Published in final edited form as:

J Korean Stat Soc. 2010 March 1; 39(1): 27–30. doi:10.1016/j.jkss.2009.10.006.

Discussion of Dette and Trampisch's paper "A general approach to D-optimal designs for weighted univariate polynomial regression models"

Weng Kee Wong, Ph.D.

Professor Weng Kee Wong, Department of Biostatistics, School of Public Health, UCLA, Los Angeles, CA 90095

Abstract

This is discussion on "A general approach to D-optimal designs for weighted univariate polynomial regression models" by Holger Dette and Matthias Trampisch.

Keywords

Polynomial regression models; D-optimal designs; G-optimal designs; heteroscedasticity; equivalence theorem

I congratulate the authors for providing an in-depth and up-to-date review of design issues for D-optimal designs for polynomial regression models with heteroscedastic errors. The authors provided new insights of the design problem with a unifying theory that I believe few can accomplish. I like the overarching thematic view and the work herein is an elegant demonstration of how mathematics can be applied to solve a well formulated statistical problem. The availability of an analytic description of the optimal design is of course most helpful for studying properties of the optimal design and in particular, its sensitivities to model assumptions. Because the optimal design is found from a fully parametric model, it is crucial to understand its robustness properties to model assumptions before implementation.

A very common criticism levied at optimal design problems is the assumption of a fully specified statistical model. In almost all cases of practical interest, the optimal design depends on the model assumptions and sometimes in a very sensitive way. This means that the D-optimal design can perform poorly for estimating all the parameters in a slightly different model. For example assuming we have homoscedastic errors, the D-optimal design for the simple liner model has zero efficiency for estimating the three coefficients in the quadratic model. The practical implication is that it can be risky to use an optimal design when there is model uncertainty and so it is good statistical practice to always check whether the optimal design is robust to important model assumptions. Indeed, any prudent user should automatically investigate robustness of the optimal design to each model assumption before implementation. If necessary, as is often the case in practice, the user should modify the optimal design, for example by adding more points, to meet other goals or concerns the

Corresponding Author: Dr Weng Kee Wong, Ph.D.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

user may have. The user may also want to modify the design weight at each point if it makes more sense for the problem at hand. The role of the optimal design is then one for calibration purposes, guiding the user on how far to stray away from the optimal design without incurring an unacceptable loss in efficiency. This important aspect about optimal design seems to be often overlooked in discussion; an optimal design provides just a guide for the user and it is expected that the user may have to modify the optimal design as the practical problem dictates. This should not be surprising because we replace the real design problem by a mathematical optimization problem and sometimes the solution to the mathematical problem may not meet all the practical requirements of the design.

In practice, we must make model assumptions before design issues can be seriously considered. There is no way around it. Frequently, uniform designs are used in practice to circumvent the model assumptions. Whether this is a successful and adequate design strategy depends on the true underlying model. Uniform designs are by no means a panacea because we also have to decide how many equally spaced points the design should have. Too many points can result in an inefficient design, see, for example, Wong and Lachenbruch (2005).

When heteroscedasticity is present, the design problem becomes more complicated. The technical difficulties can be challenging and sometimes painstaking effort has to be taken to partition the problem cleverly into suitable parts and solving them separately. We see such illustrations in this paper where the heteroscedastic D-optimal design is found on different segments of the design interval. In addition to having to deal with an invariably harder theoretical optimization problem when heteroscedasticity is present, the assumption of a known efficiency function is always subject to strong criticism. In fact, many view the assumption of a known efficiency function to be highly unrealistic in practice, and perhaps rightly so. The fact that many heteroscedastic D-optimal designs with analytical descriptions relied on efficiency functions chosen to make the mathematics work further adds to the criticisms.

The typical response seems to be that we design assuming the best guess for the efficiency function and hope that the resulting optimal design is robust to small changes in the assumed efficiency function. The efficiency function usually has unknown parameters and so we can hope that small changes in the assumed nominal parameter values do not result in a very different optimal design. We can only imagine that if the functional form of the efficiency function is mis-specified in a serious way, as when a monotonic increasing function is assumed when a decreasing function is more appropriate, the optimal design may take on a very different form. In summary, the assumption of a known efficiency function is necessary to find the optimal design but such an assumption is tenuous at best in practice. So how do we design for a heteroscedastic model without assuming a fully known efficiency function?

One approach and perhaps the first step to consider in the design for such problems is to study what I term as the critical number that each efficiency function seems to possess. This number indicates a threshold value beyond which heteroscedasticity does not matter as much from the design perspective, as long as its form is correctly specified. Specifically, let $(X, f(x), \lambda(x))$ be a given design problem where X is the known design space, $f(x)$ is the vector of known regression functions and $\lambda(x)$ is the assumed known efficiency function. The critical number of the efficiency function $\lambda(x)$ is the value c for which the D-optimal design for the problem $(X, f(x), c + \lambda(x))$ is the same as the D-optimal design for the design problem $(X, f(x), \lambda(x)=1)$. A large positive value of c indicates there are relatively less overall differences in variability in the responses across the design interval and so the design problem resembles one for designing under a homoscedastic model. As an example, suppose one assumes a simple linear model, the inverse variance of the response varies linearly

across the design interval $[-1, 1]$ and the rate of change in variability is unity. We may then assume the efficiency function is $\lambda(x) = x + c$ for some c . We do not know the value of c , but there is a value of c for which the D-optimal design for this problem is the same as the one with $\lambda(x) = 1$. We call this value of c the critical value of this $\lambda(x)$ for the linear model. For this particular $\lambda(x) = x + c$, we can verify directly using the equivalence theorem for D-optimality that the critical value is 2. The practical implication is that from the design perspective, we do not need to know the efficiency function $\lambda(x)$ completely; as long as we know that c is at least 2, then the D-optimal design for the problem $(X, f(x), \lambda(x) = x + c)$ is equally supported at both ends of the design interval $[-1, 1]$. Critical numbers for other efficiency functions are given in Table 1 for the simple linear model on the design interval $[-1, 1]$. The efficiency functions have different shapes and include symmetric and monotonic ones. Critical values for efficiency functions under a quadratic polynomial model and other types of regression models also exist.

This phenomenon is observed for other types of optimal designs, including G-optimal designs that are usually used for surface response estimation. Such an optimal design minimizes the maximum variance of the fitted response across the design space and is a special type of minimax designs. The G-optimality criterion is also a convex function in the designs but it is not differentiable. This means that we have to work with the sub-gradient of the design criterion and consequently the equivalence theorem for verifying G-optimality is more complicated, see Wong and Cook (1993), for example. Some critical values for a few efficiency functions when we have G-optimality and a simple linear model are displayed in Table 1. From the table, we note that (i) the critical numbers of the efficiency functions coincide for D and G-optimality in all our examples and (ii) some of the results in Table 1 are taken from Wong (1995).

In summary, critical number seems to have a useful role in the design for a model with heteroscedastic errors. Once this number is known, either via estimation or otherwise, the precise specification of the efficiency function is no longer required. This should help provide a partial answer to the frequently critique of having to assume a known efficiency function for the construction of an optimal design for a model with heteroscedastic errors. I am interested to know whether the authors have done work in this area and if not, how to better address this often critique of having to assume a completely known efficiency function in the construction of an optimal design.

Another interesting aspect of heteroscedastic D-optimal design is that the design itself can be non-intuitive. We first give a simple example to show that D-optimal design seems to make sense. Consider the linear polynomial model on the design interval $[-1, 1]$ with the first efficiency function in Table 1, where $\lambda(x) = c - x^2$ and $c > 1$. This suggests that we expect more variability in the response when we take observations near the extreme points of the design space. The value of c indicates the level of heteroscedasticity across the design interval. As c gets smaller and approaches 1, this variability gets arbitrarily large when we take observations at -1 and 1 . It then makes sense to back away from taking observations at -1 and 1 , the extreme points of the design space, and the theory tells us precisely how much to back away from -1 and 1 as c gets smaller. How much to back away from the endpoints? A direct calculation shows that the D-optimal design is equally supported at the extreme points if c is at least 3; otherwise it is equally supported at $-\sqrt{c/3}$ and $\sqrt{c/3}$ (Table 1). Is this design reasonable? The mathematics tells us nicely we should shift the two design points for the D-optimal design away from the endpoints and symmetrically to $-\sqrt{c/3}$ and $\sqrt{c/3}$ when the value of c decreases to 3 or less, indicating the amount of variability at the endpoints has increased. When there is smaller variability in the response across the design space, as when c is at least 3, the D-optimal design requires that we take equal number of observations at -1

and 1, which is the optimal design strategy when we have homoscedastic errors in the simple linear model.

Consider now the same setup except that $\lambda(x) = 2 + \cos(cx)$ and $c = 1, 2$ and 4 on the interval $[-1, 1]$. Figure 1 shows these efficiency functions have different shapes. How many support points does the D-optimal design have for each of these efficiency functions? The D-optimal designs for these models are displayed in Table 2. All are symmetric with one supported at 3 points and two supported at 2 points. The efficiency functions do not seem sufficient by themselves to fully explain the number of design points required in the D-optimal design. My experience with numerous other examples tells me that the number and distribution of design points can be non-intuitive in a D-optimal design and it is generally difficult to make good informed guesses on the structure of the D-optimal design based solely on the efficiency function. I would like to know whether the authors have suggestions for predetermining the number of design points and roughly how the design points are distributed based on the efficiency function alone. A better understanding of this relationship in a heteroscedastic D-optimal design problem will allow us to guess its form more accurately and find the D-optimal design quickly using the equivalence theorem.

A final question I have for the authors pertains to the equivalence theorem. The equivalence theorem is a powerful tool for verifying if a design D-optimal, and more generally for verifying any optimal design provided the criterion is convex over the set of designs in the given design space. Given a candidate design, we plot the graph of the directional derivative of the criterion and observe whether the graph lies below the horizontal axis at 0, with equality at the support points of the candidate design. Further, the equivalence theorem tells us that we can provide a lower bound for the efficiency of the candidate design by examining the height of the peak point in the graph above the horizontal axis. My experience is that the graph below the horizontal axis can vary greatly in shape, slope and depth across the design interval. I would like to know whether any useful information for the practitioners can be gained from the characteristics of the graph below the horizontal axis. I welcome the authors' comments on this aspect of the plot.

References

- Wong WK, Cook RD. Heteroscedastic G-optimal Designs. *J R S S B* 1993;55:871–880.
- Wong WK. On the Equivalence of D and G-optimal Designs in Heteroscedastic Models. *Statistics & Probability Letters* 1995;25:317–321.
- Wong, WK.; Lachenbruch, PA. Designing Studies for Dose Response in Tutorials in Biostatistics. In: Agostino, RB., editor. *Statistical Methods in Clinical Studies*. Vol. 1. John Wiley; 2005.

Figure 1.
Efficiency functions: $\lambda(x) = 2 + \cos cx$, $c=1, 2, 4$ on the design interval $[-1, 1]$.

Critical values of selected efficiency functions for D and G-optimal designs for the simple linear model on $[-1, 1]$. All optimal designs are equally supported except for the G-optimal design in (5) with mass at 1 equal to $(c-1)/2c$ if $c \geq 2$ and $(c/(c+2))^2$ if $1 < c \leq 2$.

Table 1

$\lambda(x)$	support points of D-optimal designs	support points of G-optimal designs
1 $c-x^2$	-1 1	-1 1 if $c \geq 3$ -a a if $1 < c \leq 3$
	$-\sqrt{c/3}$ $\sqrt{c/3}$	
		$a = \sqrt{1+c-1}$
2 $c- x $	-1 1	-1 1 if $c \geq 2$ -a a if $1 < c \leq 2$
	-c/2 c/2	
		$a = s - 1/s; s^3 = c + \sqrt{1+c^2}$
3 $\exp(-cx^2)$	-1 1	-1 1 if $c \leq 2$ -a a if $c > 2$
	$-\sqrt{1/(2c)}$ $\sqrt{1/(2c)}$	
		$a = \sqrt{\frac{c^2 + 4c - c}{2c}}$
4 $1/(x+c)$	-1 1	$(c+1)/2c$ $(c-1)/2c$ if $c > 1$
5 $x+c$	-1 1	-1 1 if $c \geq 2$ -c/2 1 if $1 < c \leq 2$
	$\frac{1-2c}{3}$	

Table 2

D-optimal designs for the linear model on $[-1, 1]$ for the efficiency function $\lambda(x) = 2 + \cos(cx)$ and $c = 1, 2, 4$.

$\lambda(x)$	support points (weights) of D-optimal design		
$2 + \cos(x)$	-1 (0.5)	1 (0.5)	
$2 + \cos(2x)$	$-s$ (0.5)	s (0.5); s solves $2 + \cos 2s - 2s \sin 2s = 0$; $s \approx 0.90653$	
$2 + \cos(4x)$	-1 (p)	0 (1-2p)	1 (p); $p = 0.75/(1 - \cos 4) \approx 0.4535$