Optimal Designs for Implicit Models

Mariano Amo-Salas, Alfonso Jiménez-Alcázar, and Jesús López-Fidalgo

Abstract In this paper the tools provided by the theory of the optimal design of experiments are applied to a model where the function is given in implicit form. This work is motivated by a dosimetry problem, where the dose, the controllable variable, is expressed as a function of the observed value from the experiment. The best doses will be computed in order to obtain precise estimators of the parameters of the model. For that, the inverse function theorem will be used to obtain the Fisher information matrix. Properly the *D*-optimal design must be obtained directly on the dose using the inverse function theorem. Alternatively a fictitious *D*-optimal design on the observed values can be obtained in the usual way. Then this design can be transformed through the model into a design on the doses. Both designs will be computed and compared for a real example. Moreover, different optimal sequences and their *D*-effiencies will be computed as well. Finally, *c*-optimal designs for the parameters of the model will be provided.

1 Introduction

This paper is focused on the case of nonlinear models where the explanatory variable is expressed as a function of the dependent variable or response and this function is not invertible. That is, we consider the model

$$y = \eta(x, \theta) + \varepsilon, \quad \varepsilon \sim N(0, \sigma), \tag{1}$$

M. Amo-Salas
Faculty of Medicine, University of Castilla-La Mancha, Camino de Moledores s/n 13071-Ciudad Real, Spain
e-mail: Mariano.Amo@uclm.es

A. Jiménez-Alcázar
University of Castilla-La Mancha, Environmental Sciences Institute Avda. Carlos III s/n. 45071, Toledo, Spain
e-mail: Alfonso.JAlcazar@uclm.es

J. López-Fidalgo (✉)
Higher Technical School of Industrial Engineering, University of Castilla-La Mancha, Avda. Camilo José Cela 3, 13071-Ciudad Real, Spain
e-mail: Jesus.LopezFidalgo@uclm.es

J. Kunert et al. (eds.), *mODa 11 - Advances in Model-Oriented Design and Analysis*, Contributions to Statistics, DOI 10.1007/978-3-319-31266-8_2

where y is the dependent variable, x is the explanatory variable, θ is the vector of parameters of the model and $\mu(y, \theta) = \eta^{-1}(x, \theta)$ has a known expression, but a mathematical expression of $\eta(x, \theta)$ is not available. The challenge of this situation is to find optimal experimental designs for the explanatory variable when the expression of the function $\eta(x, \theta)$ is unknown. This situation is presented in a dosimetry study which will be used as case study in this work. Firstly, the description of the case study and a general introduction to the theory of Optimal Experimental Design is given. In Sect. 2 the inverse function theorem is applied to compute the information matrix. Finally, in Sect. 3 D-optimal designs are computed and compared for the case study proposed. Moreover, arithmetic and geometric optimal sequences, c-optimal designs and their D-effiencies are computed.

1.1 Case Study Background

The use of digital radiographs has been a turning point in dosimetry. In particular, radiochromic films are very popular nowadays because of their near tissue equivalence, weak energy dependence and high spatial resolution. In this area, calibration is frequently used to determine the right dose. The film is irradiated at known doses for building a calibration table, which will be used to fit a parametric model, where the dose plays the role of the dependent variable. The nature of this model is phenomenological since the darkness of the movie is only known qualitatively. An adjustment is necessary to filter noise and interpolate the unknown doses.

Ramos-García and Pérez-Azorín [9] used the following procedure. The radiochromic films were scanned twice. The first scanning was made when a pack of films arrived and the second 24 h after being irradiated. With the two recorded images the optical density, *netOD*, was calculated as the base 10 logarithm of the ratio between the means of the pixel values before (PV_0) and after (PV) the irradiation. They used patterns formed by 12 squares of $4 \times 4\,\text{cm}^2$ irradiated at different doses. This size is assumed enough to ensure the lateral electronic equilibrium for the beam under consideration. A resolution of 72 pp, without color correction and with 48-bit pixel depth was used for the measurements. The pixel values were read at the center of every square. Then, the mean and standard error were calculated. The authors assumed independent and normally distributed errors with constant variance as well as we do in this paper.

To adjust the results to the calibration table the following model was used:

$$netOD = \eta(D, \theta) + \varepsilon,$$

where D is the dose and the error ε will be assumed normally distributed with mean zero and constant variance, σ^2. The expression of the function $\eta(D, \theta)$ is unknown but the mathematical expression of the inverse is known

$$\eta^{-1}(D, \theta) = \mu(netOD, \theta) = \alpha\; netOD + \beta\; netOD^{\gamma},\; D \in [0, B], \tag{2}$$

where $\theta = (\alpha, \beta, \gamma)^T$ are unknown parameters to be estimated using maximum likelihood (MLE).

1.2 Optimal Experimental Design: General Background

Let a general nonlinear regression model be given by Equation (1). An *exact experimental design of size n* consists of a planned collection of points x_i, $i = 1, \ldots, n$, in a given compact *design space*, $\mathscr{X}$. Some of these points may be repeated and a probability measure can be defined assigning to each different point the proportion of times it appears in the design. This leads to the idea of extending the definition of experimental design to any probability measure (*approximate design*). It can be seen that, from the optimal experimental design viewpoint, we can restrict the search to finite designs of the type

$$\xi = \begin{Bmatrix} x_1 & x_2 & \ldots & x_k \\ p_1 & p_2 & \ldots & p_k \end{Bmatrix},$$

where $x_i,\ i = 1, \ldots, k$ are the support points and $\xi(x_i) = p_i$ is the proportion of experiments made at point x_i. Thus, $p_i \geq 0$ and $\sum_{i=1}^{k} p_i = 1$.

For the exponential family of distributions the Fisher Information Matrix (FIM) of a design ξ is given by

$$M(\xi, \theta) = \sum_{x \epsilon \chi} I(x, \theta)\xi(x), \tag{3}$$

where $I(x, \theta) = \frac{\partial \eta(x,\theta)}{\partial \theta} \frac{\partial \eta(x,\theta)}{\partial \theta^T}$ is the FIM at a particular point x. If the model is nonlinear, in the sense that function $\eta(x, \theta)$ is nonlinear in the parameters, the FIM depends on the parameters and nominal values for them have to be provided.

It can be proved that the inverse of this matrix is asymptotically proportional to the covariance matrix of the parameter estimators. An optimal design criterion aims to minimize the covariance matrix in some sense and therefore the inverse of the information matrix, $\Phi[M(\xi, \theta)]$. For simplicity $\Phi(\xi)$ will be used instead of $\Phi[M(\xi, \theta)]$. In this paper two popular criteria will be used, D-optimality and c-optimality. The D-optimality criterion minimizes the volume of the confidence ellipsoid of the parameters and is given by $\Phi_D(\xi) = \det M^{-1/m}(\xi, \theta)$, where m is the number of parameters in the model. The c-optimality criterion is used to estimate a linear combination of the parameters, say $c^T\theta$, and is defined by $\Phi_c(\xi) = c^T M^{-}(\xi, \theta)c$. The superscript "–" stands for the generalized inverse class of the matrix. Although the generalized inverse is unique only for nonsingular matrices the value of $c^T M(\xi, \theta)^{-} c$ is constant for any representative of the generalized inverse class if and only if $c^T\theta$ is estimable with the design. These criterion functions are convex and non-increasing. A design that minimizes one of these functions Φ over all the designs defined on $\mathscr{X}$ is called a Φ-optimal design, or more specifically, a D- or c-optimal design. It is worth to mention here that c-optimality raises important difficulties in nonlinear models when the optimal matrix is singular. In particular the actual covariance matrix may be different from the one predicted ([8], chap. 5).

The goodness of a design is measured by its efficiency, defined by

$$\mathrm{eff}_{\Phi}(\xi) = \frac{\Phi(\xi^*)}{\Phi(\xi)}.$$

In order to check whether a particular design is optimal or not there is a celebrated equivalence theorem [4] for approximate designs and convex criteria. This theorem consists in verifying that the directional derivative is non–negative in all directions. More details on the theory of optimal experimental designs may be found, e.g., at [3, 7] or [1].

2 Inverse Function Theorem for Computing the FIM

In the general theory, the experiments are designed for the explanatory variable, x, assumed under the control of the experimenter. In the case studied in this work, the function $\eta(x,\theta)$ is unknown but we know $\mu(y,\theta) = \eta^{-1}(x,\theta)$. Therefore the FIM should be defined in terms of y instead of x. The FIM is then given by (3), in particular, for one point the FIM is

$$I(x,\theta) = \frac{\partial\eta(x,\theta)}{\partial\theta}\frac{\partial\eta(x,\theta)}{\partial\theta^T}.$$

We can calculate the FIM in terms of the response variable y through the inverse function theorem. Differentiating the equation

$$x = \mu(y,\theta) = \mu[\eta(x,\theta),\theta],$$

we obtain

$$0 = \left(\frac{\partial\mu(y,\theta)}{\partial y}\right)_{y=\eta(x,\theta)}\frac{\partial\eta(x,\theta)}{\partial\theta} + \left(\frac{\partial\mu(y,\theta)}{\partial\theta}\right)_{y=\eta(x,\theta)}.$$

Then

$$\frac{\partial\eta(x,\theta)}{\partial\theta} = -\left(\frac{\partial\mu(y,\theta)}{\partial y}\right)^{-1}_{y=\eta(x,\theta)}\left(\frac{\partial\mu(y,\theta)}{\partial\theta}\right)_{y=\eta(x,\theta)}. \tag{4}$$

Using this result the FIM can be computed and therefore optimal designs on the explanatory variable may be obtained. This is the same model to be used for design when the variable y is heteroscedastic instead of homoscedastic with a sensitivity function (inverse of the variance),

$$\left| \left(\frac{\partial \mu(y,\theta)}{\partial y} \right)^{-1}_{y=\eta(x,\theta)} \right|.$$

This makes sense since assuming the response is a trend model plus some error with constant variance implies a trend model for x, which is the inverse of the original trend model plus an error with a non-constant variance coming from the transformation of the model.

3 Optimal Designs for the Case Study

In this section, the model proposed by the case study is considered. In this model, function $\eta(D,\theta)$ is unknown but $\eta^{-1}(D,\theta) = \mu(netOD,\theta)$ is known and defined by Equation (2). Computing the regressors vector with (4),

$$\frac{\partial \eta(D,\theta)}{\partial \theta} = \left[\frac{1}{\alpha_0 + \beta_0\gamma_0 netOD^{\gamma_0-1}} \begin{pmatrix} netOD \\ netOD^{\gamma_0} \\ \beta_0 netOD^{\gamma_0} \log(netOD) \end{pmatrix} \right]_{netOD=\eta(D,\theta)},$$

where $\alpha_0, \beta_0, \gamma_0$ are some nominal values assumed for the parameters to compute the optimal design. Thus, the FIM for a design ξ is

$$M(\xi;\theta_0) = \sum_i \xi(D_i) I(D_i;\theta_0),$$

where $\theta_0^T = (\alpha_0, \beta_0, \gamma_0)$ and

$$I(D;\theta_0) = \frac{\partial \eta(D,\theta)}{\partial \theta} \frac{\partial \eta(D,\theta)}{\partial \theta^T}. \quad (5)$$

Now, the function of the original model, $\eta(D,\theta)$, needs to be plugged into these formulas instead of *netOD*, but it cannot be inverted analytically. Using the results of [9], the design space will be $\mathscr{X}_D = [0,B] = [0,972]$ and the following nominal values for the parameters will be considered: $\alpha_0 = 690, \beta_0 = 1550, \gamma_0 = 2$. For these values the inverse function will be computed numerically when needed.

Assuming the D-optimal design is a three–point design, it should have equal weights at all of them. Since D-optimality is invariant for reparametrizations, the D-optimal design is computed for variable *netOD* using matrix (5). Computing the determinant of the information matrix for an equally weighted design, $\xi_{netOD} = \{netOD_1, netOD_2, netOD_3\}$, and minimizing $\Phi_D(\xi_{netOD})$ for values of $netOD_1$,

$netOD_2$ and $netOD_3$ in the interval $\mathscr{X}_{netOD} = [0, b] = [0, 0.6]$ the following design is obtained: $\xi_{netOD} = \{0.091, 0.348, 0.6\}$. The equivalence theorem states numerically that this design is actually D-optimal.

Transforming the three points through the equation model $\mu(netOD, \theta)$, with the previous nominal values of the parameters, $D = 690netOD + 1550netOD^2$, the optimal design on D is $\xi_D = \{75.6, 427.8, 972\}$.

Now a design for *netOD* will be computed in the usual way for the function $\mu(netOD, \theta)$. This is the optimal design for a wrong MLE from the explicit inverse model. Then this design will be compared with the right one checking the loss of efficiency. We will consider *netOD* as the explanatory variable and after computing the optimal design for *netOD*, we will invert it to compute the design for D. That is, we consider $\mu(netOD, \theta)$ as the function of the original model. Using the previous nominal values the design space is then $\mathscr{X}_{netOD} = [0, 0.6]$, and the D-optimal design is obtained in a similar way as above, $\xi^I_{netOD} = \{0.137, 0.409, 0.6\}$.

The equivalence theorem states numerically that this design is actually D-optimal. At this point a design for the response, D, can be obtained by transforming the design points again using the equation model $\mu(netOD, \theta)$, $\xi^I_D = \{123.4, 541.0, 972\}$.

Apparently the design is quite different, e.g. the first points, 75.6 and 123.4, are quite different. But the efficiency of this design with respect to the optimal one, ξ_D, is rather high,

$$\text{eff}_D(\xi^I_D) = \left(\frac{\Phi_D(\xi_D)}{\Phi_D(\xi^I_D)} \right)^{\frac{1}{3}} = 0.924.$$

Usually experimenters do not like extremal designs with a few different points (three in this case). Before obtaining the data there is always a reasonable doubt about the right model to be used and therefore more different points provide more safety (frequently just subjective). In practice, it is common to find designs distributing the points equidistantly (arithmetic sequence) or with geometric decreasing or increasing distances. This is usually made in a reasonable way taking into account the experience and intuition of the experimenter, but sometimes they can be far from optimal among the different possibilities. López Fidalgo and Wong [6] optimized different types of sequences according to D-optimality, including arithmetic, geometric, harmonic and an arithmetic inverse of the trend model. In the example considered in this paper we knew by personal communication that there was particular interest in arithmetic sequences. Optimal sequences of ten points are considered for both variables, *netOD* and D.

Table 1 shows the equally weighted optimal sequence designs, including the fixed equidistant designs with the corresponding D-efficiencies. Subscripts stand for the variable in which the sequence is computed (A = arithmetic, G = geometric and E = equidistant). The last point is always the upper extreme of the design space.

The geometric sequence is quite efficient while the equidistant sequence is by far the worst design.

Table 1 Suboptimal designs according to different patterns and efficiencies (last point, 975, is omitted)

	Design points									D-eff (%)
ξ_{netOD}^{A}	49.0	107.3	176.6	257.1	348.5	451.1	564.7	689.4	825.1	78.1
ξ_{D}^{A}	57.2	158.8	260.5	362.1	463.7	565.4	667.	768.7	870.3	75.5
ξ_{netOD}^{G}	55.6	73.3	97.2	130.2	176.3	241.4	334.9	471.	671.8	76.9
ξ_{D}^{G}	52.9	73.1	101.1	139.7	193.0	266.7	368.5	509.1	703.4	77.8
ξ_{netOD}^{E}	0.0	52.8	119.5	200.0	294.2	402.2	524	659.5	808.8	71.1
ξ_{D}^{E}	0	108	216	324	432	540	648	756	864	64.9

Table 2 c-efficiencies (%) of the D-optimal design for estimating each parameter

γ	56.7
α	42.4
β	65.2

Table 3 Efficiencies (%) of the optimal designs with respect to the nominal value of γ

γ	ξ_D	ξ_D^{α}	ξ_D^{β}	ξ_D^{γ}	ξ_{netOD}^{A}	ξ_D^{A}	ξ_{netOD}^{G}	ξ_D^{G}
1.8	99.5	95.5	99.6	98.2	98.6	99.2	98.9	98.9
2.2	99.4	95.6	99.5	98.0	99.7	99.3	98.8	98.9

In the example considered here, there is special interest in accurately estimating the parameter γ. Elfving's method [2] is a graphical procedure for calculating c-optimal designs. Although the method can be applied to any number of parameters it is not used directly for more than two parameters. López-Fidalgo and Rodríguez-Díaz [5] proposed a computational procedure for finding c-optimal designs using Elfving's method for more than two dimensions.

The c-optimal designs to estimate each of the parameters of the model are

$$\xi_D^{\alpha} = \left\{ \begin{matrix} 46.25 & 439.36 & 972 \\ 0.742 & 0.186 & 0.0717 \end{matrix} \right\}, \quad \xi_D^{\beta} = \left\{ \begin{matrix} 170.7 & 972 \\ 0.622 & 0.378 \end{matrix} \right\},$$

$$\xi_D^{\gamma} = \left\{ \begin{matrix} 46.25 & 439.36 & 972 \\ 0.476 & 0.359 & 0.165 \end{matrix} \right\}.$$

Table 2 shows the c-efficiencies of the D-optimal design for estimating each parameter. These efficiencies are low, specifically the c_γ-efficiency is lower than 60 %. The c−optimal design for β is a singular two–point one.

Table 3 displays the efficiencies of the different designs computed with respect the misspecification of parameter γ. This parameter can be considered as the most important parameter. Its nominal value is $\gamma_0 = 2$. The sensitivity analysis has been performed considering a deviation of ± 10 % from this value. The efficiencies from Table 3 show that the optimal designs are rather robust with respect to the misspecification of this parameter.

4 Concluding Remarks

This work deals with the problem of a model where the function is given in implicit form. In this case the FIM could not be computed in the usual way because the expression of the function of the model is unknown. Using the inverse function theorem, the FIM can be obtained and the D-optimal design may be computed. The D-optimal design was also determined directly on the dependent variable and then it was transformed into a design on the explanatory variable. This design displayed a moderate loss of efficiency when compared with the right one in this particular case.

Dependent errors or other distribution for them can be treated as well and it is one the future research lines.

Since three–point designs may be not acceptable from a practical point of view, ten different points were forced to be in the design restricting them to follow a regular sequence. In particular, arithmetic, geometric and inverse (through the trend model) sequences were considered. All of them were more efficient than the sequence used by the researchers. The geometric sequence achived the highest efficiency.

Finally, c-optimal designs for estimating the parameters of the model were computed. The c-efficiencies of the D-optimal design were lower than 70 % and specifically the c_γ-efficiency was lower than 60 %.

Acknowledgements The authors have been sponsored by Ministerio de Economía y Competitividad and fondos FEDER MTM2013-47879-C2-1-P. They want to thank the two referees for their interesting comments.

References

1. Atkinson, A.C., Donev, A.N., Tobias, R.D.: Optimum Experimental Designs with SAS. Oxford Statistical Science Series. Oxford University Press, Oxford (2007)
2. Elfving, G.: Optimum allocation in linear regression theory. Ann. Math. Stat. **23**, 255–262 (1952)
3. Fedorov, V.V., Hackl, P.: Model-Oriented Design of Experiments. Lecture Notes in Statistics. Springer, New York (1997)
4. Kiefer, J.: General equivalence theory for optimum designs (aproximate theory). Ann. Stat. **2**(5), 848–879 (1952)
5. López-Fidalgo, J., Rodríguez-Díaz, J.M.: Elfving's methods for m-dimensional models. Metrika **59**, 235–244 (2004)
6. López Fidalgo, J., Wong, W.K.: Design for the Michaelis-Menten model. J. Theor. Biol. **215**, 1–11 (2002)
7. Pazman, A.: Foundations of Optimum Experimental Design. D. Reidel publishing company, Dordrecht (1986)
8. Pronzato, L., Pazman, A.: Design of Experiments in Nonlinear Models. Asymptotic Normality, Optimality Criteria and Small-Sample Properties. Springer, New York (2013)
9. Ramos-García, L.I., Pérez-Azorín, J.F.: Improving the calibration of radiochromic films. Med. Phys. **40**(7), 071726 (2013)

Springer

http://www.springer.com/978-3-319-31264-4

mODa 11 - Advances in Model-Oriented Design and Analysis
Proceedings of the 11th International Workshop in Model-Oriented Design and Analysis held in Hamminkeln, Germany, June 12-17, 2016
Kunert, J.; Müller, C.H.; Atkinson, A.C. (Eds.)
2016, XIV, 254 p. 33 illus., 11 illus. in color., Softcover
ISBN: 978-3-319-31264-4