ARTICLE IN PRESS

Computational Statistics and Data Analysis ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Factorial and response surface designs robust to missing observations

Marcelo A. da Silva [a], Steven G. Gilmour [b], Luzia A. Trinca [c,*]

[a] *Programa Interinstitutional de Pós-Graduação em Estatística, USP/UFSCar, São Carlos, Brazil*

[b] *Department of Mathematics, King's College London, UK*

[c] *Departamento de Bioestatística, IB, Universidade Estadual Paulista, Botucatu, Brazil*

ARTICLE INFO

Article history:
Received 29 January 2016
Received in revised form 28 April 2016
Accepted 13 May 2016
Available online xxxx

Keywords:
Compound criteria
Cook's distance
Leverage
Optimum design
Pure error

ABSTRACT

Compound optimum design criteria which allow pure error degrees of freedom may produce designs that break down when even a single run is missing, if the number of experimental units is small. The inclusion, in the compound criteria, of a measure of leverage uniformity is proposed in order to produce designs that are more robust to missing observations. By appropriately choosing the weights of each part of the criterion, robust designs are obtained that are also highly efficient in terms of other properties. Applications to various experimental setups show the advantages of the new methods.

1. Introduction

Processes, products and methods in many areas are discovered and improved by performing controlled experiments in which the levels of several continuous inputs, experimental factors, are manipulated and at least one outcome is measured. Empirical models, such as low order polynomials, relating the response to the factor levels have been extremely useful for interpreting the data from such experiments. Such models and methods are part of the large area of Response Surface Methodology. Designs for experiments in this setup are known as Response Surface (RS) designs.

It has long been recognized that the experimental design should have several good properties. In the context of RS, Box and Draper (1975) started a list that was subsequently enlarged (Box and Draper, 1987, 2007) to 14 desired properties, some of them conflicting, indicating that in practice it is wise and necessary to compromise in order to choose a good design.

On the other hand, optimum design methodologies have concentrated on variance-based criteria such as *D*-, *A*- and *I*-optimality, the so called alphabetical optimality, see Atkinson et al. (2007) for an account of design criterion definitions. The use of a single optimality criterion may lead to designs that lack practical appeal. Gilmour and Trinca (2012) re-defined the alphabetical optimality criteria such that their properties are valid under inferences based on the randomization process only. They proposed adjustments to the traditional criteria allowing for pure error degrees of freedom in order to appropriately estimate random variation, the so called *DP* and *AP* criteria for instance. However, as recognized by the authors, these criteria may produce extreme designs with no spare degrees of freedom for inclusion of additional model terms. They further proposed compound criteria that aggregate into a single function the properties reflecting four experimental objectives, including a simple component, based on degree of freedom efficiency (Daniel, 1976) to drive the

* Corresponding author.
E-mail addresses: silva.marcelo@usp.br (M.A. da Silva), steven.gilmour@kcl.ac.uk (S.G. Gilmour), ltrinca@ibb.unesp.br (L.A. Trinca).

http://dx.doi.org/10.1016/j.csda.2016.05.023
0167-9473/

design to allow some lack of fit degrees of freedom as well. The use of compound criteria as well as procedures for multiple objectives (Lu et al., 2011) has the power to produce designs that are very statistically efficient and useful for experimenters.

Concerning the extreme designs produced by using a single property, e.g. *DP*, it was pointed out by Ridout (2012) that small designs would break down in case of even one missing observation from some treatment units. Robustness to missing observations is closely related to insensitivity to wild observations, a desired design property highlighted by Box and Draper. A design is said to be *robust* to missing observations if the model parameters are still estimable, without too much loss of precision, when observations from some experimental units are not available. Just as there are different design optimality criteria for estimation and for inference, there are different criteria for robustness. Surrogate measures related to the so called *leverages*, associated with a regression model, have been used to compare designs in this sense, as well as measures related to precision.

For example, Box and Draper (1975) studied the connections of the sum of squares of leverages and other design measures and found the best replication of center points and axial point values in central composite designs (CCD). Herzberg and Andrews (1976) and Andrews and Herzberg (1979) noted that such a measure does not discriminate well between designs and proposed extended measures incorporating some probability for the event of a missing observation. Akhtar and Prescott (1986) developed an efficiency measure relating the D criterion and the leverages and compared several CCDs, while Ahmad and Gilmour (2010) studied efficiency loss with respect to several optimality criteria due to missing data from different types of points in subset designs (Gilmour, 2006) and Ahmad et al. (2012) did the same for augmented pairs designs. Related investigations were also presented by Ghosh (1982a,b) who studied robustness of certain designs under sets of s missing runs and found the maximum s value for given designs. Adding to these works, Ghosh (1983, 1989) proposed measures to study influence on estimation and prediction of observations. To the best of our knowledge, a property related to robustness to missing data has not yet been incorporated in a criterion function in order to algorithmically construct an efficient RS design robust to missing data.

In this paper we incorporate a measure for the contribution of leverages, related to Cook's distance, in a compound design criterion in order to prevent the optimal design from being too sensitive to some observations or to breakdown in case of missing data. We show through several examples that such a property is particularly important in the case of limited experimental resources. In Section 2, a criterion for assessing design robustness is developed and in Section 3 a brief description of the algorithm is presented. The proposed criterion is shown to work well in several illustrative experiments in Section 4. Some final comments are made in Section 5.

2. Efficient and robust designs

Consider a completely randomized design in which there are t treatments, the distinct combinations of levels of q quantitative factors, to be allocated to n experimental units ($t < n$), treatment r being replicated n_r times ($n_r \in \mathbb{N}$, $\sum_{r=1}^{t} n_r = n$). The underlying model for the continuous random response variable Y is

$$y_{rj} = \mu_r + \varepsilon_{rj} \quad r = 1, 2, \ldots, t; j = 1, 2, \ldots, n_r, \tag{1}$$

where, in matrix notation, $E(\boldsymbol{\varepsilon}) = \mathbf{0}$ and $\mathbf{V}(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{I}$. Once data are collected the fitting of this model allows $d = n - t$ pure error degrees of freedom to estimate σ^2 unbiasedly. As argued in Box and Draper (2007), in RS experiments we want to simplify the model and add interpretability by approximating

$$\mu_r \approx \mathbf{f}(\mathbf{x}_r)'\boldsymbol{\beta} \quad r = 1, 2, \ldots, t, \tag{2}$$

where $\mathbf{x}_r$ is the vector of levels of the q factors defining treatment r (the design experimental points), $\mathbf{f}$ is the function that expands the levels according to the desired approximating function, usually a low order polynomial, and $\boldsymbol{\beta}$ is the p-dimensional vector of parameters with its first element being the intercept denoted by β_0. In matrix notation, let $\mathbf{X}_p = [\mathbf{1}|\mathbf{X}]$ be the $n \times p$ model matrix for Eq. (2) where each row of $\mathbf{X}_p$ corresponding to treatment r is $\mathbf{f}(\mathbf{x}_r)'$, $\mathbf{1}$ is the n dimensional column vector with all elements equal to 1 and $\mathbf{X}$ is a $n \times (p-1)$ matrix.

For the *DP* design criterion (Gilmour and Trinca, 2012) we should minimize $\left(F_{p,d;1-\alpha}\right)^p / |\mathbf{X}_p'\mathbf{X}_p|$, where $F_{p,d;1-\alpha}$ is the $1-\alpha$ quantile of the F distribution with p numerator and d denominator degrees of freedom and $1-\alpha$ is the confidence level of the confidence region for the p-parameter vector $\boldsymbol{\beta}$. Other alphabetical optimalities can be defined similarly. Using the $(DP)_S$ criterion, for the case of interest in a subset of p_2 ($p_2 < p$) parameters we should minimize $\left(F_{p_2,d;1-\alpha}\right)^{p_2} |(\mathbf{M}^{-1})_{22}|$, where $\mathbf{M} = \mathbf{X}_p'\mathbf{X}_p$ and $(\mathbf{M}^{-1})_{22}$ is the portion of $\mathbf{M}^{-1}$ referring to the subset of p_2 parameters of interest. See Atkinson et al. (2007) for details on D_S and other useful design criteria. If we use $p_2 = p - 1$ which drops only the intercept (β_0) from the set of parameters of interest, minimizing $|(\mathbf{M}^{-1})_{22}|$ is equivalent to maximizing $|\mathbf{X}'\mathbf{Q}\mathbf{X}|$, where $\mathbf{Q} = \mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}'$ and $\mathbf{I}$ is the $n \times n$ identity matrix. Focusing on four design objectives, each with a priority weight κ_l ($\sum_{l=1}^{4} \kappa_l = 1$), representing

1. global F test for treatment effects in $\boldsymbol{\beta}$, with significance level α_1;
2. partial confidence intervals for each regression parameter each with confidence level of $1 - \alpha_2$;
3. point estimation of each regression parameter; and
4. lack of fit degrees of freedom,