Journal of Complexity 31 (2015) 577–589

Contents lists available at ScienceDirect

Journal of Complexity

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

Algorithms for finding generalized minimum aberration designs

Dursun A. Bulutoglu [a,*], Kenneth J. Ryan [b]

[a] *Department of Mathematics and Statistics, Air Force Institute of Technology, Wright–Patterson Air Force Base, OH 45433, USA*

[b] *Department of Statistics, West Virginia University, Morgantown, WV 26506, USA*

ARTICLE INFO

Article history:
Received 17 September 2014
Received in revised form
22 November 2014
Accepted 9 December 2014
Available online 3 January 2015

Keywords:
Constraint programming
Directed enumeration constraints
Extension algorithm
Heuristic search
Isomorphism rejection
`Nauty`

ABSTRACT

Statistical design of experiments is widely used in scientific and industrial investigations. A generalized minimum aberration (GMA) orthogonal array is optimum under the well-established, so-called GMA criterion, and such an array can extract as much information as possible at a fixed cost. Finding GMA arrays is an open (yet fundamental) problem in design of experiments because constructing such arrays becomes intractable as the number of runs and factors increase. We develop two directed enumeration algorithms that call the integer programming with isomorphism pruning algorithm of Margot (2007) for the purpose of finding GMA arrays. Our results include 16 GMA arrays that were not previously in the literature, along with documentation of the efficiencies that made the required calculations possible within a reasonable budget of computer time. We also validate heuristic algorithms against a GMA array catalog, by showing that they quickly output near GMA arrays, and then use the heuristics to find near GMA arrays when enumeration is computationally burdensome.

Published by Elsevier Inc.

1. Introduction

This work tailors some state-of-the-art methods from operations research to find solutions in a fundamental class of problems in design of experiments. The main contribution of this paper is two directed enumeration algorithms that call the Margot [13] integer linear programming (ILP) solver.

* Corresponding author.
E-mail addresses: Dursun.Bulutoglu@afit.edu (D.A. Bulutoglu), kjryan@mail.wvu.edu (K.J. Ryan).

http://dx.doi.org/10.1016/j.jco.2014.12.001
0885-064X/Published by Elsevier Inc.

These algorithms were used to extend the known catalog of optimum orthogonal arrays (OAs) with 16 new entries. We also use heuristic search algorithms for finding optimum or near-optimum OAs when exact methods require too much computation. Algorithm performance (i.e., speed and ability to find optimum solutions) is documented.

1.1. Orthogonal arrays and the GMA criterion

A factorial design $\boldsymbol{Y}$ with N runs and k factors each having s-levels is an orthogonal array of strength t, $1 \le t \le k$, denoted by OA(N, k, s, t), if each of the s^t level combinations appears exactly N/s^t times when $\boldsymbol{Y}$ is projected onto any t factors. The index λ of an OA(N, k, s, t) is defined as N/s^t. An OA(N, k, s, t) is universally optimal for estimating the model containing all main effects and all interactions having $\lfloor t/2 \rfloor$ factors or less; see Cheng [7] and Mukerjee [16].

The design obtained by permuting factors or runs as well as levels in a subset of factors in an OA(N, k, s, t) is also an OA(N, k, s, t). Let such operations be called *isomorphism operations*. Two OA(N, k, s, t) are called *isomorphic* if one can be obtained from the other by applying a sequence of isomorphism operations. Assuming the hierarchical ordering principle (see Section 3.5 of Wu and Hamada [23]), two OA(N, k, s, t) are compared under model uncertainty using the *generalized minimum aberration* (GMA) criterion developed in Xu and Wu [26]. Let $\boldsymbol{Y} = [y_{ij}]$ be a 2-level design with entries ± 1 having N runs and k factors, and let $l = \{i_1, i_2, \ldots, i_r\} \subseteq \mathbb{Z}_k := \{1, \ldots, k\}$ be a nonempty subset of r factors. The GMA criterion is based on the concept of the J-characteristics

$$J_r(l) := \sum_{i=1}^{N} \prod_{j \in l} y_{ij}$$

of Tang and Deng [22]. Note $0 \le |J_r(l)| \le N$, and a larger $|J_r(l)|$ implies a stronger degree of aliasing among the factors in l. An average aliasing among all subsets of r factors is

$$A_r(\boldsymbol{Y}) := \frac{1}{N^2} \sum_{\{l \subseteq \mathbb{Z}_k : |l| = r\}} J_r(l)^2,$$

and GWP$(\boldsymbol{Y}) := (A_1(\boldsymbol{Y}), A_2(\boldsymbol{Y}), \ldots, A_k(\boldsymbol{Y}))$ is the *generalized word length pattern* (GWP) of $\boldsymbol{Y}$. The GMA criterion selects designs that sequentially minimize the GWP. A design with the same first non-zero GWP entry as a GMA design is a *weak* GMA design.

The general concept of GWP for s-level designs is computed as follows. Let $d_{ij}(\boldsymbol{Y})$ be the number of columns at which the ith and jth rows of $\boldsymbol{Y}$ differ, and define

$$B_r(\boldsymbol{Y}) := N^{-1} |\{(i, j) : d_{ij}(\boldsymbol{Y}) = r,\ i, j = 1, \ldots, N\}|$$

for $r = 0, \ldots, k$. The *distance distribution* $(B_0(\boldsymbol{Y}), B_1(\boldsymbol{Y}), \ldots, B_k(\boldsymbol{Y}))$ of $\boldsymbol{Y}$ determines the GWP and vice versa; the direct relationships provided in Xu and Wu [26] are:

$$A_j(\boldsymbol{Y}) = N^{-1} \sum_{i=0}^{k} P_j(i, s, k) B_i(\boldsymbol{Y})$$

$$B_j(\boldsymbol{Y}) = N s^{-k} \sum_{i=0}^{k} P_j(i, s, k) A_i(\boldsymbol{Y})$$

for $j = 0, \ldots, k$, where $A_0(\boldsymbol{Y}) = 1$ and $P_j(x, s, k) := \sum_{i=0}^{j} (-1)^i (s-1)^{j-i} \binom{x}{i} \binom{k-x}{j-i}$ are the *Krawtchouk polynomials*. When computing the Krawtchouk polynomials, the recursion

$$P_j(x, s, k) = P_j(x-1, s, k) - P_{j-1}(x-1, s, k) - (s-1) P_{j-1}(x, s, k)$$

with initial values $P_0(x, s, k) = 1$ and $P_j(0, s, k) = (s-1)^j \binom{k}{j}$ is useful.

1.2. Finding GMA designs

In general, finding GMA designs is a very difficult problem. Butler [5,6] theoretically constructed 2-level GMA designs. Butler's proofs for establishing that the constructed designs were GMA involved