

Minimizing Weighted Mean of Inefficiency for Robust Designs[†]

Han Son Seo¹⁾

Abstract

This paper addresses issues of robustness in Bayesian optimal design. We may have difficulty applying Bayesian optimal design principles because of the uncertainty of prior distribution. When there are several plausible prior distributions and the efficiency of a design depends on the unknown prior distribution, robustness with respect to misspecification of prior distribution is required. We suggest a new optimal design criterion which has relatively high efficiencies across the class of plausible prior distributions. The criterion is applied to the problem of estimating the turning point of a quadratic regression, and both analytic and numerical results are shown to demonstrate its robustness.

Keywords: Bayesian design; optimality; robustness.

1. Introduction

Assume that for an independent variable x a response y is measured in accordance with the statistical model

$$y = f(x, \theta) + e,$$

where f is an unknown function and the e 's are uncorrelated real-valued random variables having mean zero and constant variance. The experimenter is allowed to take N independent observations on y at $x_{(1)}, \dots, x_{(N)}$. The optimal design problem is : how should we select the x 's ?

Obviously the optimal design problem depends on the statistical function f . In addition, even when f is known, it may still depend on the unknown parameter θ . In fact, when we are interested in estimating a nonlinear combination of coefficients of a linear model, or, in general, in designing experiments for nonlinear models, the efficiency of a design depends on the values of the unknown parameters. The Bayesian approach

[†] This paper supported by Konkuk University in 2006.

1) Professor, Department of Applied Statistics, Konkuk University, 1 Hwayang-Dong, Gwangjin-Gu, Seoul 143-701, Korea. E-mail: hseo@konkuk.ac.kr

uses prior distributions for the unknown parameter in designing optimal experiment. But in reality there are several plausible prior distributions for a single parameter. In this situation the robustness of a Bayesian design becomes an important matter. There has been previous work on robustness of Bayesian optimal design. Dette (1990) suggested Bayesian optimal designs which are robust to model choice in polynomial regression. Seo (2002, 2006) deals with robustness aspects by specifying design assumptions in several prior distributions, one of which is assumed to be more favored than others. DasGupta and Studden (1991) defined three measures of robustness, (a) minimizing the maximum inefficiency, (b) minimizing the range of the Bayes risks and, (c) minimizing the diameter of the set of Bayes estimates. In this paper we propose a new Bayesian optimal design criterion by minimizing the weighted mean of inefficiency. The measure of likelihood of a prior distribution is expressed as assigned weight to the prior distribution.

The rest of the article is organized as follows. In Section 2 a general equivalence theorem is established which gives several alternative criteria for establishing that a particular design is optimal. Section 3 proposes the new criterion. Section 4 investigate the robustness of the proposed criterion by applying the criterion to the problem of estimating the turning point of a quadratic regression. Both analytic and numerical results are shown. Section 5 includes a summary of results and concluding remarks.

2. General Equivalence Theorem

Let Ξ be the set of all probability measures on X . We consider convex functions ϕ on Ξ and if we assume X is compact then Ξ is compact in weak convergence topology. An optimal design is a measure in Ξ which minimizes $\phi(\eta)$. In this section we present Whittle's (1973) general equivalence theorem, which gives several alternative criteria for optimality. Characterization of optimal designs can be formulated in several ways. For linear design it is convenient to characterize the optimal design in terms of properties of the information matrix (Silvey, 1980). For non-linear design, we characterize optimal designs directly in terms of properties of the design measures themselves. In particular, we use directional derivatives of the criterion with respect to design measures to characterize the optimal design.

Definition 2.1 *For two measures η_1 and η_2 in Ξ , the directional derivative at η_1 in the direction of η_2 is denoted by $F(\eta_1, \eta_2)$ and is defined, when the limit exists, by*

$$F(\eta_1, \eta_2) = \lim_{\varepsilon \downarrow 0} \varepsilon^{-1} [\phi((1 - \varepsilon)\eta_1 + \varepsilon\eta_2) - \phi(\eta_1)].$$

The extreme points of Ξ are the measures which put point mass at a single point x in X . We denote such a measure by η_x . If $F(\eta_1, \eta_2)$ is linear in η_2 then ϕ is said to be differentiable and

$$F(\eta_1, \eta_2) = \int F(\eta_1, \eta_2) \eta_2(dx).$$

We use $d(\eta, x)$ to denote $F(\eta, \eta_x)$. With this notation, we now give Whittle's general equivalence theorem which characterizes optimal designs for this set-up. Whittle (1973) presented this theorem in the context of linear optimal design. Chaloner and Larntz (1989) give additional conditions to permit use of the theorem for non-linear designs. We state the theorem in its extended version that applies to non-linear designs.

Theorem 2.1 (Whittle, 1973). *Assume that X is compact, that directional derivatives exist and are continuous in x , that there is at least one measure in Ξ for which ϕ is finite, and that ϕ is such that if $\eta_i \rightarrow \eta$ in weak convergence then $\phi(\eta_i) \rightarrow \phi(\eta)$.*

- (a) *If ϕ is convex, then a ϕ -optimal design η_0 can be equivalently characterized by any of the three conditions*
 - (i) η_0 minimizes $\phi(\eta)$,
 - (ii) η_0 maximizes $\inf_{x \in X} d(\eta, x)$,
 - (iii) $\inf_{x \in X} d(\eta_0, x) = 0$.
- (b) *The point (η_0, η_0) is a saddle point of F in that $F(\eta_0, \eta_1) \geq 0 = F(\eta_0, \eta_0) \geq F(\eta_2, \eta_0)$ for all $\eta_1, \eta_2 \in \Xi$.*
- (c) *If ϕ is differentiable, then the support of η_0 is contained in the set of X for which $d(\eta_0, x) = 0$ almost everywhere in η_0 measure.*

Proof: See Whittle (1973). □

3. Minimum Weighted Mean Of Inefficiency Design

Denote the prior distribution for θ as $\Delta(\theta)$ and let $\phi(\eta, \Delta)$ be a certain function of interest for the design η evaluated for prior distribution Δ . The goal of Bayesian design with the prior distribution Δ for the parameter θ is to minimize $\phi(\eta, \Delta)$ over η .

Definition 3.1 *If there exist a design η^* which minimizes $\phi(\eta, \Delta)$ among all designs η then η^* is called a B-optimal design (for the prior distribution Δ).*

To assess the relative worth of a design η against a B-optimal design for prior distribution Δ , we use the efficiency of a design defined by DasGupta and Studden (1991).

Definition 3.2 *The efficiency of a design η with respect to the prior distribution Δ is defined by*

$$EFF(\eta, \Delta) = \frac{\phi(\eta^*, \Delta)}{\phi(\eta, \Delta)},$$

where η^* is a B-optimal design for prior distribution Δ .

We assume that there are a finite number of prior distributions for θ , denoted as Δ_i $i = 1, \dots, n$, but we do not single out one as being more plausible than others. Under this situation we suggest a design criterion of minimizing the weighted mean of inefficiency to produce a robust design. Weight δ_i represents how much the prior Δ_i is favored to the others.

Definition 3.3 For given $\delta = \{\delta_1, \delta_2, \dots, \delta_n\}$, where $\delta_i \in (0, 1)$ and $\sum_{i=1}^n \delta_i = 1$, $\eta^\#$ is a B_δ -optimal design (for the set of prior distributions $\{\Delta = \Delta_1, \Delta_2, \dots, \Delta_n\}$) if $\eta^\#$ minimizes $\Phi^\delta(\eta, \Delta)$, defined by

$$\Phi^\delta = \sum_{i=1}^n \frac{\delta_i}{EFF(\eta, \Delta_i)}. \quad (3.1)$$

It can be easily verified that the directional derivative for the B_δ -optimal design criterion is expressed as

$$F(\eta_1, \eta_2) = \sum_{i=1}^n \frac{\delta_i}{\Phi(\eta_i^*, \Delta_i)} F_i(\eta_1, \eta_2), \quad (3.2)$$

where η_i^* is B-optimal design for prior distribution Δ_i and F_i is the directional derivative with the prior distribution for θ being Δ_i .

4. Application: Turning Point Problem

In this section the proposed design criterion $\Phi^\delta(\eta, \Delta)$ in (3.1) is applied to the turning point problem of a quadratic regression. Before we derive an optimal design we need to fix the criterion $\Phi(\eta, \Delta)$ of a B-optimal design in the definition 3.1.

For the specification of function $\Phi(\eta, \Delta)$ in Bayesian optimal designs we take usual squared error loss $L(\tilde{g}, g) = (\tilde{g} - g)^2$ as loss function for estimating $g(\theta)$ with $\tilde{g}(\theta)$. Then the usual criterion $\Phi(\eta, \Delta)$ for choosing an optimal design corresponds to the approximate expected posterior variance of $g(\theta)$. This criterion is generalized by Chaloner and Larntz (1989). When several functions of θ , $g(\theta) = (g_1(\theta), \dots, g_k(\theta))^T$, are of interest the criterion is expressed as the expected weighted trace of the product of a symmetric matrix and the inverse of the information matrix,

$$\phi^*(\eta) = \begin{cases} \mathbf{E}_\theta(\text{tr} B(\theta) M(\theta, \eta)^{-1}), & M(\theta, \eta) \text{ is nonsingular for all value of } \theta, \\ \infty, & M(\theta, \eta) \text{ is singular for all value of } \theta, \end{cases} \quad (4.1)$$

where $B(\theta) = C(\theta)C(\theta)^T$ is a symmetric p by p matrix, $C(\theta)$ is the p by k matrix with (i, j) the component $\partial g_j(\theta)/\partial \theta_i$ and M is the Fisher information matrix,

$$[M(\theta, \eta)]_{ij} = - \int (\partial^2 \log(P(y|\theta, x)/\partial \theta_i \partial \theta_j)) \eta(dx).$$

If we let $\Phi(\eta, \Delta) = \mathbf{E}_\theta(\text{tr} B(\theta) M(\theta, \eta)^{-1})$ as in (4.1) the key assumption in Whittle's theorem that the criterion function should be convex are easily verified. The design

criterion $\Phi^\delta(\eta, \Delta)$ for B_δ -optimal designs is a convex function of η in Ξ and the directional derivative of ϕ^* at η_1 in the direction of η_2 , where η_1 and η_2 , are measures in Ξ , is

$$F(\eta_1, \eta_2) = \phi^*(\eta_1) - \mathbf{E}_\theta(\text{tr} B(\theta) M(\theta, \eta_1)^{-1} M(\theta, \eta_2) M(\theta, \eta_1)^{-1}). \quad (4.2)$$

F_i in (3.2) also becomes

$$F(\eta_1, \eta_2) = \Phi(\eta_1, \Delta_i) - \mathbf{E}_\theta(\text{tr} B(\theta) M(\theta, \eta_1)^{-1} M(\theta, \eta_2) M(\theta, \eta_1)^{-1}). \quad (4.3)$$

The model we consider for the application of the optimal design criterion is the quadratic model,

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + e_i,$$

where $\beta = (\beta_0, \beta_1, \beta_2)$ are unknown coefficients and the errors e_i are independent, normally distributed with mean zero and variance σ^2 . The turning point of the model is defined as $\gamma = -\beta_0/2\beta_2$, the value of x at which the expected value of y is a maximum or minimum depending on the sign of β_2 . The design region X is restricted to the interval $[-1, 1]$ without loss of generality. It can be shown that the B -optimal design depends on β_0, β_2 and γ only through the first two moments of distribution of γ . Therefore the prior for γ is summarized as the vector, $\Delta = (m, v)$, where $m = \mathbf{E}(\gamma)$ and $v = \mathbf{V}(\gamma)$.

Theorem 4.1 (Mandal, 1978). *For a large sample size, if σ^2 and β_2 are assumed to be independent of γ then we have the following results.*

(a) *The design criterion becomes*

$$\phi(\eta, \Delta) = \frac{1}{d^2} \left[\frac{1}{\mu^2} + \frac{\frac{4}{d^2}v + \left\{ \frac{2}{d}(m - c) - \frac{1}{\mu_2}(\mu'_3 - \mu'_2\mu'_1) \right\}^2}{\mu'_4 - \mu'^2_2 - \frac{1}{\mu_2}(\mu'_3 - \mu'_2\mu'_1)^2} \right],$$

where

$x_{(1)}$ = minimum value of supporting points of a design η ,

$x_{(N)}$ = maximum value of supporting points of a design η ,

$$z = \frac{2x - x_{(1)} - x_{(N)}}{x_{(N)} - x_{(1)}}, \quad \mu'_r = \int z^r \eta(dz), \quad \mu_r = \int (z - \mu_1)^r \eta(dz),$$

$$c = \frac{x_{(1)} + x_{(N)}}{2}, \quad d = \frac{x_{(N)} - x_{(1)}}{2}.$$

(b) *A B -optimal design for the special case of the prior centered at 0, i.e. $\Delta = (0, v)$ is*

$$\eta^*(-1) = \eta^*(1) = \mu'^*_2/2 \text{ and } \eta^*(0) = 1 - \mu'^*_2,$$

where $\mu'^*_2 = \{1 + 2(v^{-1} + 4)^{-1/2}\}^{-1}$.

For further details on the B-optimal design, see Chaloner (1989) and El-Krunz and Studden (1991).

We now derive B-optimal designs analytically for the turning point problem for a set of prior distributions having means satisfying a particular condition. This condition is trivially met if all prior distributions have a common mean at the center of the design region.

Theorem 4.2 *For the turning point problem, consider prior distributions $\Delta_i = (m_i, v_i)$ for $v_i > 0$ and $m_i \in \mathbb{R}$ $i = 1, \dots, n$. For given $\delta = \{\delta_1, \delta_2, \dots, \delta_n\}$, where $\delta_i \in (0, 1)$ and $\sum_{i=1}^n \delta_i = 1$, if $\sum_{i=1}^n \delta_i \left(\prod_{t \neq i}^n \Phi(\eta_t^*, \Delta_t) \right) m_i = 0$, where η_t^* is B-optimal design for Δ_t , then the δ -mixture B $_\delta$ -optimal design (for the set of prior distributions $\Delta = \Delta_1, \Delta_2, \dots, \Delta_n$) is given by*

$$\eta^\#(-1) = \eta^\#(1) = \frac{1}{2} \left(1 + 2 \left(\frac{M}{L} + 4 \right)^{-\frac{1}{2}} \right)^{-1}$$

and

$$\eta^\#(0) = 1 - \left(1 + 2 \left(\frac{M}{L} + 4 \right)^{-\frac{1}{2}} \right)^{-1},$$

where

$$M = \sum_{k=1}^n \delta_k \left(\prod_{t \neq k}^n \Phi(\eta_t^*, \Delta_t) \right), \quad L = \sum_{i=1}^n \delta_i \left(\prod_{j \neq i}^n \Phi(\eta_j^*, \Delta_j) \right) (v_i + m_i^2).$$

Proof: We know that $\Phi(\eta, \Delta_i) = \int_{-\infty}^{\infty} g(\eta, \theta) \rho_i(\theta) d(\theta)$ for some function or g , where $\rho_i(\theta)$ is the probability density function of i -th prior distribution.

$$\begin{aligned} \Phi^\delta(\eta) &= \sum_{i=1}^n \delta_i \frac{\Phi(\eta, \Delta_i)}{\Phi(\eta_i^*, \Delta_i)} \\ &= \sum_{i=1}^n \frac{\delta_i}{\Phi(\eta_i^*, \Delta_i)} \int_{-\infty}^{\infty} g(\eta, \theta) \rho_i(\theta) d\theta \\ &= \int_{-\infty}^{\infty} g(\eta, \theta) \left\{ \sum_{i=1}^n \frac{\delta_i \rho_i(\theta)}{\Phi(\eta_i^*, \Delta_i)} \right\} d\theta \\ &= \int_{-\infty}^{\infty} g(\eta, \theta) \left\{ \sum_{i=1}^n \frac{\delta_i \left(\prod_{t \neq i}^n \Phi(\eta_t^*, \Delta_t) \right) \eta_i(\theta)}{\prod_{j=i}^n \Phi(\eta_j^*, \Delta_j)} \right\} d\theta \\ &= \frac{\sum_{k=1}^n \delta_k \left(\prod_{t \neq k}^n \Phi(\eta_t^*, \Delta_t) \right)}{\prod_{j=i}^n \Phi(\eta_j^*, \Delta_j)} \int_{-\infty}^{\infty} g(\eta, \theta) \frac{\sum_{i=1}^n \delta_i \left(\prod_{t \neq i}^n \Phi(\eta_t^*, \Delta_t) \right) \rho_i(\theta)}{\sum_{k=1}^n \delta_k \left(\prod_{t \neq k}^n \Phi(\eta_t^*, \Delta_t) \right)} d\theta. \end{aligned}$$

$\Phi^\delta(\eta)$ is proportional to $\int_{-\infty}^{\infty} g(\eta, \theta) \frac{\sum_{i=1}^n \delta_i \left(\prod_{t \neq i}^n \Phi(\eta_t^*, \Delta_t) \right) \rho_i(\theta)}{\sum_{k=1}^n \delta_k \left(\prod_{t \neq k}^n \Phi(\eta_t^*, \Delta_t) \right)} d\theta$.

Now

$$\frac{\sum_{i=1}^n \delta_i \left(\prod_{t \neq i}^n \Phi(\eta_t^*, \Delta_t) \right) \rho_i(\theta)}{\sum_{k=1}^n \delta_k \left(\prod_{t \neq k}^n \Phi(\eta_t^*, \Delta_t) \right)}$$

is another p.d.f of θ with expected value of

$$\frac{\sum_{i=1}^n \delta_i \left(\prod_{t \neq i}^n \Phi(\eta_t^*, \Delta_t) \right) m_i}{\sum_{k=1}^n \delta_k \left(\prod_{t \neq k}^n \Phi(\eta_t^*, \Delta_t) \right)} = 0$$

and variance of

$$\begin{aligned} & \frac{\sum_{i=1}^n \delta_i \left(\prod_{t \neq i}^n \Phi(\eta_t^*, \Delta_t) \right) \mathbf{E}_i(\theta^2)}{\sum_{k=1}^n \delta_k \left(\prod_{t \neq k}^n \Phi(\eta_t^*, \Delta_t) \right)} - \left(\frac{\sum_{i=1}^n \delta_i \left(\prod_{t \neq i}^n \Phi(\eta_t^*, \Delta_t) \right) m_i}{\sum_{k=1}^n \delta_k \left(\prod_{t \neq k}^n \Phi(\eta_t^*, \Delta_t) \right)} \right)^2 \\ &= \frac{\sum_{i=1}^n \delta_i \left(\prod_{t \neq i}^n \Phi(\eta_t^*, \Delta_t) \right) \mathbf{E}_i(\theta^2)}{\sum_{k=1}^n \delta_k \left(\prod_{t \neq k}^n \Phi(\eta_t^*, \Delta_t) \right)} \\ &= \frac{\sum_{i=1}^n \delta_i \left(\prod_{t \neq i}^n \Phi(\eta_t^*, \Delta_t) \right) (v_i + m_i^2)}{\sum_{k=1}^n \delta_k \left(\prod_{t \neq k}^n \Phi(\eta_t^*, \Delta_t) \right)^2} = \frac{L}{M}, \end{aligned}$$

where $\mathbf{E}_i(\theta^2)$ is the second moment of the i^{th} prior distribution. By applying Theorem 4.1 which gives the B -optimal design for a centered prior distribution, we have the theorem. \square

As noted above, when all prior distributions have their mean at the center of the design region, the above theorem applies.

Example 4.1 All four prior distributions having their mean at the center of the design region are considered. Four prior distributions are $\Delta_1 = (0, .03)$, $\Delta_2 = (0, .07)$, $\Delta_3 = (0, .15)$, $\Delta_4 = (0, .90)$. Given weights B_6 -optimal designs are derived by applying the Theorem 4.1. Table 1 displays efficiencies for B_6 -optimal designs for several sets of weights. To denote relative weights for each prior distribution we use integer numbers. For example (1234) represents assigning weight 1/10 to Δ_1 , 2/10 to Δ_2 , 3/10 to Δ_3 and 4/10 to Δ_4 .

$B_{(1234)}$ -optimal design could be considered a robust design compared to the B -optimal design for Δ_1 , Δ_2 and Δ_4 but not for Δ_3 . We use the term “robust” to mean the design loses relatively little compared to each B -optimal design, and gains a lot relative to some. Some sets of weights, for example, $B_{(1111)}$ and $B_{(4323)}$ give designs that are

Table 4.1: $\eta(0)$ (mass assigned to 0) and efficiencies of B_6 -optimal designs compared to B -optimal designs based on analytic results for prior distributions $\Delta_1 = (0, .03)$, $\Delta_2 = (0, .07)$, $\Delta_3 = (0, .15)$, $\Delta_4 = (0, .90)$.

B -design	δ	$\eta(0)$	EFFICIENCIES			
			$v_1(0.03)$	$v_2(0.07)$	$v_3(0.15)$	$v_4(0.90)$
$\eta^*(\Delta_1)$		0.2466000	1	0.9728118	0.9128410	0.7891656
$\eta^*(\Delta_2)$		0.3186651	0.9766429	1	0.9830795	0.9052679
$\eta^*(\Delta_3)$		0.3797959	0.9299636	0.9843830	1	0.9670393
$\eta^*(\Delta_4)$		0.4693980	0.8338256	0.9164029	0.9687717	1
B_6 -design	(1111)	0.3577687	0.9489643	0.9933891	0.9978927	0.9485570
	(1234)	0.3926280	0.9179255	0.9775745	0.9993099	0.9758818
	(4321)	0.3213290	0.9750357	0.9999675	0.9845667	0.9086486
	(4323)	0.3475527	0.9569943	0.9963334	0.9954362	0.9385517
	(2223)	0.3705265	0.9382272	0.9885998	0.9996317	0.9597732

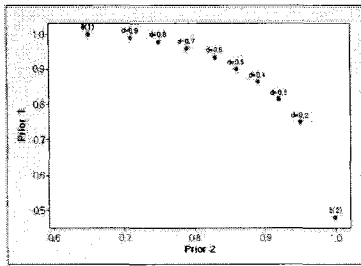
robust in the sense that extreme efficiencies are moderated at a small cost of efficiency for other prior distribution.

When the mean of a prior distribution is not the center of the design region, the Theorem 4.1 can not be applied. For those cases numerical methods are used to derive B_δ -optimal designs. We can derive directional derivatives of B - and B_δ -optimal designs in the direction of η_x , $d(\eta, x)$ from (3.2), (4.2) and (4.3) and verify that they are fourth degree polynomials. By the general equivalence Theorem 2.1-(a)-(iii), we know that the number of supporting points of optimal design is three. Simplex algorithm (Nelder and Mead, 1965) are used to find the best three points design.

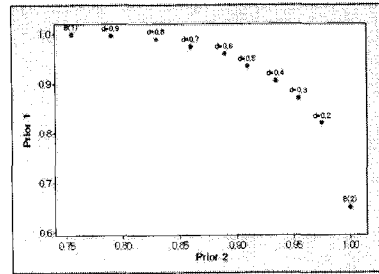
Example 4.2 For simplicity, we only work with two prior distributions. With two prior distributions, we consider three cases such that

- a) Both two prior distribution are certain, *i.e.*, have small variances : $\Delta_1 = (-.2, .07)$
 $\Delta_2 = (.5, .07)$.
- b) The one prior distribution is certain (small variance) but the other prior distribution is not (large variance) : $\Delta_1 = (-.2, .3)$ $\Delta_2 = (.5, .07)$.
- c) Neither of two prior distributions is certain : $\Delta_1 = (-.2, .30)$ $\Delta_2 = (.5, .30)$.

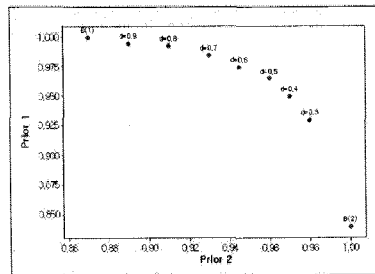
Figure 4.1 compares the efficiencies of B - and B_δ -optimal designs and indicates that the B_δ -optimal design criterion yields robust designs for all three cases with various values of weights.



(a)



(b)



(c)

Figure 4.1: Efficiencies of B_δ -optimal designs compared B -optimal designs based on numerical results for prior distributions (a) $\Delta_1 = (-.2, .07)$ $\Delta_2 = (.5, .07)$ (b) $\Delta_1 = (-.2, .3)$ $\Delta_2 = (.5, .07)$ (c) $\Delta_1 = (-.2, .30)$ $\Delta_2 = (.5, .30)$, $d = 0.1$ stands for B_δ -optimal designs with $\delta = (.1, .9)$.

5. Discussion

Robustness in the Bayesian optimal designs means maintaining reasonably high efficiencies for all competing prior distributions. For that purpose a new optimal design criterion is proposed. When it is applied to the turning point problem in the quadratic regression analytic results can be derived for centered-mean prior distribution. By calculating directional directive and using the general equivalence theorem the maximum number of design points and design's optimality can be verified. Numerical solutions for design criteria for non centered-mean prior distribution also show similarly good results. Robustness discussed in this paper can be applied to other optimality criteria. A natural extension would be applying the idea to the D -optimality.

References

- Chaloner, K. (1989). Bayesian design for estimating the turning point of a quadratic regression. *Communication in Statistics and Methods*, **18**, 1385–1400.
- Chaloner, K. and Larntz, K. (1989). Optimal Bayesian design applied to logistic regression experiments. *Journal of Statistical Planning and Inference*, **21**, 191–208.
- DasGupta, A and Studden, W. J. (1991). Robust Bayesian experimental designs in normal linear models. *The Annals of Statistics*, **19**, 1244–1256.
- Dette, H (1990). A generalization of D- and D1-optimal designs in polynomial regression. *The Annals of Statistics*, **18**, 1784–1804.
- El-Krunz, M. Sadi and Studden, W. J. (1991). Bayesian optimal designs for linear regression models. *The Annals of Statistics*, **19**, 2183–2208.
- Mandal, N. K. (1978). On estimation of the maximal point of a single factor quadratic response function. *Calcutta Statistical Association Bulletin*, **27**, 119–125.
- Nelder, J. A. and Mead, R. (1965). A simplex method for function minimization. *Computer Journal*, **7**, 308–313.
- Seo, H. S. (2002). Restricted Bayesian optimal designs in turning point problem. *Journal of the Korean Statistical Society*, **30**, 163–178.
- Seo, H. S. (2006). Mixture Bayesian Robust Design. *Journal of the Korean Society for Quality Management*, **34**, 48–53.
- Silvey, S. D. (1980). *Optimal Design*. Chapman & Hall/CRC, London.
- Whittle, P. (1973). Some general points in the theory of optimal experimental design. *Journal of the Royal Statistical Society, Ser. B*, **35**, 123–130.

[Received November 2007, Accepted December 2007]