Journal of the Korean Statistical Society 45 (2016) 210–220

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

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

Two-level screening designs derived from binary nonlinear codes

T.I. Katsaounis [a,*], Manohar L. Aggarwal [b]

[a] Department of Mathematics, The Ohio State University, Mansfield, OH 44906, USA

[b] Department of Mathematics, University of Memphis, Memphis, TN 38152, USA

ARTICLE INFO

Article history:
Received 24 October 2014
Accepted 7 October 2015
Available online 29 October 2015

AMS 2000 subject classifications:
primary 62K15
secondary 62K05

Keywords:
Nonregular factorial design
Screening experiment
Nonlinear code
Generalized aberration
Uniform design
Hamming code

ABSTRACT

Nonregular fractional factorial designs can provide economical designs in screening experiments. In this paper, two criteria are proposed for evaluating the projectivity and uniformity properties of projections onto active factors in two-level nonregular fractional factorial designs. Moreover, two-level nonregular fractional factorial designs derived from binary nonlinear codes with 12, 24, 32 and 40 codewords and various lengths are evaluated using the new criteria. Such designs are also evaluated under the known $E(s^2)$ criterion for optimal designs in screening experiments, and are compared to Plackett–Burman designs or to projections of Plackett–Burman designs. Results show that some binary nonlinear codes can provide useful two-level nonregular fractional factorial designs in screening experiments. A search method is proposed for finding good designs with a large number of factors, starting from a good design with the same number of runs but with a smaller number of factors.

1. Introduction

In some experiments there is a large number of potential factors that can influence the response variable. If only a small number of these factors are *active*, a screening experiment is performed in order to estimate the main effects of active factors and possibly few two-factor interactions that involve active factors. A factor is considered active if it is involved in at least one non negligible factorial effect. A common model for analyzing observations in a screening experiment is given by $\mathbf{Y} = \mathbf{X}\beta + \epsilon$, where $\mathbf{Y}$ is the vector of observations, $\mathbf{X}$ is the model matrix (which includes the mean, main effects of active factors and a few possibly two-factor interactions of active factors), β is the vector of parameters, and ϵ is the random error vector. The terms in the error vector are assumed to be distributed according to a normal distribution with mean equal to zero and (constant) variance equal to σ^2. See, for example, Dean and Lewis (2006) for a recent review of literature about designs and analysis in screening experiments.

An important issue in selecting a design in a screening experiment is estimability of factorial effects of interest. Two-level regular FFDs provide good designs in screening experiments in this regard. A regular FFD is a fraction of the total factor level combinations that can be obtained through a defining relation among factors. In a regular FFD any two factorial effects are either completely aliased or orthogonal to each other. Estimability of factorial effects is determined from the length of the shortest word (term) in its defining relation, called resolution (Box & Hunter, 1961). In a regular FFD with resolution $a = 2t$

* Corresponding author. Tel.: +1 419 755 4032; fax: +1 419 755 4367.
E-mail address: katsaounis.1@osu.edu (T.I. Katsaounis).

http://dx.doi.org/10.1016/j.jkss.2015.10.001

(or $a = 2t + 1$) all main effects and interactions involving at most t factors are estimable, under the assumption that all interactions involving more than t (or $t+1$) factors are negligible. Estimable factorial effects can be estimated independently of each other. As an example, the regular fraction of the full factorial 2^5 consisting of the eight runs (combinations of factor levels $-1, 1$) $(-1\ -1\ -1\ -1\ -1, 1\ -1\ -1\ 1\ 1, -1\ 1\ -1\ 1\ -1, -1\ -1\ 1\ -1\ 1, 1\ 1\ -1\ -1\ 1, 1\ -1\ 1\ 1\ -1, -1\ 1\ 1\ 1\ 1, 1\ 1\ 1\ -1\ -1)$, can be obtained through the defining relation $1 = (2345) = -(124) = -(135)$, usually written as $I = 2345 = -124 = -135$. That is, its eight runs (where a run is denoted by (12 345)) satisfy these equations (note that the factor levels can also be coded as 0, 1). The length of the shortest word in its defining relation is three, therefore it has resolution three. Thus, for example, the main effect of factor 1 is completely aliased with the two-factor interactions of factors 2,4 and 3,5, and the main effect of factor 2 is completely aliased with the three-factor interaction of factors 3,4,5. However, all main effects are estimable, under the assumption that all interactions involving more than two factors are negligible. A limitation of two-level regular FFDs is that they have run size (n) that equal a power of two.

On the other hand, two-level nonregular FFDs have flexible run size. A difficulty is that in a nonregular FFD two factorial effects can be completely or partially aliased. As a result, the aliasing pattern among factorial effects can be complex, and cannot be characterized through a defining relation (see, for example, Phoa, Xu, & Wong, 2009 for a recent study of nonregular FFDs). However, two-level nonregular FFDs can provide useful designs in screening experiments because they have good projections (subdesigns of small subsets of factors). Projections in a nonregular FFD can be characterized via the concept of *projectivity*. A FFD is said to have projectivity q if for every subset of q factors a complete factorial design possibly with some combinations replicated is produced (see Box & Tyssedal, 1996, and Cheng, 1995). As an example, the 12-run Plackett–Burman design with eleven factors, which is generated by cyclic permutation of the elements of the row vector (1 1 0 1 1 1 0 0 0 1 0) and adding a row of 0's (Plackett & Burman, 1946), has projectivity $q = 3$ (i.e. $t = 1$). It can be used to estimate all main effects of up to eleven factors, under that assumption that all interactions involving more than two factors are negligible. It can also be used to identify two active factors and estimate their main effects and interaction (see Dean & Lewis, 2006, Chapter 7). In general, in a FFD that has projectivity $q = 2t$ (or $q = 2t + 1$) all main effects and interactions that involve at most t factors are estimable, under the assumption that all interactions that involve more than t (or $t + 1$) factors are negligible. Notice that, in a two-level nonregular FFD in a screening experiment we want all main effects and interactions that involve at most t active factors to be estimable (assuming that all interactions that involve more than t (or $t + 1$) active factors are negligible), therefore only q-dimensional projections that involve active factors need to form a complete factorial design possibly with some runs replicated.

When it is not known in advance which factors are active, and thus which factorial effects actually have a significant impact on the response variable, we want a design that performs well even if the assumed model is not the true model. Therefore, another important issue in selecting a design in a screening experiment is robustness to model misspecifications. Choosing a design with points distributed uniformly on the experimental region, called uniform design, can address this issue. As Fang (1980) discussed, a uniform design can provide important information about the effects under study, even if the assumed model is not the true model, using a relatively small number of runs. Although uniformity of design points and estimability of factorial effects are different concepts, there are connections between them that have been discussed by various authors (see, for example, Fang, Lin, Winker, & Zhang, 2000, and Qin & Ai, 2007). In particular, Fang et al. (2000) showed that orthogonal FFDs can have uniform one- and two-dimensional projections. In an orthogonal FFD all main effects are estimable, under the assumption that interactions that involve more than two factors are negligible (see also Section 3.1). They concluded that, in general, uniform distribution of design points can be easily achieved in low dimensional projections in a FFD. Notice that, in our setting, we want to achieve a uniform distribution of design points in the q-dimensional projections that involve active factors.

Many of the existing criteria for ranking ($n \times p$) two-level nonregular FFDs, in terms of projectivity or uniformity, take into account all design projections onto dimensions $1, \ldots, p$. Thus, using such criteria to find a good FFD can be very inefficient. In our setting, a good two-level nonregular FFD allows estimation of factorial effects that involve active factors and has uniform projections onto active factors. In this paper, two ranking criteria for selecting good two-level nonregular FFDS are proposed, which take into account only relevant design projections. Moreover, the performance of binary nonlinear codes as two-level nonregular FFDs in screening experiments is investigated using the new criteria. Our motive for this investigation is the relationship between codes and screening designs. For example, the regular fraction discussed above is the shortened Hamming code of length 5 (listed in http://neilsloane.com/oadir/oa.8.5.2.2.txt), and the 12-run Plackett–Burman is (equivalent to) the Hadamard code with 12 codewords and length 11 (given, for example, in McWilliams & Sloane, 1977, p.39); Hadamard codes relate to a famous class of designs in screening experiments called Hadamard designs (see, for example, Assmus Jr. & Key, 1992). Note also that, the regular fraction discussed above can be obtained as the projection of the 8×8 Hadamard matrix, listed in http://neilsloane.com/hadamard/had.8.txt (with the first column of 1's omitted), onto the first five columns. In addition, the Plackett–Burman designs are a special case of Hadamard designs (see, for example, Dean & Lewis, 2006, Chapter 7). However, there is very little known about whether general nonlinear codes can be used as screening designs.

The rest of this paper is organized as follows. In Section 2, some fundamental connections between (non)linear codes and (non)regular FFDs are given. In Section 3.1, a criterion that measures the average projectivity of q-dimensional projections onto active factors in a two-level nonregular FFD is given. In Section 3.2, a criterion that measures the average uniformity of q-dimensional projections onto active factors in a two-level nonregular FFD is given. In Section 4, some existing binary nonlinear codes with 12, 24, 32, and 40 codewords and various lengths are ranked as two-level nonregular FFDs using