x})$ is the relationship between the expectation of the response variable y and k independent variables $\boldsymbol{x} = (x_1, x_2, \cdots, x_k)'$. In practice the response function is either very complicated or unknown. Suppose that we wish to design an experiment for estimating mean responses over some specified experimental region S. Then an experiment is usually designed for a multiple linear regression model for $\eta(\boldsymbol{x})$, which can be written as

$$y = f(\boldsymbol{x})'\boldsymbol{\beta} + \varepsilon, \tag{1.1}$$

where $f(\boldsymbol{x})$ is a vector of functions of $\boldsymbol{x}$, $\boldsymbol{\beta}$ is an unknown parameter vector and ε is an error term with mean 0 and variance σ^2. This approach always causes some concerns about bias due to departure from model (1.1). A sensible design strategy is therefore to design an experiment for model (1.1) so that the precision of the least squares estimator of $f(\boldsymbol{x})'\boldsymbol{\beta}$ is robust against the model departure. In order to take the model departure into account, most of the previous works on robust designs assume that the true model is

$$y = f(\boldsymbol{x})'\boldsymbol{\beta} + z(\boldsymbol{x}) + \varepsilon,$$

1) Professor, Department of Statistics, Gyeongsang National University, Chinju, 660-701, Korea.

where $z(\boldsymbol{x})$ belongs to some specified class $\mathcal{F}$ of functions of $\boldsymbol{x}$. Refer to Weins (1992) for the discussion on several types of $\mathcal{F}$. The unknown response function is mostly approximated by some low order polynomial model. We thus assume that $f(\boldsymbol{x})'\boldsymbol{\beta}$ is a polynomial of order d. That is, $f(\boldsymbol{x})$ is the vector of p multiple monomials $\prod_{i=1}^{k} x_i^{\alpha_i}$ up to degree d. If $\eta(\boldsymbol{x})$ has $(d+1)$ continuous derivatives on S, then for each $\boldsymbol{x}\in S$ there exists $\boldsymbol{\gamma}$ such that $z(\boldsymbol{x})=\boldsymbol{h}(\boldsymbol{x})'\boldsymbol{\gamma}$ where $\boldsymbol{h}(\boldsymbol{x})$ is the vector of q multiple monomials of degree $(d+1)$. Furthermore $\boldsymbol{h}(\boldsymbol{x})'\boldsymbol{\gamma}$ is bounded above. We therefore assume that the true model is

$$y=f(\boldsymbol{x})'\boldsymbol{\beta}+\boldsymbol{h}(\boldsymbol{x})'\boldsymbol{\gamma}+\varepsilon, \tag{1.2}$$

and $\boldsymbol{h}(\boldsymbol{x})'\boldsymbol{\gamma}$ belongs to

$$\mathcal{F}=\{\boldsymbol{h}(\boldsymbol{x})'\boldsymbol{\gamma} \mid |\boldsymbol{h}(\boldsymbol{x})'\boldsymbol{\gamma}|\le\delta \text{ for all } \boldsymbol{x}\in S\}, \tag{1.3}$$

where δ is assumed known. It was shown by Park (1995) that class (1.3) is equivalent to $\Gamma=\{\boldsymbol{\gamma} \mid \boldsymbol{\gamma}'\boldsymbol{\gamma}\le\tilde{\delta}\}$ where $\tilde{\delta}=\delta^2/h_{\max}$ and $h_{\max}=\max_S \boldsymbol{h}(\boldsymbol{x})'\boldsymbol{h}(\boldsymbol{x})$.

We consider the problem of constructing the optimal designs for polynomial regression model $y=f(\boldsymbol{x})'\boldsymbol{\beta}+\varepsilon$ under the assumption that the true model is (1.2) where $\boldsymbol{\gamma}\in\Gamma$. The optimal designs depend on the shape of S and the corresponding design objective. Two types of S usually adopted are the hypercube $\{\boldsymbol{x} \mid |x_i|\le 1,\ i=1,2,\cdots,k\}$ and the hypersphere $\{\boldsymbol{x} \mid \boldsymbol{x}'\boldsymbol{x}\le 1\}$, which are respectively referred to as the cuboidal region and the spherical region. Both types of S are considered in this paper. The design objective of this paper is the estimation of mean responses. Box and Draper (1959) advocated the AMSE criterion for similar design problems. Refer to Park (1995) for some comments on the previous works based on the AMSE criterion and the minimax MSE designs for parameter estimation. We suggest in Section 2 the minimax AMSE optimality minimizing the maximum AMSE over Γ. Section 3 derives a class of first order minimax AMSE designs. Specific first order minimax AMSE designs are chosen in Section 4 by optimizing the secondary criterion that maximizes the power of the lack of fit test.

2. Optimal design problem

The design problem is to choose n, not necessarily distinct, design points in S. A design

ξ can be regarded as a probability measure on S defined by $\xi(\boldsymbol{x})=N(\boldsymbol{x})/n$ where $N(\boldsymbol{x})$ is the repetition of the design point $\boldsymbol{x}$. We extend this definition to include all probability measures on S. This is the so-called approximate design theory. Henceforth ξ and Ξ will denote an arbitrary probability measure and the set of all probability measures on S. Let

$$M_{11}(\xi)=\int_S f(\boldsymbol{x})f(\boldsymbol{x})'\,d\xi, \qquad M_{12}(\xi)=\int_S f(\boldsymbol{x})h(\boldsymbol{x})'\,d\xi, \qquad M_{22}(\xi)=\int_S h(\boldsymbol{x})h(\boldsymbol{x})'\,d\xi,$$

$\tilde{M}(\xi)=M_{11}^{-1}(\xi)M_{12}(\xi)$ and $\mu_{ij}=M_{ij}(\xi_u)$ where ξ_u is the uniform probability measure on S. Under the postulated model the mean response at an arbitrary point $\boldsymbol{x}$ in S is estimated by $f(\boldsymbol{x})'\boldsymbol{b}$, where $\boldsymbol{b}$ is the least squares estimator of $\boldsymbol{\beta}$. AMSE of $f(\boldsymbol{x})'\boldsymbol{b}$, denoted by $J(\boldsymbol{\gamma},\xi)$, is then obtained as

$$J(\boldsymbol{\gamma},\xi)=tr(\mu_{11}M_{11}^{-1}(\xi))+\frac{n}{\sigma^2}\boldsymbol{\gamma}'D(\xi)\boldsymbol{\gamma},$$

where $D(\xi)=\tilde{M}'(\xi)\mu_{11}\tilde{M}(\xi)-2\tilde{M}'(\xi)\mu_{12}+\mu_{22}$ and tr denotes the trace. We propose the designs that $\sup_\Gamma J(\boldsymbol{\gamma},\xi)=\inf_\Xi \sup_\Gamma J(\boldsymbol{\gamma},\xi)$. Such designs will be called the minimax AMSE designs. Park (1995) showed that it suffices to restrict our attention to $\Gamma_0=\{\boldsymbol{\gamma}\mid\boldsymbol{\gamma}'\boldsymbol{\gamma}=\delta\}$. The minimax AMSE designs are therefore the designs such that $\sup_{\Gamma_0} J(\boldsymbol{\gamma},\xi)=\inf_\Xi \sup_{\Gamma_0} J(\boldsymbol{\gamma},\xi)$. It can be further shown that

$$\sup_{\Gamma_0} J(\boldsymbol{\gamma},\xi)=tr(\mu_{11}M_{11}^{-1}(\xi))+\nu\, ch_{\max}(D(\xi))$$

where $\nu=n\delta/\sigma^2$ is the relative importance of bias versus variance and $ch_{\max}$ denotes the maximum characteristic root. Unfortunately due to the lack of convexity, it is almost impossible to solve the optimization problem directly. Noting that Γ_0 is invariant with respect to orthogonal transformations and following the discussion in Kiefer (1959), we further suggest that the minimax AMSE designs be selected from Ξ_0, the set of invariant designs. The optimization problem is now reduced to the minimization of $\sup_{\Gamma_0} J(\boldsymbol{\gamma},\xi)$ with respect to $\xi\in\Xi_0$.

3. Optimal first order minimax AMSE designs

In this section we derive the minimax AMSE invariant designs for $d=1$. The optimal dsigns for $k=1$ are first obtained. If $k=1$, then $f(\boldsymbol{x})'=(1, x_1)$, $h(\boldsymbol{x})=x_1^2$ and the two types of S under consideration are identical. $M_{ij}(\xi)$ of first order invariant designs and μ_{ij} are obtained as

$$M_{11}(\xi)=\begin{pmatrix}1 & 0\\ 0 & m_2\end{pmatrix}, M_{12}(\xi)=\begin{pmatrix}m_2\\ 0\end{pmatrix}, \quad \mu_{11}=\begin{pmatrix}1 & 0\\ 0 & a_2\end{pmatrix}, \quad \mu_{12}=\begin{pmatrix}a_2\\ 0\end{pmatrix} \text{ and } \mu_{22}=a_2, \tag{3.1}$$

where $m_2=\int_S x_1^2 d\xi$, $a_2=\int_S x_1^2 d\xi_u$ and $a_4=\int_S x_1^4 d\xi_u$. The first order minimax AMSE invariant designs ξ^* are therefore characterized by the second design moment m_2. Since $a_2=1/3$ and $a_4=1/5$,

$$\sup_{\Gamma_0} J(\gamma, \xi)=1+\frac{1}{3m_2}+\nu\left\{\left(m_2-\frac{1}{3}\right)^2+\frac{4}{45}\right\}. \tag{3.2}$$

Consequently if $\nu\leq 1/4$, then the optimal value of m_2, m_2^*, is 1. Otherwise, m_2^* is the soultion of

$$6\nu m_2^3-2\nu m_2^2-1=0, \tag{3.3}$$

which is obtained by equating the derivative of (3.2) with respect to m_2 to zero.

Suppose that $k\geq 2$. Then

$$f(\boldsymbol{x})'=(1, x_1, \cdots, x_k) \text{ and } h(\boldsymbol{x})=(x_1^2, \cdots, x_k^2, x_1x_2, \cdots, x_{k-1}x_k).$$

$M_{ij}(\xi)$ and μ_{ij} are given by

$$M_{11}(\xi)=\begin{pmatrix}1 & 0'\\ 0 & m_2 I_k\end{pmatrix}, \quad M_{12}(\xi)=\begin{pmatrix}m_2 1_k' & 0'\\ 0 & 0\end{pmatrix}, \tag{3.4}$$

$$\mu_{11}=\begin{pmatrix}1 & 0'\\ 0 & a_2 I_k\end{pmatrix}, \quad \mu_{12}=\begin{pmatrix}a_2 1_k' & 0'\\ 0 & 0\end{pmatrix}, \text{ and } \mu_{22}=\begin{pmatrix}(a_4-a_{22})I_k+a_{22}J_k & 0\\ 0 & a_{22}I_{k(k-1)/2}\end{pmatrix},$$

where $m_2=\int_S x_i^2 d\xi$, $a_2=\int_S x_i^2 d\xi_u$, $a_{22}=\int_S x_i^2 x_j^2 d\xi_u$, $a_4=\int_S x_i^4 d\xi_u$, I_k is the identity matrix of order k, $\mathbf{1}_k$ is the vector of ones and $J_k = k^{-1}\mathbf{1}_k\mathbf{1}_k'$. Therefore

$$D(\xi) = \begin{pmatrix} (a_4 - a_{22})I_k + (m_2^2 - 2a_2m_2 + a_{22})J_k & 0 \\ 0 & a_{22}I_{k(k-1)/2} \end{pmatrix}.$$

Consequently

$$\sup_{\Gamma_0} J(\gamma, \xi) = 1 + \frac{k}{3m_2} + \nu \max\left\{\frac{1}{9}, \frac{4}{45} + k\left(m_2 - \frac{1}{3}\right)^3\right\}$$

for the cuboidal region and

$$\sup_{\Gamma_0} J(\gamma, \xi) = 1 + \frac{k}{(k+2)m_2} + \nu \max\left\{\frac{2}{(k+2)(k+4)}, k\left(m_2 - \frac{1}{k} + 2\right)^2 + \frac{4}{(k+2)^2(k+4)}\right\}$$

for the spherical region. We now present theorems for finding m_2^*. It is not difficult to verify the theorems algebraically. Thus the proofs are omitted. The value of ν_{max} are tabulated in Table 1.

Theorem 1. Suppose that S is the cuboidal region. Let $\nu_{max} = \dfrac{45k\sqrt{5k}}{2(1+\sqrt{5k})^2}$, then m_2^* is

(i) 1 if $\nu \leq \frac{1}{4}$;

(ii) the solution of the equation (3.3) if $\frac{1}{4} \leq \nu \leq \nu_{max}$;

(iii) $\dfrac{\sqrt{5k}+1}{3\sqrt{5k}}$ if $\nu_{max} \leq \nu$.

Theorem 2. Suppose that S is the spherical region. Let $\nu_{max} = \dfrac{(k+2)^2(k+4)^{3/2}}{2^{3/2}(\sqrt{k+4}+\sqrt{2})^2}$, then m_2^* is

(i) $\frac{1}{k}$ if $k \leq 3$ and $\nu \leq \frac{k^3}{4}$;

(ii) the solution of equation $2(k+2)\nu m_2^3 - 2\nu m_2^2 - 1 = 0$ if $k \leq 3$ and $\frac{k^3}{4} \leq \nu \leq \nu_{max}$;

(iii) $\dfrac{\sqrt{k+4}+\sqrt{2}}{(k+2)\sqrt{k+4}}$ if $k \leq 3$ and $\nu \geq \nu_{max}$;

(iv) $\frac{1}{k}$ if $k \geq 4$.

Table 1. ν_{max} for the cuboidal and spherical regions

	ν_{max}	
k	cuboidal region	spherical region
2	8.2139	5.5692
3	11.0093	9.9311
4	13.4414	
5	15.6250	
6	17.6245	
7	19.4803	
8	21.2197	

4. Secondary criteria for first order minimax AMSE designs

The previous section derived the first order minimax AMSE invariant designs ξ^*, which are characterized by only the second design moment m_2. Therefore there are many optimal designs. This provides us with a room for applying other desirable optimality criteria. That is, we can consider a secondary criterion in selecting a specific first order ξ^*. Even though the minimax AMSE designs aim at protecting the experimenter from the worst possible model departure, it is still desirable to perform the lack of fit test. The power of the lack of fit test depends on the noncentrality parameter $\gamma' L(\xi)\gamma$, where

$$L(\xi) = M_{22}(\xi) - M_{12}(\xi)' M_{11}^{-1}(\xi) M_{12}(\xi). \tag{4.1}$$

Then it is reasonable to maximize $\gamma' L(\xi)\gamma$ in some sense. Atkinson and Fedorov (1975a) suggested T-optimality that maximizes $\gamma' L(\xi)\gamma$. Atkinson and Fedorov (1975b), Jones and Mitchell (1978) and DeFeo and Myers (1992) considered the criteria maximizing the minimum and average of $\gamma' L(\xi)\gamma$ over some specified region of γ. Since we are assuming that $\gamma'\gamma \le \delta$, we consider two criteria that maximize the minimum and average of $\gamma' L(\xi^*)\gamma$ over $\gamma'\gamma = \lambda$ for any $\lambda \le \delta$. The minimum and average of $\gamma' L(\xi^*)\gamma$ over $\gamma'\gamma = \lambda$ are $\lambda\, ch_{min}(L(\xi^*))$ and $\lambda q^{-1} tr(L(\xi^*))$. Thus the two criteria are equivalent to the maximization of $ch_{min}(L(\xi^*))$ and $tr(L(\xi^*))$, which will be referred to as T_1- and T_2-optimalities. There is another interpretation for the T_i-optimalities. Sometimes we may set δ to the largest value of $\gamma'\gamma$ that we can tolerate. That is, if $\gamma'\gamma \ge \delta$, the postulated model is considered as an inadequate model. In these circumstances the lack of fit test is mandatory and it is sensible to maximize the minimum or average of $\gamma' L(\xi^*)\gamma$ over the contour $\gamma'\gamma=\delta$. This approach also results in the T_i-optimalities.

Let m_{22} and m_4 denote the fourth design moments $\int_S x_i^2 x_j^2 d\xi$ and $\int_S x_i^4 d\xi$ for the invariant designs. Design moments up to degree four are concerned with the T_i- optimalities and only m_2, m_{22} and m_4 are nonzero. Denote the optimum values of m_{22} and m_4 by m_{22}^* and m_4^*. Since m_2 was determined already, m_{22} and m_4 will be determined by the secondary criteria. If $k=1$, $M_{11}(\xi^*)$ and $M_{12}(\xi^*)$ are given by (3.1) with m_2 replaced by m_2^* and $M_{22}(\xi^*)=m_4$. Therefore $ch_{\min}(L(\xi^*))=tr(L(\xi^*))=m_4-m_2^*$. Since $m_4\le m_2^*$, the value of m_4 maximizing $ch_{\min}(L(\xi^*))$ and $tr(L(\xi^*))$ is m_2^*. As a result the T_1-optimal first order ξ^* is identical with the T_2-optimal first order ξ^*. Equality of two design moments m_2 and m_4 implies that two vertices of S and the center point constitute the support of the T_i-optimal first order ξ^*. The weights for each vertex and the center point are respectively $m_2^*/2$ and $(1-m_2^*)$. According to the results given in the previous section, $m_2^*=1$ if $\nu\le 1/4$. Then the T_i-optimal first order ξ^* consists of only two vertices.

Only the cuboidal region is considered in the rest of this section. Suppose that $k\ge 2$. $M_{11}(\xi^*)$ and $M_{12}(\xi^*)$ of the invariant designs are then given by (3.4) with m_2 substituted with m_2^*. Corresponding $M_{22}(\xi^*)$ and $L(\xi^*)$ are

$$M_{22}(\xi^*)=\begin{pmatrix}(m_4-m_{22})I_k+m_{22}J_k & 0' \\ 0 & m_{22}I_{k(k-1)/2}\end{pmatrix} \qquad (4.2)$$

and

$$L(\xi^*)=\begin{pmatrix}(m_4-m_{22})I_k+(m_{22}-m_2^{*2})J_k & 0 \\ 0 & m_{22}I_{k(k-1)/2}\end{pmatrix}. \qquad (4.3)$$

Consequently

$$ch_{\min}(L(\xi^*))=\min\{(m_4-m_{22}),(m_4-m_{22})+k(m_{22}-m_2^{*2}),m_{22}\}$$

and

$$tr(L(\xi^*))=k(m_4-m_2^{*2})+\frac{k(k-1)}{2}m_{22}.$$

We first derive the T_2-optimal first order ξ^*. $tr(L(\xi^*))$ is maximized when both m_{22} and m_4 are as large as possible. Since $m_{22} \leq m_4 \leq m_2^*$, the optimum values of m_{22} and m_4 are therefore m_2^*. Corresponding maximum of $tr(L(\xi^*))$ is $km_2^*\{(k+1)-2m_2^*\}/2$. Next we consider the T_1-optimal first order ξ^*. The following theorem, which can also be proved easily, determines the values of m_{22} and m_4 maximizing $ch_{\min}(L(\xi))$ for given m_2.

Theorem 3. If the second design moment m_2 of a first order invariant design ξ is less than or equal to $1/2$, $ch_{\min}(L(\xi))$ achieves its maximum $m_2/2$ when $m_{22}=m_2/2$ and $m_4=m_2$. Otherwise, $ch_{\min}(L(\xi))$ achieves its maximum $m_2(1-m_2)$ when $m_{22}=m_4=m_2$.

If $m_2^* \geq 1/2$, the T_1-optimal first order ξ^* is identical to the T_2-optimal first order ξ^* since $m_2^*=m_{22}^*=m_4^*$ for both secondary criteria. We can obtain from equation (3.3) that $m_2^* \geq 1/2$ is when $\nu \leq 4$. Therefore, if ν seems to be less than or equal to 4, any of the two secondary criteria may be employed.

Finally we illustrate how to construct the T_i-optimal first order ξ^*. One common design moment condition required by the T_i-optimal first order ξ^* is $m_2^*=m_4^*$. This condition implies that the T_i-optimal first order ξ^* puts all mass on the points having coordinates 0 and ± 1 only, i.e., on A_j's where A_j denotes the set of points with j coordinates equal to 0 and the remaining coordinates equal to ± 1. If m_{22}^* is also equal to m_4^* and m_2^*, the points with positive mass are only A_0 and A_k, which are respectively the vertices and the center point. Therefore the T_1-optimal first order ξ^* for $m_2^* \geq 1/2$ and the T_2-optimal first order ξ^* are 2^{k-f} (fractional) factorial designs plus the center point. Corresponding weights allocated to each vertex and the center point are $m_2^*/2^{k-f}$ and $(1-m_2^*)$. If $\nu \leq 1/4$, $m_2^*=1$ by Theorem 1 and the support of the T_i-optimal first order ξ^* is composed of only the vertices of S. Theorem 3 shows that if $m_2^* \leq 1/2$, m_{22}^* of T_1-optimal first order ξ^* is $m_2^*/2$. Then, as in Galil and Kiefer (1977), we can always find a T_1-optimal first order ξ^* with mass $q_j/\binom{n}{j}2^{k-j}$ at each point of A_j by solving the equations

$$\sum_{j=0}^{k} q_j = 1, \quad \sum_{j=0}^{k} \frac{(k-j)}{k} q_j = m_2^* \quad \text{and} \quad \sum_{j=0}^{k} \frac{(k-j)(k-j-1)}{k(k-1)} q_j = m_{22}^*$$

subject to the constraints $q_j \geq 0$. A simple and well-known method is to use A_0, A_{k-1} and A_k, which are respectively the vertices, the $2k$ axial points $(\pm 1, 0, \cdots, 0)$, $\cdots, (0, \cdots, 0, \pm 1)$ and the center point. Such design are called the central composite designs. Corresponding weights for each vertex, the center point and each axial point are respectively $m_2^*/2^{k-f+1}$, $(1-m_2^*)$ and $m_2^*/(4k)$.

References

[1] Atkinson, A.C. and Fedorov, V.V. (1975a). The Design of Experiments for Discriminating Between Two Rival Models, *Biometrika,* Vol. 62, 57-70.

[2] Atkinson, A.C. and Fedorov, V.V. (1975b). Optimal Design: Experiments for Discriminating Between Several Models, *Biometrika,* Vol. 62, 289-303.

[3] Box, G.E.P. and Draper, N.R. (1959). A Basis for the Selection of A Response Surface Design, *Journal of the American Statistical Association,* Vol. 54, 622-654.

[4] DeFeo, P. and Myers, R.H. (1992). A New Look At Experimental Design Robustness, *Biometrika,* Vol. 79, 375-380.

[5] Galil, Z. and Kiefer, J. (1977). Comparison of Designs for Quadratic Regression on Cubes, *Journal of Statistical Planning and Inference,* Vol. 1, 121-132.

[6] Jones, E.R. and Mitchell, T.J. (1978). Design Criteria for Detecting Model Inadequacy, *Biometrika,* Vol. 65, 541-551.

[7] Kiefer, J. (1959). Optimal Experimental Designs (with discussion), *Journal of Royal Statistical Society Ser. B*, Vol. 21, 272-319.

[8] Park, J.Y. (1995). Minimum Mean Squared Error Invariant Designs for Polynomial Approximation, The Korean Communications in Statistics, Vol. 2, 376-386.

[9] Weins, D.P. (1992). Minimax Designs for Approximately Linear Regression, *Journal of Statistical Planning and Inference*, Vol. 31. 353-371.