Journal of Statistical Planning and Inference ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

D-optimal asymmetric orthogonal array plus p run designs

E.M. Bird, Deborah J. Street *

University of Technology Sydney, NSW 2007, Australia

ARTICLE INFO

Article history:
Received 14 January 2015
Received in revised form 12 August 2015
Accepted 3 September 2015
Available online xxxx

MSC:
62K05

Keywords:
Fractional factorial design
Mixed-level array
Augmented runs
Adjoined runs
Augmented OA
Asymmetric array

ABSTRACT

In a recent paper, Chatzopoulos et al. (2011) identified some sufficient conditions to describe a set of p runs which, when adjoined to a symmetric orthogonal array, would result in a Type 1 optimal orthogonal array plus p run design. In this paper we find conditions for sets of p runs which, when adjoined to an asymmetric orthogonal array, result in a D-optimal orthogonal array plus p run design.

1. Introduction

Orthogonal arrays (OAs) are a class of fractional factorial designs, and are optimal according to a range of optimality criteria. This makes it tempting to construct fractional factorial designs by adjoining additional runs to an OA when the number of runs available for the experiment is only slightly larger than the number in the OA. Several researchers have investigated the performance of fractional factorial designs obtained by adjoining one, or more, runs to an orthogonal array. We summarise their work below.

When adding one additional run, the most general result was obtained by Mukerjee (1999) who gave a condition that must be satisfied to obtain an OA augmented with one additional run which is optimal, in the set of all fractional factorial designs with that number of runs, with respect to every generalised Type 1 criterion. Chai et al. (2002) have investigated the construction of orthogonal array plus one run designs when the conditions identified by Mukerjee (1999) are not satisfied. As a result, optimal resolution 3 orthogonal array plus one run designs with up to 100 runs in the original OA have been identified, except perhaps for some with 72 runs.

Other authors have investigated the form of optimal orthogonal array plus p run designs when $p > 1$. Hedayat and Zhu (2003) determined the best sets of runs to adjoin to saturated D-optimal resolution 3 binary orthogonal arrays for the estimation of the main effects only model. Tsai and Liao (2011) developed sufficient conditions for a partially replicated parallel-flats design, for a mix of binary and ternary factors, to be D-, A- or E-optimal. They considered the optimal regular designs for a specified set of model terms, which may be more extensive than a main effects only model. Chatzopoulos et al. (2011) identified some sufficient conditions to describe a set of p runs which, when adjoined to a symmetric orthogonal array, would result in a Type 1 optimal orthogonal array plus p run design for a main effects only model. Other results on

* Corresponding author.
E-mail address: Deborah.Street@uts.edu.au (D.J. Street).

http://dx.doi.org/10.1016/j.jspi.2015.09.001
0378-3758/

Please cite this article in press as: Bird, E.M., Street, D.J., D-optimal asymmetric orthogonal array plus p run designs. J. Statist. Plann. Inference (2015), http://dx.doi.org/10.1016/j.jspi.2015.09.001

binary factors are available in the literature. For example, Butler and Ramos (2007) (augmenting a resolution 5 design), Tsai et al. (2012) (deleting factors from a Hadamard matrix and adjoining runs to the resulting binary OA), and Tsai and Liao (2014) (using extended minimum aberration as the design comparison criterion).

Throughout this paper we focus on determining the D-optimal design, for the estimation of the main effects only model, amongst the set of orthogonal array plus p run designs. We provide an algorithm to find the best pairs of runs to adjoin to any OA, which we have applied to all OAs with $N \leq 100$. For larger values of p we characterise the theoretically best sets of runs to adjoin and give some examples of the best realisable sets of runs.

In Section 2 we give the notation that we will use in this paper. In Section 3 we give a bound for the determinant of an augmented OA. In Sections 4 and 5 we look at sets of p runs for $p = 2$ and $p = 3$ respectively, and in Section 6 we consider some general constructions. We close with a brief discussion.

2. Notation

We will consider an OA with m factors, where there are m_i factors with s_i levels, $1 \leq i \leq k$. Thus $m = \sum_{i=1}^{k} m_i$. We will represent the factor levels by the elements of Z_{s_i} and we will represent the treatment combinations by m-tuples with entries from the appropriate Z_{s_i}. If the OA has N runs and is of strength t we will denote this by writing OA$[N, s_1^{m_1} \times s_2^{m_2} \times \cdots \times s_k^{m_k}, t]$.

We will let $\mathbf{r}_i$, $1 \leq i \leq p$, be the adjoined runs and these may or may not be from the OA. (In general we are more likely to want to repeat runs of the OA so that we can get an estimate of pure error.)

We let I_s be the identity matrix of order s and $\mathbf{1}_s$ be the $s \times 1$ vector with all elements equal to 1. Since we want to estimate the main effects only model, we want to replace each level of each factor with appropriate entries from a set of orthogonal polynomials. To do this we follow the approach of Chatzopoulos et al. (2011) and define P_s to be a contrast matrix of order $(s-1) \times s$ that satisfies $P_s P_s' = sI_{s-1}$ and $P_s \mathbf{1}_s = 0$. Then for each run $\mathbf{t}$ of the complete factorial we associate a row vector with first entry 1 and in which each level of $\mathbf{t}$ is replaced by the transpose of the corresponding column of the matrix P_{s_i}. We denote this extended row vector by $\mathbf{q}$ and we observe that each extended row vector is of length $\alpha = 1 + \sum_i (s_i - 1)m_i$. We use Q for the $N \times \alpha$ matrix of the extended row vectors associated with the runs in the orthogonal array. So Q is the model matrix for the main effects only model. Thus $M = Q'Q$ is the information of the orthogonal array and hence we know that $M = NI_\alpha$.

3. A bound on the determinant of the information matrix of the augmented design

Suppose that we adjoin p runs to an OA$[N, s_1^{m_1} \times s_2^{m_2} \times \cdots \times s_k^{m_k}, t]$. Then we will write the model matrix of the augmented design as $Q_A = \begin{bmatrix} Q \\ A \end{bmatrix}$ with corresponding information matrix $M_A = Q'Q + A'A = NI_\alpha + A'A$.

A design is said to be *D-optimal* in a class of competing designs when its information matrix has the largest determinant from amongst the designs in the class. In this paper the class of competing designs is the set of orthogonal array plus p run designs.

To determine the D-optimal design, we first find a bound for $\det(M_A)$.

Following Hedayat and Zhu (2003), we can see that

$$\det(M_A) = \det(NI_\alpha + A'A) = \det(NI_\alpha)\det(I_p + A(NI_\alpha)^{-1}A') = N^{\alpha-p}\det(NI_p + AA').$$

Since $NI_p + AA'$ is positive definite and the diagonal entries of AA' are α, Hadamard's inequality shows that $\det(NI_p + AA')$ is bounded above by $(N+\alpha)^p$. Thus the upper bound on $\det(M_A)$ is $N^{\alpha-p}(N+\alpha)^p$.

The form of this determinant confirms that adjoining any single run is equally good according to the D-optimality criterion since $\mathbf{q} \cdot \mathbf{q}'$ takes the same value for any run from the complete factorial.

Suppose that we adjoin two runs, $\mathbf{r}_1$ and $\mathbf{r}_2$, with corresponding extended row vectors $\mathbf{q}_1$ and $\mathbf{q}_2$. We let d_i be the Hamming distance between $\mathbf{r}_1$ and $\mathbf{r}_2$ for the m_i factors with s_i levels, $1 \leq i \leq k$. Then

$$\begin{aligned} \mathbf{q}_1 \cdot \mathbf{q}_2' &= 1 + \sum_{i=1}^{k} [(s_i - 1)(m_i - d_i) + (-1)d_i] \\ &= 1 + \sum_{i=1}^{k} [(s_i - 1)m_i - s_i d_i]. \end{aligned} \tag{3.1}$$

To find the D-optimal OA plus 2 run design, we need to maximise the determinant of the 2×2 matrix $NI_2 + AA'$ which can be written as

$$\begin{bmatrix} N+\alpha & \mathbf{q}_1 \cdot \mathbf{q}_2' \\ \mathbf{q}_1 \cdot \mathbf{q}_2' & N+\alpha \end{bmatrix}.$$

Thus the D-optimal design is obtained by adding a pair of rows with the smallest realisable value of $|\mathbf{q}_1 \cdot \mathbf{q}_2'|$.

Consider larger values of p. The diagonal entries of $NI_p + AA'$ are all equal to $(N+\alpha)$, so the trace of this matrix is $(N+\alpha)p$. Since we want to maximise the determinant of this matrix, the arithmetic–geometric mean inequality says that we want