Biometrika (1992), **79**, 4, *pp.* 847–9
Printed in Great Britain

A note on quadratic designs for nonlinear regression models

BY TIMOTHY E. O'BRIEN

Department of Statistics, North Carolina State University, Raleigh, North Carolina 27695, U.S.A.

SUMMARY

Hamilton & Watts (1985) state that their quadratic design criterion results in designs with only p points of support. In this note we demonstrate how their design criterion may be used to obtain designs with $p+1$ design points, thus allowing for a check of the adequacy of the assumed model.

Some key words: Intrinsic nonlinearity; Model misspecification; Nonlinear regression; Optimal design; Parameter-effects nonlinearity.

1. INTRODUCTION

The first real advance in nonlinear design methodology was given by Box & Lucas (1959) with the introduction of a local D-optimal design criterion. Local D-optimal designs minimize the volume of the linear approximation to the exact confidence region for θ, typically at some initial estimate, $\hat{\theta}$. Demonstrating that this first-order approximation can often be quite poor, Hamilton & Watts (1985) introduce a second-order volume approximation. Whereas D-optimal designs work with a tangent plane approximation, quadratic designs have the distinct advantage of taking into account the curvature of the expectation surface.

Although Atkinson (1988) and Chaloner & Larntz (1989) give examples of Bayesian D-optimal designs with more than p support points, Atkinson & Hunter (1968), M. J. Box (1968) and Vila (1991) give evidence that in most practical situations non-Bayesian D-optimal designs typically result in replicates of only p support points. Further, Hamilton & Watts state that their quadratic criterion also results in replicates of only p points regardless of the value of n. A major disadvantage of designs with only p support points is that these designs provide no ability to check for lack of fit of the hypothesized model, a concern raised by Box & Lucas (1959) and Cochran (1973).

The purpose of this note is to demonstrate how the quadratic design criterion of Hamilton & Watts may be used to obtain optimal designs with $p+1$ support points. An example is given to illustrate this technique.

2. BACKGROUND

For the expectation function, $E(y)=\eta(\theta)$, where θ is a $p\times 1$ parameter vector, $\hat{V}$ and $\hat{W}$ are the first and second derivative arrays of η with respect to θ evaluated at $\hat{\theta}$, having dimensions $n\times p$ and $n\times p\times p$, respectively. Also, $\hat{V}$ has the QR decomposition $\hat{V}=\hat{Q}\hat{R}=[\hat{U}\,|\,\hat{N}]\hat{R}=\hat{U}\hat{L}^{-1}$, where the columns of $\hat{U}$ form an orthonormal basis for the tangent plane to the expectation surface at $\eta(\hat{\theta})$, and the columns of $\hat{N}$ form an orthonormal basis for the space orthogonal to the tangent plane. Using a quadratic approximation, Hamilton & Watts (1985) show that the volume of a $100(1-\alpha)\%$ confidence region for θ is approximately equal to

$$v=c|\hat{V}'\hat{V}|^{-\frac{1}{2}}|\hat{D}|^{-\frac{1}{2}}\{1+k^2\operatorname{tr}(\hat{D}^{-1}\hat{M})\}, \tag{1}$$

where c and k are constants relative to the design, $\hat{M}$ is a function of parameter effects curvature, and $\hat{D}$ measures the intrinsic curvature in the direction of the residual vector, $\hat{\varepsilon}$. Here $\hat{D}=I_p-\hat{B}$, and $\hat{B}=\hat{L}'[\hat{\varepsilon}'][\hat{W}]\hat{L}$. Claiming that (1) cannot be used as a design criterion as it requires knowledge of the unknown residuals, the authors replace the residual vector in (1) by a vector of zeros, so that $\hat{D}=I_p$, and obtain a criterion that seeks designs which minimize the volume

$$v'=c|\hat{V}'\hat{V}|^{-\frac{1}{2}}\{1+k^2\operatorname{tr}(\hat{M})\}. \tag{2}$$

Such designs are functions of $\hat{\theta}$, σ, and of the particular parameterization used and ignore the intrinsic nonlinearity of the expectation surface.

Hamilton & Watts illustrate the use of their criterion with the intermediate product model function

$$\eta(\theta, x_i) = \theta_1\{\exp(-\theta_2 x_i) - \exp(-\theta_1 x_i)\}/(\theta_1 - \theta_2) \quad (\theta_1, \theta_2, x_i > 0), \tag{3}$$

using initial parameter values $\hat{\theta} = (0{\cdot}7, 0{\cdot}2)'$. In addition to the original form (3), the authors also use the following parameterizations:

ratio: $\phi_1 = \theta_1, \quad \phi_2 = \theta_2/\theta_1$;
logarithm: $\phi_1 = \log\theta_1, \quad \phi_2 = \log\theta_2$;
peak: $\phi_1 = (\log\theta_1 - \log\theta_2)/(\theta_1 - \theta_2), \ \phi_2 = \exp(-\theta_2\phi_1)$.

They also report that optimal n-point designs for $n = 3, \ldots, 7$ using (2) results in replicated two-point designs. For example, for $\hat{\theta} = (0{\cdot}7, 0{\cdot}2)'$, $\sigma = 0{\cdot}1$, and $n = 3$, the design (1·05, 5·83, 5·83) minimizes the area (2) with a value of $v' = 0{\cdot}1985$.

3. QUADRATIC DESIGNS WITH $p+1$ POINTS OF SUPPORT

For designs with $n > p$ points, Hamilton & Watts' criterion uses (2) to approximate (1) since $\hat{\varepsilon}$ is unknown. We now show that, whenever the number of design points is equal to $p+1$, we can still use equation (1) to obtain quadratic designs since we do know the direction and length of $\hat{\varepsilon}$.

For any n greater than p, the residual vector is always orthogonal to the tangent plane. Since the columns of the $n \times (n-p)$ matrix $\hat{N}$ form an orthonormal basis for the space orthogonal to the tangent plane, it follows that we can write $\hat{\varepsilon} = \hat{N}\alpha$, where α is some $(n-p) \times 1$ vector. For the particular case where $n = p+1$, $\hat{N}$ is an n-dimensional vector, and $\alpha = \sigma$ since the norm of $\hat{\varepsilon}$ is σ. Here $\hat{N}$ is easily obtained by premultiplying any vector not in the tangent plane by the orthogonal projection matrix, $I_n - \hat{U}(\hat{U}'\hat{U})^{-1}\hat{U}'$, and normalizing. It follows that $\hat{\varepsilon}$ is unique except for its sign, which turns out to be inconsequential since both positive and negative versions of $\hat{\varepsilon}$ yield the same design. Hence, equation (1) can be used to generate $(p+1)$-point quadratic designs for given initial parameter estimates.

To illustrate the above methodology, we again consider the intermediate product model function (3), with initial estimates $\hat{\theta} = (0{\cdot}7, 0{\cdot}2)'$ and $\sigma = 0{\cdot}1$. The 3-point design which minimizes the area (1) can be obtained by taking the vector $\hat{N}$ equal to the cross-product of the two columns of $\hat{U}$, and by using a minimization routine analogous to the one described by Hamilton & Watts (1985, p. 244). The resultant design is (1·02, 4·72, 6·81) with $v = 0{\cdot}1895$, which represents a 4·5% area reduction over the corresponding design obtained by Hamilton & Watts given above. When σ was increased to 0·15, the percentage area reduction increased to 8·2%. Further, in all examples investigated, changing the sign of $\hat{\varepsilon}$ resulted only in a rearrangement of the same design; e.g. (1·02, 6·81, 4·72) instead of (1·02, 4·72, 6·81). Thus, although the orientation of the expectation surface and the residual vector changes with this sign change, the design remains unchanged.

To investigate the changes in the 3-point designs obtained from minimizing (1) due to parameterization and noise level, designs were obtained using the original, ratio, logarithm, and peak forms of (3) given above, and for $\sigma = 0$, 0·025, 0·05, 0·075, 0·10. The results are listed in Table 1. When $\sigma = 0$, all the optimal designs are the same as the D-optimal design, (1·23, 6·86, 6·86), since in this case (1) reduces to $c|\hat{V}'\hat{V}|^{-\frac{1}{2}}$, the first-order volume approximation.

Increasing σ causes the optimal design to move away from this point in different directions, depending on the parameterization. Note that for values of σ above 0·20 in this example the volume approximation (1) breaks down since the Jacobian approximation becomes unstable.

Table 1. *Quadratic optimal designs for varying parameterizations and noise levels*

σ	Original	Ratio	Logarithm	Peak
0·00	(1·23, 6·86, 6·86)	(1·23, 6·86, 6·86)	(1·23, 6·86, 6·86)	(1·23, 6·86, 6·86)
0·025	(1·21, 6·41, 7·08)	(1·22, 6·37, 7·06)	(1·23, 6·51, 7·17)	(1·25, 6·55, 7·21)
0·05	(1·15, 5·79, 7·11)	(1·20, 5·60, 7·09)	(1·25, 6·11, 7·43)	(1·30, 6·28, 7·59)
0·075	(1·09, 5·17, 7·01)	(1·16, 4·58, 6·97)	(1·26, 5·67, 7·66)	(1·38, 6·02, 7·97)
0·10	(1·02, 4·72, 6·81)	(1·09, 3·59, 6·55)	(1·28, 5·12, 7·84)	(1·47, 5·69, 8·32)

4. Conclusion

Whereas D-optimal designs take account of neither the intrinsic nor the parameter-effects nonlinearity of the model, and quadratic designs based on (2) ignore intrinsic nonlinearity, $(p+1)$-point quadratic designs based on (1) take all of the curvature of the expectation surface into account. The technique used to obtain $(p+1)$-point quadratic designs is relatively simple; and the derived ability to check the adequacy of the assumed model is often paramount.

Acknowledgements

The author wishes to express his appreciation to G. Peter Clarke, Linda M. Haines and John O. Rawlings for helpful contributions.

References

Atkinson, A. C. (1988). Recent developments in the methods of optimum and related experimental designs. *Int. Statist. Rev.* **56**, 99-115.

Atkinson, A. C. & Hunter, W. G. (1968). The design of experiments for parameter estimation. *Technometrics* **10**, 271-89.

Box, G. E. P. & Lucas, H. L. (1959). Design of experiments in nonlinear situations. *Biometrika* **46**, 77-90.

Box, M. J. (1968). The occurrence of replications in optimal designs of experiments to estimate parameters in non-linear models. *J. R. Statist. Soc.* B **30**, 290-302.

Chaloner, K. & Larntz, K. (1989). Optimal Bayesian design applied to logistic regression experiments. *J. Statist. Plan. Inf.* **21**, 191-208.

Cochran, W. G. (1973). Experiments for nonlinear functions. *J. Am. Statist. Assoc.* **68**, 771-81.

Hamilton, D. C. & Watts, D. G. (1985). A quadratic design criterion for precise estimation in nonlinear regression models. *Technometrics* **27**, 241-50.

Vila, J. P. (1991). Local optimality of replications from a minimal D-optimal design in regression: A sufficient and quasi-necessary condition. *J. Statist. Plan. Inf.* **29**, 261-77.

[*Received February* 1992. *Revised June* 1992]