Adding points to D-optimal designs

Carlos de la Calle-Arroyo, Mariano Amo-Salas, Jesús López-Fidalgo, Licesio J. Rodríguez-Aragón

Universidad de Castilla-La Mancha
Universidad de Navarra

May 17, 2021

The model

A regression model can be expressed as

$$y(x) = \eta(x;\theta) + \varepsilon \quad \varepsilon \sim \mathcal{N}(\mu, \sigma^2(x)),$$

with θ the unknown parameters of the model, of interest for the experimenter.

Approximate Designs

A design is a collection of points to take a measure and the number of observations to be taken at those points.
A design seen as a probability measure over the design space is called an **approximate design**. It consists on the support points (x_i) and their corresponding proportions $(\xi(x_i))$:

$$\xi = \left\{ \begin{array}{ccc} x_1 & \dots & x_k \\ \xi(x_1) & \dots & \xi(x_k) \end{array} \right\}.$$

Approximate designs have properties that allow us easier calculations.

FIM

Under the normality assumption, the Fisher Information Matrix (FIM) of a design ξ is

$$M(\xi, \theta) = \sum_{x \in \mathcal{X}} f(x) f^t(x) \xi(x),$$

with $f(x) = \partial\eta(x, \theta)/\partial\theta$, the partial derivative vector of the model for each parameter, a first order Taylor expansion if the model is nonlinear.

D-optimality

While searching for optimal designs, a criterion to minimise $M^{-1}(\xi, \theta)$, proportional to the variance-covariance matrix of the estimates of the model, needs to be selected. In this work, this criterion is D-optimality, with the criterion expression:

$$\phi_D[M(\xi, \theta)] = |M(\xi, \theta)|^{-1/m}$$

and efficiency function

$$\mathrm{eff}_D(\xi_1, \xi_2) = \left(\frac{|M(\xi_1,\theta)|}{|M(\xi_2,\theta)|}\right)^{-1/m}$$

D-optimality Equivalence Theorem

The General Equivalence Theorem [3] shows that a design $\xi^\star$ is ϕ_D-optimal if and only if,

$$\psi(x, \xi^\star) = f^t(x) M^{-1}(\xi) f(x) - m \geq 0, \quad x \in \mathcal{X},$$

with the equality holding at the support points of $\xi^\star$.
The sensitivity function is

$$d(x, \xi) = f^t(x) M^{-1}(\xi) f(x)$$

Determinant recursive formula

The recursive formula for the determinant of the design $\xi_{n+1} = (1-\alpha)\xi_n + \alpha\xi_{x_{n+1}}$, adding a unipunctual design supported at x_{n+1} to a non-singular design ξ_n [2, p. 153], is given by

$$|M(\xi_{n+1})| = |M(\xi_n)|(1-\alpha)^m \left(1 + \frac{\alpha d(x_{n+1}, \xi_n)}{1-\alpha}\right),$$

which, with a bit of algebra gives an expression that relates the efficiency of ξ_{n+1} with the sensitivity function [1]

$$d(x, \xi_{n+1}) = \frac{1-\alpha}{\alpha}\left(\left(\frac{\text{eff}}{1-\alpha}\right)^m - 1\right).$$

Why modify the optimal design?

Optimal designs tend to require too few points, frequently very extreme. This often means:

- Impossibility to perform a lack-of-fit test.
- Impossibility to verify the adequacy of the experimental observations for the model.
- Preference for designs with a higher number of points.

The aim of this study is to improve the D-optimal design from the experimenter's point of view, building D-augmented designs. **D-augmented designs** are designs with more support points than the optimum but still guarantee a good enough efficiency.

Theorem

Let ξ be a non-singular design and $x_1, x_2, \ldots, x_s \in \mathcal{X}$ such that $d(x_i, \xi) \geq d(x_1, \xi)$ for all $i = 2, \ldots, s$. Then, let $\alpha_i \in (0, 1)$ for all $i = 1, \ldots, s$, and $\alpha = \sum_{i=1}^{s} \alpha_i < 1$ and $\xi_1 = (1 - \alpha)\xi + \alpha\xi_{x_1}$.

Then $\text{eff}(\xi_1, \xi) \leq \text{eff}(\xi_s, \xi)$ for $\xi_s = (1 - \alpha)\xi + \sum_{i=1}^{s} \alpha_i \xi_{x_i}$.

Application

The previous result allows us to:

1. Select an initial design.
2. Choose the combined **weight** of the new points to be added.
3. Choose a valid minimum **efficiency** for the resulting design.
4. Calculate the regions of candidate points.
5. Add points to the design in the desired proportion, with the guaranteed efficiency.

Shinny app

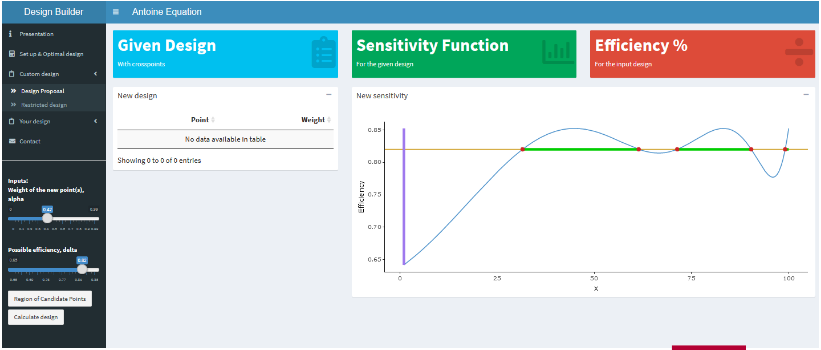
To illustrate this work, a Shiny App has been developed. It allows to calculate D-augmented design for a few models.
It can be visited at https://kezrael.shinyapps.io/AddPoints/ .

D-Optimal Design

First, the user selects the model, the space of the design and nominal values, if needed, and calculates the D-optimal design

Region of candidate points

Then, by selecting the weight and the efficiency, the user generates the region of candidates points



D-augmented design

Lastly, the user either adds points to create a D-augmented design or gets a pre-build D-augmented design

[1] S. Argumedo-Galván and V. López-Ríos. Metodología para incrementar el número de puntos experimentales en un diseño $D-$óptimo. *Ing. y Cien.*, 10:181–201, 2014.

[2] A. Atkinson, A. Donev, and R. Tobias. *Optimum Experimental Designs, With SAS*. Oxford Statistical Science Series. OUP Oxford, 2007.

[3] J. Kiefer and J. Wolfowitz. The equivalence of two extremum problems. *Canad. J. Math.*, 12:363–366, 1960.

Thanks for you attention!

For further information you can find me at

carlos.callearroyo@uclm.es