International Journal of Pure and Applied Mathematics
Volume 39 No. 1 2007, 83-89

D-OPTIMAL DESIGN WITH RANDOM BLOCK EFFECTS IN HETEROSCEDASTIC MODELS

Qiang Guo[1], Jian-Guo Liu[2 §], Da-Tian Niu[3]
[1,3]School of Science
Dalian Nationalities University
Dalian, 116600, P.R. CHINA
[2]Institute of System Engineering
Dalian University of Technology
Dalian, 116024, P.R. CHINA
e-mail: liujg004@yahoo.com.cn

Abstract: D-optimal regression designs under random block effects in heteroscedastic models are considered. After taking the homeomorphism transformation to the trace function, we give the condition to check the D-optimality of the design and the equation system of the positions and the powers of the design points. Finally, we give the numerical results and the efficiency analysis.

AMS Subject Classification: 65Y05, 65Y10, 49J35, 90C25, 90C30, 68W10, 68W25
Key Words: D-optimal, approximate design, information matrix, heteroscedastic

1. Introduction

D-optimal regression design is an important part of the experiment designs. The D-optimal regression designs in homogeneous models have been studied extensively. There have been lots of important achievements in this field. The Equivalence Theory of Kiefer and Wolfowitz [6] provides major impetus to the research of optimal designs. Most of literature on homogeneous models can be found in Pukelsheim [7]. Khuri [5] discussed the response surface models with random block effects, but did not consider the design aspect expect for

Received: April 25, 2007

§Correspondence author

orthogonal blocking. Cheng [3], Atkins and Cheng [1] studied the experiment described by Chasalow [2], and gave the explicit solution of the positions and the powers of the design points in homogeneous models. The heteroscedastic models are common in practice, which include homogeneous models. The traditional method is to use the homogeneous models to approach the heteroscedastic ones by taking logarithm to the latter. The shortcoming of the method is that the errors are large when the efficiency is near 1. Wong [8], Guo [4] derived some results on optimal regression designs for heteroscedastic models. In this article, we study the D-optimal designs with random block effects in heteroscedastic models directly and give the explicit construction of D-optimal design with blocks of size two for quadratic regression on $[-1, 1]$, and present an analytic proof of the D-optimality. In next section, we give the heteroscedastic regression model with blocks.

2. The Heteroscedastic Regression Model

Consider the optimization problem $\min_{x\in æ} \|\nu(x)\|^2$ based on the statistical model

$$y(x) = \mathbf{f}^{\mathrm{T}}(x)\beta + \nu(x), \tag{1}$$

where $x \in \omega \subset R^n$, the design region ω is a compact set in R^n, vector $\mathbf{f}(x) = (\mathbf{f}_1(x), \mathbf{f}_2(x), \cdots, \mathbf{f}_t(x))^{\mathrm{T}} \in R^t$, and $\beta = (\beta_1, \beta_2, \cdots, \beta_t) \in R^t$ is an unknown parameters vector, $y(x)$ is the response at the x-level of the independent variables. The errors $\nu(x)$ are independent by means zero and variances proportional to $1/\lambda(x)$, where $\lambda(x)$ is called the efficiency function (Pazman, 1986). We consider the heteroscedastic model, i.e., $\lambda(x)$ is not constant on ω.

We assume that there are N observations. Let $\mathbf{Y} = (y_1, y_2, \cdots, y_N)^{\mathrm{T}}$ be the vector of observations at x_1, x_2, $\cdots$, x_N (not necessarily all distinct). Then the former model can be expressed as

$$\mathrm{E}(\mathbf{Y}) = \mathbf{A}\beta = (a_{mn})_{N\times t}\beta, \tag{2}$$

$$\mathrm{cov}(\mathbf{Y}) = \mathrm{diag}_{m\,=\,1,\,2,\cdots,\,N}\left(\frac{\sigma^2}{\lambda(x_m)}\right), \tag{3}$$

where $a_{mn} = f_n(x_m)$, $m = 1, 2, \cdots, N$; $n = 1, 2, \cdots, t$. Assume that $\lambda_j(x)$, $\lambda_e(x)$ are known, and $\lambda(x)$ is constructed by $\lambda_e(x)$ and $\lambda_j(x)$. The variances of random errors are proportional to $\frac{1}{\lambda_e(x)}$, while the random block errors are proportional to $\frac{1}{\lambda_j(x)}$. Let the $[i + (j - 1)k]$-th entries of $y(x)$ is the i-th observation in the j-th block, where the block size is k. Then (3) can be expressed as following

$$\mathrm{cov}(\mathbf{Y}) = \mathrm{cov}(\nu(x)) + I_b \otimes \mathrm{cov}(\theta_j), \tag{4}$$

where $\mathrm{cov}(\nu(x)) = \sigma_e^2 \mathrm{diag}_{m=1,\cdots,N}(\frac{1}{\lambda_e(x_m)})$, $\mathrm{cov}(\theta_j) = \sigma_b^2 \mathrm{diag}_{i=1,\cdots,k;j=1,\cdots,b}\,(\frac{1}{\lambda_j(x_i+(j-1)k)})$.
Assume that $\sigma_b^2 = a\sigma_e^2$, where $a \in R^+$. Then one has

$$\mathrm{cov}(\mathbf{Y}) = \mathbf{B}(x)\sigma_e^2 = [\frac{1}{\lambda_e(x_{i+(j-1)k})} + \frac{a}{\lambda_j(x_{i+(j-1)k})}]\sigma_e^2.$$

Equation (4) covers the model $\mathbf{Y} = \mathbf{A}\beta + \theta + \nu(x)$, where $\theta = (\theta_1, \cdots, \theta_b)^{\mathrm{T}} \in R^{b,1}$ is a vector of random block effects, $\nu(x) \in R^{bk,1}$ is a vector of random errors, and $\mathrm{E}(\theta) = 0, \mathrm{E}(\nu(x)) = 0$, $\mathrm{cov}(\theta, \nu(x)) = 0$. The problem of determining optimal designs (selecting N points $x_1, \cdots, x_N$ and grouping them into b blocks of size k) for estimating the unknown parameters $\beta_1, \cdots, \beta_t$ will be considered. The covariance matrix of their generalized least square estimators $\widehat{\beta_1}, \cdots, \widehat{\beta_t}$ is equal to $\sigma_e^2[\mathbf{A}^{\mathrm{T}}\mathbf{B}^{-1}(x)\mathbf{A}]^{-1}$. A design is called *D-optimal* if it minimizes the determinant of $[\mathbf{A}^{\mathrm{T}}\mathbf{B}^{-1}(x)\mathbf{A}]^{-1}$. The inverse of $\mathbf{A}^{\mathrm{T}}\mathbf{B}^{-1}(x)\mathbf{A}$ is called the *information matrix* of the design. It is possible to study the efficiencies of such designs only if the optimal designs were determined.

3. D-optimal Designs with Random Block-Effects

When we use regression model to approach the response surface, the higher the order of the model is, the better the result is. In general, it is sufficient to use quadric model to approach response surface. To quadric model ($t = 2$), suppose there are relations $x_0 = -z_0$ and $y_0 = 0$ in the three D-optimal regression design points x_0, y_0, z_0, then we can obtain their value. The D-optimal designs with blocks of size two for quadric regression on $[-1, 1]$ in homogeneous models are presented by Cheng [3], Atkins and Cheng [1]. In their papers, the blocks of D-optimal design ξ^* are (x_0, k), $(z_0, -k)$, (z_0, x_0), the powers of blocks are ε, ε and $1 - 2\varepsilon$, respectively. Similar to it, we do the design of heteroscedastic model. So the design is denoted by ξ^*, too.

To unify the formulations of the optimal design points of different errors distribution combinations, we take homeomorphism transformation to the trace function $g(u, v; x_0, y_0, z_0)$, and give the mapping $T : (x_0, y_0, z_0) \mapsto (1, 0, -1)$, keeping the properties of the functions fixed. This leads into the operation of right scalar multiplication $(f\lambda)(x)$. For any convex function $f : R^n \to R^1$ and $\lambda \in [0, \infty)$, we have

$$(f\lambda)(x) = \lambda f(\lambda^{-1}x). \tag{1}$$

If the trace function $g(u,v)$ is convex, by (1) one has

$$(g\lambda)(u,v;x_0,y_0,z_0)=\lambda g(\lambda^{-1}u,\lambda^{-1}v;\lambda^{-1}x_0,\lambda^{-1}y_0,\lambda^{-1}z_0).$$

Let $\lambda^{-1}x_0=1,\lambda^{-1}y_0=0,\lambda^{-1}z_0=-1$. Since $x_0=-z_0$ and $y_0=0$, we have $\lambda=|x_0|$. Thus the mapping is

$$(g|x_0|)(u,v;x_0,y_0,z_0)=|x_0|g(|x_0|^{-1}u,|x_0|^{-1}v;1,0,-1). \tag{2}$$

Theorem 1. *Suppose $t=2$. For any efficiency function $\lambda(x)$, if $\mu(x)=1/\lambda(x)$ is even function on $[-1,1]$, $\lambda(x)$ is convex and $\mu'(x)$ is negative on $[0,1]$, there exists a k satisfying*

$$\begin{cases} (3\varepsilon^2-4\varepsilon+1)\mu(k)+(2\varepsilon-3\varepsilon^2)k^2\mu(1)=0, & (A)\\ 2k[2(1-\varepsilon)\mu(k)+(3k^2-1)\varepsilon\mu(1)]\mu(k) & \\ \qquad -(k^2-1)[(1-\varepsilon)\mu(k)+2k^2\varepsilon\mu(1)]\mu'(k)=0. & (B)\end{cases}$$

Set $D=(1-k^2)^2[(2\varepsilon-\varepsilon^2)k^2\mu(1)+(1-\varepsilon^2)\mu(k)]-2[(1-\varepsilon)\mu(k)+\varepsilon k^2\mu(1)][(1-\varepsilon)\mu(k)+\varepsilon k^4\mu(1)]/\mu(0)$. If $\frac{1}{3}<\varepsilon<\frac{1}{2}$ and $D>0$, there exists a D-optimal design ξ^* with weight ε, ε, $1-2\varepsilon$ on the points $(-1,k)$, $(-k,1)(k>0)$ and $(-1,1)$, respectively; otherwise, the design ξ_B with weight $\frac{1}{3}$ on each of the points $(-1,0)$, $(0,1)$ and $(-1,1)$ is D-optimal.

Proof. The inverse matrix of $\mathbf{M}(\xi^*)$ is proportional to

$$\begin{bmatrix} R & 0 & T\\ 0 & S & 0\\ T & 0 & U\end{bmatrix}. \tag{3}$$

Let $\mu(x)=1/\lambda(x)$, then we have

$$\begin{aligned} R=&\ [(1-\varepsilon)\mu(k)+\varepsilon k^2\mu(1)][(1-\varepsilon)\mu(k)+\varepsilon k^4\mu(1)]/(\mu(1)^2\mu(k)^2),\\ T=&\ -[(1-\varepsilon)\mu(k)+\varepsilon k^2\mu(1)]^2/(\mu(1)^2\mu(k)^2),\\ S=&\ \varepsilon(1-\varepsilon)(1-k^2)^2\mu(1)\mu(k)/(\mu(1)^2\mu(k)^2),\\ U=&\ [(1-\varepsilon)\mu(k)+\varepsilon k^2\mu(1)][(1-\varepsilon)\mu(k)+\varepsilon\mu(1)]/(\mu(1)^2\mu(k)^2).\end{aligned}$$

The trace function is

$$g(u,v)\doteq \operatorname{tr}\left[\begin{pmatrix}1 & u & u^2\\ 1 & v & v^2\end{pmatrix}\mathbf{M}^{-1}(\xi^*)\begin{pmatrix}1 & 1\\ u & v\\ u^2 & v^2\end{pmatrix}\begin{pmatrix}\lambda(u) & 0\\ 0 & \lambda(v)\end{pmatrix}\right]. \tag{4}$$

Let $v=\beta u$. Since $\mu(x)$ is even function, we have

$$g(u,\beta u) = R\{\frac{1}{\mu(u)} + \frac{1}{\mu(\beta u)}\} + (2T+S)\{\frac{1}{\mu(u)} + \frac{\beta^2}{\mu(\beta u)}\}u^2 + U\{\frac{1}{\mu(u)} + \frac{\beta^4}{\mu(\beta u)}\}u^4. \quad (5)$$

$g(u,\beta u)$ is convex since $U > 0$, $\lambda(u)$, u^4 and u^2 are all convex functions on $u \in [-1,0]$. Thus the maximum of $g(u,\beta u)$, $\beta \in [0,1]$ occurs either at a point of the form $(-1,-\beta)$ or at $(0,0)$. By the definition of trace function $g(u,v)$, one has that

$$[g(-1,-k) - g(0,0)]\mu^2(1)\mu^2(k) = (1-k^2)^2[(2\varepsilon-\varepsilon^2)k^2\mu(1) + (1-\varepsilon^2)\mu(k)] - 2[(1-\varepsilon)\mu(k) + \varepsilon k^2\mu(1)] [(1-\varepsilon)\mu(k) + \varepsilon k^4\mu(1)]/\mu(0). \quad (6)$$

Set $D = (1-k^2)^2[(2\varepsilon-\varepsilon^2)k^2\mu(1)+(1-\varepsilon^2)\mu(k)]-2[(1-\varepsilon)\mu(k)+\varepsilon k^2\mu(1)][(1-\varepsilon)\mu(k) + \varepsilon k^4\mu(1)]/\mu(0)$. If $D > 0$, we have $g(-1,-k) > g(0,0)$, then the maximum of $g(u,\beta u)$ is attained at $u = -1$; otherwise, the maximum is attained at $u = 0$. The maximum of $g(-1,-\beta)$ over $\beta \in [0,1]$ is obtained at $\beta = k(k > 0)$. If $\mu(x)$ is even function and $\mu'(1)$ is negative, it can be proved that $\mu'(-1) > 0$ and $\mu'(0) = 0$.

Then we have

$$\begin{cases} \frac{\partial g(-1,v)}{\partial v}\mid_{v=-1} & < \quad 0, \\ \frac{\partial g(-1,v)}{\partial v}\mid_{v=0} & = \quad \mu(0) - \mu'(0)R = \mu(0) > 0. \end{cases} \quad (7)$$

There must be a point $k \in (0,1)$ at which $\frac{\partial g(-1,v)}{\partial (v)}\mid_{v=-k}= 0$, i.e., $g(-1,-\beta)$ reaches its maximum at $\beta = k$. By symmetry theory, we have that $g(-1,-k) = g(k,1)$. It suffices to show that the maximum is obtained at $(-1,1)$ according to the multivariate version of equivalence theory. Our aim is to choose k which satisfy

$$\begin{cases} \frac{\partial}{\partial v} g(-1,v)\mid_{v=-k}= 0, \\ g(-1,1) = g(-1,-k). \end{cases} \quad (8)$$

Since

$$[g(-1,1) - g(-1,-k)]\mu(1)^2\mu(k)^2 = (1-k^2)^2[(-3\varepsilon^2 + 4\varepsilon - 1)\mu(k) + (3\varepsilon^2 - 2\varepsilon)k^2\mu(1)],$$

by the second formulation of (8), we have that

$$(3\varepsilon^2 - 4\varepsilon + 1)\mu(k) + (2\varepsilon - 3\varepsilon^2)k^2\mu(1) = 0. \quad (9)$$

We can obtain that $(3\varepsilon-2)(3\varepsilon-1)<0$ by combining (9) with $\lambda(x)>0$. Since $\varepsilon<\frac{1}{2}$, we have that $\varepsilon\in(\frac{1}{3},\frac{1}{2})$. On the other hand, one has

$$\frac{\partial g(-1,v)}{\partial v}\,|_{v=-k}=\frac{1}{\mu^2(k)}\{-[2(2T+S)k+4Uk^3]\mu(k) + \mu'(k)[R+(2T+S)k^2+Uk^4]\}. \quad (10)$$

By the first formulation of (8), we have following equation

$$2k[2(1-\varepsilon)\mu(k)+(3k^2-1)\varepsilon\mu(1)]\mu(k) + (1-k^2)[(1-\varepsilon)\mu(k)+2k^2\varepsilon\mu(1)]\mu'(k)=0. \quad (11)$$

(9), (11) respectively are (A), (B). □

4. Numerical Experiments

We take an example in the following to illustrate the Theorem 1. If $\lambda_j(x)=1$, $\lambda_e(x)=\exp\{cx^2\}(c>0)$, $\mu(x)=a+\exp\{-cx^2\}$, then we have

a	0.3	0.5	0.7	0.9	1.1	1.3
k	±0.4864	±0.4030	±0.3310	±0.2613	±0.1851	±0.0725
ε	0.3519	0.3473	0.3405	0.3397	0.3367	0.3338
D	-3.5755	-2.4576	-1.6579	-1.0422	-0.5259	-0.0809
eff	0.8593	0.9318	0.9691	0.9881	0.9971	1

Table 1: The corresponding ε, k, D to different a and $c=4$

In Table 1, $eff=\frac{|M(\xi_B)|}{|M(\xi^*)|}$, stand for the efficiency. The solutions that $k=0$, $\varepsilon=0.3333$ to every a are omitted. When $a=0.9$, $D<0$ and $\epsilon\in(\frac{1}{3},\frac{1}{2})$, so the D-optimal design ξ^* is the design with wight $0.3397, 0.3397, 0.3206$ on each of the blocks $(-1,-0.2613)$, $(0.2613,1)$ and $(-1,1)$, respectively. When $a=0.3$, comparing to the design ξ_B, the design ξ^* improve the efficiency 16.5/100, the design ξ^* is very efficient and necessary.

References

[1] J.E. Atkins, C.S. Cheng, Optimal regression designs in the presence of random block effects, *J. Statist. Plan. Infer.*, **77** (1999), 321-335.

[2] S.D. Chasalow, *Exact Optimal Response Surface Designs With Random Block Effects*, Ph.D. Thesis, University of California, Berkeley (1992).

[3] C.S. Cheng, Optimal regression designs under random block effect models, *Statist. Sinica*, **5** (1995), 485-497.

[4] Q. Guo, Z.Q. Xia, The choice of experiment design in heteroscedastic, *Sys. Engineering*, **115** (2003), 124-128.

[5] A.I. Khuri, Response surface models with random block effects, *Technometrics*, **34** (1992), 26-37.

[6] J. Kiefer, J. Wolfowitz, The equivalence of two extremum problems, *Canada J. Math.*, **14** (1960), 363-366.

[7] F. Pukelsheim, *Optimal Design of Experiments*, John Wiley, New York (1993).

[8] K.W. Wong, R.D. Cook, Heteroscedastic G-optimal designs, *J. R. Statist. Soc., B*, **4** (1981), 871-880.