A New Robust Design Strategy for Sigmoidal Models Based on Model Nesting

Timothy E. O'Brien

Department of Statistics, Washington State University

1 Introduction

For a given process, researchers often have a specific nonlinear model in mind and perhaps a reasonable initial estimate of the p model parameters. In this situation, optimal design theory produces designs which typically have only p support points even when the final sample size (n) is chosen to exceed the number of parameters. Since p-point designs assume that the model function is known with complete certainty and provide no opportunity to test for the adequacy of the assumed model, they are clearly not "optimal" in most practical settings.

The focus of this paper is to provide an algorithm to obtain efficient designs with "extra" design points, or "robust" designs, by nesting a given model function (the "original" model function) in a larger one (the "super-model") which reduces to the original model for certain parameter choices. This design approach, called the nesting design strategy, has been applied to linear models in [1], [2], [7], [14], and [15], and only in very simple instances to nonlinear models in [3] and [6]. The application to nonlinear models is more difficult since the super-model is often less apparent (and more *ad hoc*) and requires a keen understanding of the nature of the various model functions; for this reason, the discussion here is limited only to sigmoidal growth models (e.g., those given in Chapter 4 of [11] and Chapter 7 of [13]).

2 Sigmoidal Growth Models

Most growth models in current use fall into one of three families: the Weibull, the Log-Logistic, and the Richards. Special cases of these models include the Logistic, Gompertz, Michaelis-Menton, Mischerlich, and Simple Exponential models, and often several model functions from these families can be used

In Dutter, R and Grossman, W., eds. (1994)
COMPSTAT Physica-Verlag: Heidelberg pp 97-102

to adequately describe a given set of data. For example, the two-parameter Weibull function (W2),

$$\eta_{w2} = \exp\{-[(x/\theta_1)^2]^{\theta_2}\}, \tag{1}$$

and the two-parameter Log-Logistic model function (LL2),

$$\eta_{ll2} = \frac{1}{1+(x/\theta_3)^{\theta_4}}, \tag{2}$$

behave quite similarly over $\Re_+$ for certain parameter choices. Since optimal designs for either of these model functions typically have only two support points (see [9]), functions which generalize the Weibull, Log-Logistic, and Richards families are required.

One important generalization of (1) is the three-parameter humped Weibull model function (HW3),

$$\eta_{hw3} = \exp\left\{-[(\frac{x-\theta_3}{\theta_1})^2]^{\theta_2}\right\}. \tag{3}$$

In some instances, the HW3 model function fits the data obtained in ozone dose response studies better that the W2 function (see [12]), although the biological interpretation of the corresponding "hump-effect" is not readily apparent. In addition, an important generalization of the Weibull, Log-Logistic and Richards model functions is the six-parameter Eclectic model (E6),

$$\eta_{e6} = \frac{\theta_1}{\left\{1+\frac{1}{\theta_6}\exp\left[\frac{\left(\frac{x-\theta_2}{\theta_3}\right)^{\theta_4\theta_5}-1}{\theta_5}\right]\right\}^{\theta_6}}, \tag{4}$$

studied in [10]. Conditions under which this function reduces to the Weibull, Log-Logistic and Richards family members are given in [10]; to illustrate one of these cases, note that the three-parameter Log-Logistic model function (LL3),

$$\eta_{ll3} = \frac{\theta_1}{1+(x/\theta_3)^{\theta_4}}, \tag{5}$$

is obtained from (4) by taking $\theta_2 = 0, \theta_5 \to 0$, and $\theta_6 = 1$.

3 Optimal Design Theory

The design problem for the homoskedastic Gaussian nonlinear model

$$y_i = \eta(\mathbf{x}_i, \phi) + \epsilon_i \qquad \epsilon_i \sim iid\ N(0,\sigma^2) \ \text{ for i} = 1,\dots,\text{n}$$

typically involves choosing an n-point design, ξ, to estimate some function of the p-dimensional parameter vector, ϕ, with high efficiency. This design associates the design weights $\omega_1, \omega_2, ..., \omega_n$ with the design points (or vectors) $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n$, respectively, and the corresponding (Fisher) information matrix is given by

$$\mathbf{M}(\xi, \phi) = \sum_{i=1}^{n} \omega_i \frac{\partial \eta(\mathbf{x}_i)}{\partial \phi} \frac{\partial \eta(\mathbf{x}_i)}{\partial \phi'} = \mathbf{V}'\Omega\mathbf{V},$$

where $\mathbf{V}$ is the n×p Jacobian of η and $\Omega = \mathrm{diag}\{\omega_1, \ldots, \omega_n\}$.

First-order optimal designs typically minimize some convex function of $\mathbf{M}^{-1}$. For example, designs which minimize the determinant $|\mathbf{M}^{-1}(\xi, \phi^{\circ})|$ are called locally D-optimal, where the term "locally" is used to emphasize the fact that an initial estimate of the parameter vector has been used. Other first order design criteria are discussed in [2] and [4], and second-order design criteria–or procedures which provide designs that attempt to reduce curvature in addition to efficiently estimating parameters–are presented in [8], [9], and [10]. Since optimal designs typically have only p support points regardless of the design criteria and final sample size used (see [9] and [16]), we seek a practical design algorithm which efficiently estimates the model parameters and also provides "extra" design points to check for model mis-specification.

4 A First-Order Nesting Design Strategy

Suppose that a researcher feels that the original model function $\eta_{or}(\phi_1)$ adequately describes a given process, but desires a design which, in addition to efficiently estimating the p_1 model parameters of η_{or}, also may be used to check for the adequacy of the assumed model. As a first step, we search for a relevant super-model, $\eta_{sm}(\phi_1, \phi_2)$, such that η_{sm} reduces to η_{or} when the p_2-vector ϕ_2 is equal to some (possibly extended) real vector. It is important that the super-model contains a reasonable generalization of the original model; thus, for example, by nesting the Log-Logistic model function in the Eclectic function, departures from the Log-Logistic model function in the direction of the Weibull and Richards functions may be detected.

A measure of the inefficiency that the design ξ has in estimating ϕ_1 in η_{or} is given by $|\mathbf{M_{11}}^{-1}|$, and a measure of it's inefficiency regarding detecting departures from η_{or} in the direction of η_{sm} is given by $\left|\left(\mathbf{M_{22}} - \mathbf{M_{21}}\, \mathbf{M_{11}^{-1}}\, \mathbf{M_{12}}\right)\right.$ where $\mathbf{M}_{ij} = \mathbf{V}_i\Omega\mathbf{V}_j$ (i,j = 1,2); see [2]. We combine these measures into the single (first-order) inefficiency measure,

$$\psi_1(\xi, \lambda, \phi) = \frac{\lambda}{p_1} |\mathbf{M_{11}}^{-1}| + \frac{1-\lambda}{p_2} \left|\left(\mathbf{M_{22}} - \mathbf{M_{21}}\, \mathbf{M_{11}^{-1}}\, \mathbf{M_{12}}\right)^{-1}\right|, \quad (6)$$

and seek designs to minimize ψ_1 for given choices of λ and $\phi = (\phi_1', \phi_2')'$, designs which are called locally D_λ-optimal here. Note that λ controls the amount of information obtained regarding estimation of ϕ_1 relative to detecting departures from the original model function, so that when we take $\lambda = 1$, we obtain information only regarding ϕ_1, whereas when we choose $\lambda = 0$, we obtain information only about departures from the original function. Our recommendation is to obtain the locally D_λ-optimal design, ξ_λ, such that it's D-efficiency relative to the locally D-optimal design, ξ_D, $\{\mathbf{M_{11}}^{-1}(\xi_D) / \mathbf{M_{11}}^{-1}(\xi_\lambda)\}$ is around 0.95.

5 Examples

5.1 Example 1

An environmental scientist believes that the W2 function in (1) with $\theta_1 = \theta_2 = 1$ adequately describes the dose response relationship of a given cultivar exposed to ozone (x), but wishes to allow for the hump-effect of the HW3 function in (3). Our procedure provides the locally D_λ-optimal design using the super-model (3) with $\theta_3 = 0$ and $\lambda = .95$, which associates the weights $\omega = 0.10$, 0.43 and 0.47 with the points x = 0.33, 0.63 and 1.30, respectively. This design is preferred to the locally D-optimal design for estimating the W2 model parameters, ξ_D, since this latter design has only two support points (at x = 0.59 and 1.28 each with equal weight), yet is "close" to ξ_D since it's D-efficiency is 95%.

5.2 Example 2

An animal scientist feels that the LL3 model function in (5) with $\theta_1 = \theta_3 = 1$, and $\theta_4 = 2$ reasonably describes the amount of food left in a cow's rumin x minutes after ingestion, but seeks a design with more than three support points so as to check for model-mis-specification. Using the E6 model function in (4) as the super-model with $\theta_2 = 0, \theta_5 = .001$, $\theta_6 = 1$ and $\lambda = .95$, the corresponding locally D_λ-optimal design, ξ_λ, places the weights $\omega = 0.18$, 0.13 0.05, 0.30, 0.14 and 0.20 at the points x = 0, 0.04, 0.23, 0.60, 1.35 and 1.97, and results in a D-efficiency of nearly 97%. This design is preferred to the locally D-optimal design for estimating the LL3 model parameters since this latter design has only three support points (at x = 0, 0.59 and 1.68 each with equal weight). It is important to note that since the E6 function also has the Weibull and Richards families as special cases, ξ_λ protects against departures from the LL3 function in the direction of practically all other sigmoidal curves.

6 A Second-Order Nesting Design Strategy

The first-order nesting strategy presented above is easily extended to provide efficient robust designs with reduced marginal curvature by using a penalty function approach. Based on Clarke's criterion for the seriousness of curvature (in [5]), we let $k_i = \max\{|mc_i| - 0.10, 0\}$, $l_i = \max\{|mc_i| - 0.30, 0\}$, and $\pi(\xi, \phi) = \sum_{i=1}^{p_1} \exp\{\alpha_i k_i + \beta_i l_i\}$; here mc_i is the marginal curvature associated with the i^{th} parameter in η_{or}, $\pi(\xi, \phi)$ is our curvature penalty function, and the α_i's and β_i's are chosen to emphasize certain components of the marginal curvature vector over others. Thus,

$$\psi_2(\xi, \gamma, \lambda, \phi) = \gamma \, \psi_1(\xi, \lambda, \phi) + (1 - \gamma) \, \pi(\xi, \phi),$$

for ψ_1 in (6), is a second-order inefficiency measure, and $\gamma \in [0, 1]$ controls the degree of emphasis placed on parameter estimation relative to curvature reduction. Designs which minimize ψ_2 for specific choices of γ, λ, and ϕ, called locally D_γ-optimal here, are usually preferred to locally D_λ-optimal designs when curvature is a concern. For example, the marginal curvatures associated with θ_3 using the D_λ-optimal design given in Example 2 is 0.3066, indicating serious curvature by Clarke's criterion. In contrast, the locally D_γ-optimal design obtained by minimizing ψ_2 (with $\gamma = .5$, $\lambda = .95$, each $\alpha_i = 2$, and each $\beta_i = 0$) is such that no component of the marginal curvature vector exceeds 0.10. Further, this latter design, which associates the weights ω = 0.08, 0.16 0.12, 0.23, 0.21 and 0.20 with the points x = 0, 0.05, 0.13, 0.75, 1.53 and 1.91, is also efficient in estimating the LL3 model parameters since it's D-efficiency is 90%.

7 Discussion

The design strategies presented here provide researchers with a reasonable compromise between so-called "optimal" designs, which typically cannot be used to check whether the assumed model is indeed valid, on the one hand, and designs comprised of arbitrarily chosen design points (for example, those with a geometric spacing of points), on the other. The first nesting design strategy is intended to be used with "close-to-linear" nonlinear models (see [11]); when designs with reduced curvature are desired, the second nesting design strategy should be used. Although our focus here has only been on sigmoidal growth curves, the application of the above procedures to other classes of nonlinear models is obvious provided relevant super-models can be found.

References

[1] Atkinson, A.C. (1972). Planning experiments to detect inadequate regression models. *Biometrika* **59**, 275-293.

[2] Atkinson, A.C. and Donev, A.N. (1992). *Optimum Experimental Designs.* Oxford: Clarendon Press.

[3] Box, G.E.P. and Lucas, H.L. (1959). Design of experiments in non-linear situations. *Biometrika* **46**, 77-90.

[4] Chaloner, K. and Larntz, K. (1989). Optimal Bayesian design applied to logistic regression experiments. *J. Stat Plann. Infer.* **21**, 191-208.

[5] Clarke, G.P.Y. (1987). Marginal curvatures and their usefulness in the analysis of nonlinear regression models. *J. Amer Stat. Assoc.* **82**, 844-850.

[6] Cochran, W.G. (1973). Experiments for nonlinear functions. *J. Amer. Stat. Assoc.* **68**, 771-781.

[7] DeFeo, P. and Myers, R.H. (1992). A new look at experimental design robustness. *Biometrika* **79**, 375-380.

[8] Hamilton, D.C. and Watts, D.G. (1985). A quadratic design criterion for precise estimation in nonlinear regression models. *Technometrics* **27**, 241-250.

[9] O'Brien, T.E. (1992). A note on quadratic designs for nonlinear regression models. *Biometrika* **79**, 847-849.

[10] O'Brien, T.E. (1993). Design strategies for nonlinear regression models. Ph.D. dissertation, North Carolina State University, Raleigh, N.C.

[11] Ratkowsky, D.A. (1983). *Nonlinear Regression Modeling.* New York: Marcel Dekker.

[12] Reinert, R.A., Rufty, R.C. and Eason, G. (1988). Interaction of tobacco etch or tobacco vein mottling virus and ozone on biomass changes in Burley tobacco. *Envir. Pollution.* **53**, 209-218.

[13] Seber, G.A.F. and Wild, C.J. (1989). *Nonlinear Regression.* New York: Wiley.

[14] Stigler, S.M. (1971). Optimal experimental design for polynomial regression. *J. Amer. Stat. Assoc.* **66**, 311-318.

[15] Studden, W. J. (1982). Some robust-type D-optimal designs in polynomial regression. *J. Amer. Stat. Assoc.* **77**, 916-921.

[16] Vila, J.P. (1991). Local optimality of replications from a minimal D-optimal design in regression: a sufficient and quasi-necessary condition. *J. Stat. Plann. Infer.* **29**, 261-277.