Statistica Sinica **8**(1998), 527-533

EXACT A-OPTIMAL DESIGNS FOR QUADRATIC REGRESSION

Fu-Chuen Chang and Yu-Rong Yeh

National Sun Yat-sen University

Abstract: For quadratic regression on $[-1,1]$, exact n-point A-optimal designs are determined when $n \equiv 0,\ 1,\ 3 \pmod 4$. Some conjectures for the exact A-optimal designs when $n \equiv 2 \pmod 4$ are stated. For the case $[-a,a]$, $a>0$, the properties of exact A-optimal designs are discussed on grounds of an intensive numerical study.

Key words and phrases: A-optimal, exact design, quadratic regression.

1. Introduction

Consider the quadratic regression model

$$y(x) = \beta_0 + \beta_1 x + \beta_2 x^2, \qquad x \in [a,b],$$

and suppose that the experimenter takes n uncorrelated observations with expectation $y(x_1),\ldots,y(x_n)$ and variance $\sigma^2>0$. Such a vector $\mathbf{x}_n = (x_1,\ldots,x_n)$ is called an exact n-point design ξ_n. Denote

$$M(\xi_n) = \begin{pmatrix} \mu_0 & \mu_1 & \mu_2 \\ \mu_1 & \mu_2 & \mu_3 \\ \mu_2 & \mu_3 & \mu_4 \end{pmatrix},$$

where $\mu_k = \sum_{i=1}^n x_i^k/n$, $0 \le k \le 4$, and $\Xi_n = \{\xi_n : M(\xi_n) \text{ is nonsingular}\}$. A design ξ_n for which n_i components are equal to a_i, $1 \le i \le m$, $n_1+\cdots+n_m = n$, will be denoted by

$$\xi_n = \begin{pmatrix} a_1 & \cdots & a_m \\ n_1 & \cdots & n_m \end{pmatrix}.$$

Note that $\xi_n \in \Xi_n$ if and only if ξ_n has at least three different design points.

It is well known that for $\xi_n \in \Xi_n$ the covariance matrix of the least squares estimates $\hat\beta(\xi_n) = (\hat\beta_0(\xi_n),\hat\beta_1(\xi_n),\hat\beta_2(\xi_n))^T$ for the parameter vector $\beta = (\beta_0,\beta_1,\beta_2)^T$ is given by

$$\mathrm{Cov}(\hat\beta(\xi_n)) = \frac{\sigma^2}{n} M^{-1}(\xi_n) = \frac{\sigma^2}{n}(m^{ij}(\xi_n))_{0\le i,j\le 2};$$

in particular, $\mathrm{Var}(\hat\beta_i(\xi_n)) = \frac{\sigma^2}{n} m^{ii}(\xi_n), 0 \le i \le 2$, and $\sum_{i=0}^2 \mathrm{Var}(\hat\beta_i(\xi_n)) = \frac{\sigma^2}{n} tr\ M^{-1}(\xi_n)$. The purpose of this note is to determine the n-point A-optimal design ξ_n^* which minimizes the sum of variances of the estimators $\hat{\beta}_i(\xi_n)$, i.e.

$$tr\ M^{-1}(\xi_n^*) = \min_{\xi_n \in \Xi_n} tr\ M^{-1}(\xi_n).$$

In Section 2 we present the exact n-point A-optimal design for the case $[a, b] = [-1, 1]$.

The exact n-point D-optimal designs for quadratic regression were solved by Gaffke and Krafft (1982). Some numerically approximate ϕ_p-optimal designs for quadratic regression were studied by Preitschopf and Pukelsheim (1987). Constantine, Lim and Studden (1987) derived a necessary condition for an exact design to be admissible under the polynomial regression model. They showed that Salaevskii's conjecture (1966) holds true for cubic regression, but not for quartic regression. Recently Krafft and Schaefer (1995) obtained the exact n-point *EMM*-designs (Elfving-minimax designs) on the interval $[-1, 1]$. Imhof (1996) independently obtained the exact n-point A-optimal designs for quadratic regression when $n = 4k-1, 4k, 4k+1$ and $n = 4k+2$ $(k > 3)$.

2. n-Point A-Optimal Designs on $[-1, 1]$

We are going to prove the following theorem.

Theorem 2.1. *Let $n = 4k+q$, $k \in N$ and $q \in \{-1, 0, 1\}$. Then the exact n-point A-optimal designs for quadratic regression on $[-1, 1]$ are given by*

$$\xi_n^* = \begin{pmatrix} -1 & 0 & 1 \\ k & 2k-1 & k \end{pmatrix} \qquad \text{if } n = 4k-1,$$

$$\xi_n^* = \begin{pmatrix} -1 & 0 & 1 \\ k & 2k & k \end{pmatrix} \qquad \text{if } n = 4k,$$

$$\xi_n^* = \begin{pmatrix} -1 & 0 & 1 \\ k & 2k+1 & k \end{pmatrix} \qquad \text{if } n = 4k+1.$$

Two lemmas will be needed in the proof of Theorem 2.1.

Lemma 2.2. *If M and $\bar{M}$ are positive definite matrices of the form*

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \qquad \text{and} \qquad \bar{M} = \begin{pmatrix} M_{11} & 0 \\ 0 & M_{22} \end{pmatrix}$$

where $M_{11} \in R^{r \times r}$ and $M_{22} \in R^{s \times s}$, then $tr\ M^{-1} \geq tr\ \bar{M}^{-1}$, with equality if and only if $M_{12} = 0$.

The proof of Lemma 2.2 follows easily from the fact that

$$\begin{aligned} tr\ M^{-1} &= tr\ M_{11}^{-1} + tr\ A_1^T A_1 + tr\ M_{22}^{-1} + tr\ A_2^T A_2 \\ &\geq tr\ M_{11}^{-1} + tr\ M_{22}^{-1}, \end{aligned}$$

where $A_1=(M_{22}-M_{21}M_{11}^{-1}M_{12})^{-1/2}M_{21}M_{11}^{-1}$ and $A_2=(M_{11}-M_{12}M_{22}^{-1}M_{21})^{-1/2}$ $M_{12}M_{22}^{-1}$. For a reference of the inverse formula for a partitioned matrix, the reader is referred to Graybill (1983), Theorem 8.2.1.

Lemma 2.3. *Let S_n be the range of the first two moments of the n design points on $[0,1]$, i.e. $S_n = \{(\tau_1,\tau_2)|\tau_1 = \sum_{i=1}^n z_i/n, \tau_2 = \sum_{i=1}^n z_i^2/n, z_i \in [0,1], i = 1,\ldots,n\}$. Then $S_n = \{(\tau_1,\tau_2)|\tau_1^2 \le \tau_2 \le (k/n)+n(\tau_1-k/n)^2; k = \lfloor n\tau_1\rfloor\}$ where $\lfloor\alpha\rfloor$ is the greatest integer which is smaller than or equal to α.*

Proof. Obviously one has $0 \le \tau_1 \le 1$. Now fixing $\tau_1 = \tau_1^0$, τ_2 is a continuous function defined in the intersection of $[0,1]^n$ and the plane $z_1+z_2+\cdots+z_n = n\tau_1^0$ which is a convex region. Therefore, the intersection of S_n and the straight line $\tau_1 = \tau_1^0$ is a segment of the straight line $\tau_1 = \tau_1^0$. By Jensen's inequality one gets $\tau_2 \ge (\tau_1^0)^2$ and equality can be obtained for $z_1 = z_2 = \cdots = z_n = \tau_1^0$. Since τ_2 is convex, its maximum must be attained on the boundary of the domain. Suppose that the maximizer has k components 1, then it has $n-k-1$ components 0 and one component $x \in [0,1]$. From the relation $\tau_1^0 = k/n + x/n$ and $\max_{\tau_1=\tau_1^0}\tau_2 = k/n + n(\tau_1^0-k/n)^2$, Lemma 2.3 is proved.

Proof of Theorem 2.1. The proof of the case $n = 4k$ is trivial because it is well known that the approximate A-optimal design for quadratic regression on $[-1,1]$ is given by $\xi^*(\pm1) = 1/4$, and $\xi^*(0) = 1/2$, (Pukelsheim and Torsney (1991)). The proofs of the other two cases $n = 4k \pm 1$ are similar. Thus, we will only prove the case $n = 4k-1$ and defer the proof for the case $n = 4k+1$ to the Appendix.

Note that $tr\ M^{-1}(\xi_n)$ is invariant in permutation of regressors. Therefore, we may assume $f(x) = (x,1,x^2)^T$, and the information matrix

$$M(\xi_n) = \begin{pmatrix} \mu_2 & \mu_1 & \mu_3 \\ \mu_1 & \mu_0 & \mu_2 \\ \mu_3 & \mu_2 & \mu_4 \end{pmatrix}.$$

Putting $\mu_1 = \mu_3 = 0$ in M, one gets

$$\bar{M}(\xi_n) = \begin{pmatrix} \mu_2 & 0 & 0 \\ 0 & \mu_0 & \mu_2 \\ 0 & \mu_2 & \mu_4 \end{pmatrix}.$$

Take a square transformation $z = x^2$ which transforms a design ξ on [-1,1] to a design η_n on [0,1]. Then one has $\tau_1 = \sum_{i=1}^n z_i/n = \mu_2$, and $\tau_2 = \sum_{i=1}^n z_i^2/n = \mu_4$ where $z_i = x_i^2,\ 1,\ldots,n$. Let

$$\phi(\tau_1,\tau_2) = tr\ \bar{M}^{-1}(\xi_n) = \frac{1}{\mu_2} + \frac{\mu_4+1}{\mu_4-\mu_2^2} = \frac{1}{\tau_1} + \frac{\tau_2+1}{\tau_2-\tau_1^2}.$$

From Lemma 2.2 one has $tr\ M^{-1}(\xi_n) \geq tr\ \bar{M}^{-1}(\xi_n)$. To find the minimizer of $tr\ M^{-1}(\xi_n)$ over Ξ, it suffices to show that there exists a design ξ_n such that $tr\ M^{-1}(\xi_n) = \min_{(\tau_1,\tau_2)\in S_n} \phi(\tau_1, \tau_2)$ where the set S_n is given in Lemma 2.3 and the plots of S_n when $n = 5, 6, 7$ and 8, are depicted in Figure 1.

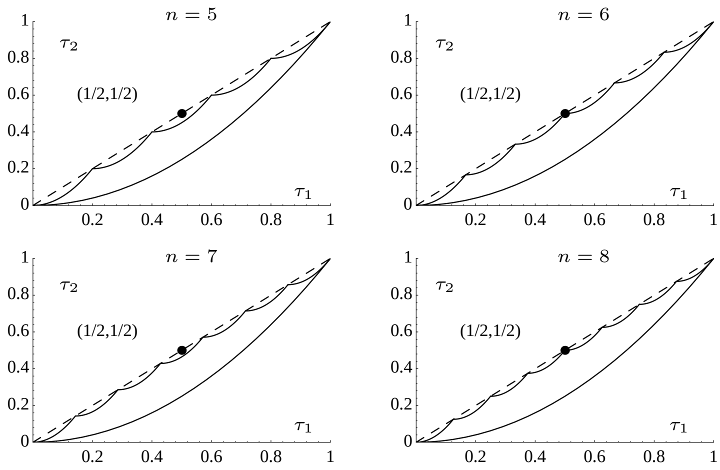

Figure 1. Moment set S_n on [0,1], for $n = 5, 6, 7$ and 8.

Note that the minimizer of $\phi(\tau_1, \tau_2)$ over S_n must occur on the arc $((m + x)/n, (m + x^2)/n)$, $0 \leq x \leq 1$, by Lemma 2.3 and the fact that for a fixed τ_1 and $\tau_2 \geq \tau_1^2$, $\phi(\tau_1, \tau_2)$ decreases in τ_2, and $\phi(\tau, \tau) = 2/[\tau(1-\tau)]$ decreases on $(0, 1/2]$, increases on $[1/2, 1)$. Now we are going to prove that $\min_{(\tau_1,\tau_2)\in S_n} \phi(\tau_1, \tau_2) = \phi(\tau_1^*, \tau_2^*)$ where $(\tau_1^*, \tau_2^*) = (m/n, m/n)$ or $((m + 1)/n, (m + 1)/n)$.

Plugging $(\tau_1, \tau_2) = ((m + x)/n, (m + x^2)/n)$ into $\phi(\tau_1, \tau_2)$, one gets

$$\varphi(x) = \frac{n[2mn + (n - m)x + (n + m - 1)x^2 + x^3]}{(m + x)[m(n - m) - 2mx + (n - 1)x^2]},$$

and $d\varphi(x)/dx = g(x)/h(x)$, where

$$\begin{aligned} g(x) = {} & n(-1 - 2m + 2n - n^2)x^4 + 2n(-m - 3m^2 + n + 2mn - n^2)x^3 \\ & +2mn(-3m^2 + 4n + mn - 3n^2)x^2 + 2m^2n(m - m^2 + 5n - n^2)x \\ & +m^2n(m^2 + 4mn - n^2), \end{aligned}$$

and $h(x) = (m + x)^2[m(n - m) - 2mx + (n - 1)x^2]^2 > 0$. Substituting $n = 4k - 1$ and $m = 2k - 1$, one obtains

$$\begin{aligned} g(x) = {} & -2(2k - 1)(4k - 1)^2x^4 - 4(3k - 1)(2k - 1)x^3 \\ & +2(4k - 1)(2k - 1)(52k^2 - 46k + 9)x^2 \end{aligned}$$

$$-4(2k-1)^2(4k-1)(10k^2-17k+4)x$$
$$+4(2k-1)^2(4k-1)(5k^2-5k+1),$$

and

$$\frac{d^2}{dx^2}g(x) = -24(2k-1)(4k-1)^2\left[x+\frac{3k-1}{2(4k-1)}\right]^2$$
$$-2(2k-1)(416k^3-499k^2+182k-21) < 0, \quad \text{for all } k \geq 1.$$

Moreover, for any $k \geq 1$, one has

$$g(0) = 4(2k-1)^2(4k-1)(5k^2-5k+1) > 0$$
$$\text{and} \quad g(1) = 4(1-4k)k^2(-1+4k)(-3+10k) < 0.$$

It follows that on $[0,1]$, $\varphi(x)$ is increasing first and then decreasing. Therefore, one gets $\min_{x\in[0,1]}\varphi(x) = \varphi(0) = \varphi(1) = (4k-1)^2/[k(2k-1)]$. Thus, for the optimal moments one has $\mu_2 = \mu_4 \in \{(2k-1)/n, 2k/n\}$ and $\mu_1 = \mu_3 = 0$. However $\mu_2 = \mu_4$ is only possible when $\mathbf{x} \in \{-1,0,1\}^n$ and since μ_1 vanishes, -1 and 1 must appear equally often in $\mathbf{x}$. The proof of the statement is complete.

Remark. For the case $n=4k+2$, the technique used in the proof of Theorem 2.1 does not go through. On grounds of an intensive numerical study, one observes that there are two optimal designs $\xi_n^* = \begin{pmatrix} -1 & x_0 & 1 \\ k & 2k+1 & k+1 \end{pmatrix}$ and $\begin{pmatrix} -1 & -x_0 & 1 \\ k+1 & 2k+1 & k \end{pmatrix}$ where x_0 is the unique zero in $(0,1)$ of

$$p(x) = (1+2k)x^4 - 4(1+2k)^2x^3 + 6(1+2k)x^2 - 4(1+8k+8k^2)x + (1+2k).$$

This conjecture with $k > 3$ is proved by Imhof (1996).

Remark. From a numerical study, all of the n-point A-optimal designs on $[a,b]$ appear to have three support points and contain at least one end point. If the design interval is $[-a,a], a > 0$, then the optimal designs seem to enjoy the following interesting properties:

1. When $a < a_n^*$, there is a unique symmetric optimal design supported at 0 and $\pm a$, except for a few cases like $n = 4k+2$.
2. When $a > a_n^*$, there are two asymmetric optimal designs:

$$\begin{pmatrix} u & v & a \\ 1 & n-2 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -a & -v & -u \\ 1 & n-2 & 1 \end{pmatrix},$$

   where $-a < u < 0 < v < a$ and $v < -u$.
3. When $a = a_n^*$, there are two optimal designs, a symmetric and an asymmetric one.

4. When a is close to a_n^* and $a < a_n^*$, the optimal designs have the form of $\begin{pmatrix} -a & 0 & a \\ 1 & n-2 & 1 \end{pmatrix}$.
5. The numerical values of a_n^* for some cases of n are given in Table 1 . One can see that a_n^*/n is greater than 1.135 and increases very slowly in n.

Table 1. Critical values a_n^*

n	3	4	5	6	7	8	9	10	20
a_n^*	3.405	4.608	5.876	7.174	8.486	9.807	11.133	12.461	25.804
a_n^*/n	1.135	1.151	1.175	1.196	1.212	1.226	1.237	1.246	1.290
n	30	40	50	60	70	80	90	100	200
a_n^*	39.171	52.542	65.915	79.289	92.663	106.038	119.413	132.787	266.528
a_n^*/n	1.306	1.306	1.318	1.321	1.324	1.325	1.327	1.328	1.333

A referee asked for the proofs of property 3 and the existence of $\lim_{n\to\infty} a_n^*/n$. Property 3 may be partially explained by properties 1-2 and the fact that $tr\ M^{-1}(\xi_n^*)$ is continuous in a if one scales the optimal design on $[-a_n, a_n]$ when a_n is close to a_n^* to a design on $[-a_n^*, a_n^*]$. The proof of the existence of $\lim_{n\to\infty} a_n^*/n$ seems to be impossible. More details on numerical results are given in Chang and Yeh (1996).

Acknowledgements

This research was supported by a grant from the National Science Council of the Republic of China. The authors are indebted to a referee for his constructive comments on a earlier version of this paper.

Appendix

Proof of the case $n = 4k + 1$ in Theorem 2.1.

Let $d\varphi(x)/dx = g(x)/h(x)$. If $n = 4k + 1$ and $m = 2k$, then

$$\begin{aligned} g(x) = & -4k(1+4k)^2x^4 - 4k(1+4k)(1+6k)x^3 + 4k(1+4k)(1-6k-52k^2)x^2 \\ & +16k^2(1+4k)(2+7k-10k^2)x + 4k^2(1+4k)(-1+20k^2), \end{aligned}$$

and

$$\begin{aligned} \frac{d^2}{dx^2}g(x) = & -48k(1+4k)^2\left[x + \frac{1+6k}{4(1+4k)}\right]^2 \\ & +k(11+20k-500k^2-1664k^3) < 0, \quad \text{for all } k \geq 1. \end{aligned}$$

Moreover, for $k \geq 1$, one has

$$g(0) = 4k^2(1+4k)(-1+20k^2) > 0 \text{ and } g(1) = -4k(1+2k)^2(1+4k)(1+5k) < 0.$$

It follows that on $[0,1]$, $\varphi(x)$ is increasing first and then decreasing. Thus, one gets $\min_{x\in[0,1]}\varphi(x) = \varphi(0) = \varphi(1) = (4k+1)^2/[k(2k+1)]$. Hence, the optimal moments $\mu_2 = \mu_4 \in \{2k/n, (2k+1)/n\}$ and $\mu_1 = \mu_3 = 0$. However $\mu_2 = \mu_4$ is only possible when $\mathbf{x} \in \{-1,0,1\}^n$ and since μ_1 vanishes, -1 and 1 must appear equally often in $\mathbf{x}$. The assertion of the theorem is proved.

References

Chang, Fu-Chuen and Yeh, Yu-Rong (1996). Exact A-optimal exact designs for quadratic regression. Technical Report. Dept. Applied Mathematics, National Sun Yat-sen University.

Constantine, K. B., Lim, Y. B. and Studden, W. J. (1987). Admissible and optimal exact designs for polynomial regression. *J. Statist. Plann. Inference* **16**, 15-32.

Gaffke, N. and Krafft, O. (1982). Exact D-optimum designs for quadratic regression. *J. Roy. Statist. Soc. Ser. B* **44**, 394-397.

Graybill, F. A. (1983). *Matrices with Applications in Statistics.* Wadsworth, Belmont, CA.

Imhof, L. (1996). A-optimum exact designs for quadratic regression. Technical Report. Institut für Statistik, RWTH Aachen.

Krafft, O. and Schaefer, M. (1995). Exact Elfving-minimax designs for quadratic regression. *Statist. Sinica* **5**, 475-484.

Preitschopf, F. and Pukelsheim, F (1987). Optimal designs for quadratic regression. *J. Statist. Plann. Inference* **16**, 213-218.

Pukelsheim, F. and Torsney, B. (1991). Optimal weights for experimental designs on linearly independent support points. *Ann. Statist.* **19**, 1614-1625.

Salaevskii, O. V. (1966). The problem of the distribution of observations in polynomial regression. *Proc. Steklov Inst. Math.* **79**, 147-166.

Department of Applied Mathematics, National Sun Yat-sen University, Kaohsiung, Taiwan.
E-mail: changfc@math.nsysu.edu.tw

(Received July 1996; accepted May 1997)