Journal of Complexity 30 (2014) 98–110

Contents lists available at ScienceDirect

Journal of Complexity

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

Construction of uniform designs without replications

Bochuan Jiang [a,b], Mingyao Ai [a,*]

[a] LMAM, School of Mathematical Sciences and Center for Statistical Science, Peking University, Beijing 100871, China

[b] Science and Technology on Complex Land Systems Simulation Laboratory, Beijing 100012, China

ARTICLE INFO

Article history:
Received 3 September 2012
Accepted 10 October 2013
Available online 23 October 2013

Keywords:
Complementary design
Discrepancy
Rank vector
Threshold-accepting algorithm
Uniformity

ABSTRACT

A uniform design scatters its design points evenly on the experimental domain according to some discrepancy measure. In this paper all the design points of a full factorial design can be split into two subdesigns. One is called the complementary design of the other. The complementary design theories of characterizing one design through the other under the four commonly used discrepancy measures are investigated. Based on these complementary design theories, some general rules for searching uniform designs through their complementary designs are proposed. An efficient method to check if a design has repeated points is introduced and a modified threshold-accepting algorithm is proposed to search uniform or nearly uniform designs without replications. The new algorithm is shown to be more efficient by comparing with other existing methods. Many new uniform or nearly uniform designs without replications are tabulated and compared.

1. Introduction

Uniform design has been widely used especially for computer experiments since it was proposed by Fang [1]. Its main idea is to scatter its design points evenly on the experimental domain according to some discrepancy measure. The commonly used measures of non-uniformity include the centered L_2-discrepancy (CD), the wrap-around L_2-discrepancy (WD) and the symmetric L_2-discrepancy (SD) introduced by Hickernell [9,10] and the discrete discrepancy (DD) proposed by Hickernell and Liu [11]. For a comprehensive discussion about the relationships among them refer to Fang, Li and Sudjianto [2].

* Corresponding author.
E-mail address: myai@math.pku.edu.cn (M. Ai).

0885-064X/$ – see front matter
http://dx.doi.org/10.1016/j.jco.2013.10.005

A design D in a specified design space $\mathcal{D}$ is said to be a uniform design under some discrepancy if it minimizes the discrepancy among all designs in $\mathcal{D}$.

Many construction methods of uniform or nearly uniform designs have been proposed. They are broadly classified into two categories: combinatorial algebra methods and algorithm optimizations. Based on the techniques of combinatorial algebras and number theories, good lattice method (Fang and Wang [8]), Latin square method (Fang, Shiu and Pan [6]) and balanced incomplete block design method (Lu and Meng [13]) were introduced. Along the line of algorithm optimizations, the threshold accepting heuristic (Winker and Fang [15]), simulated annealing (Morris and Mitchell [14]), stochastic evolutionary (Jin, Chen and Sudjianto [12]) and balance-pursuit heuristic (Fang, Tang and Yin [7]) were adopted. For detailed reviews and discussion of all kinds of constructions of uniform or nearly uniform designs refer to Fang, Li and Sudjianto [2] and the related references therein.

In this paper, a full factorial design is split into two subdesigns according to the design points. One is called the complementary design of the other. The complementary design theories of characterizing one design through the other under different discrepancy measures are investigated, which are very powerful for searching uniform or nearly uniform designs when their complementary designs are small. Furthermore, an efficient algorithm to check if a design has repeated points is proposed, and correspondingly the threshold-accepting (TA) algorithm is modified to search more uniform designs without replications.

The remainder of this paper is organized as follows. Section 2 presents the quadratic forms of the four discrepancy measures CD, WD, SD and DD. Section 3 establishes some relationships between the discrepancy measures of one design and its complementary design. Based on these complementary design theories, some general rules for searching uniform designs through their complementary designs are proposed in Section 4. In Section 5, a modified TA algorithm is provided to construct uniform designs without replications. The performance of the modified TA algorithm is displayed in Section 6 by comparing with other two existing methods. Many new uniform or nearly uniform designs without replications are tabulated and compared in Section 7. Section 8 concludes this article with some remarks.

2. Quadratic forms of discrepancies

Some notation and definitions are introduced here. Let $A \otimes B$ denote the Kronecker product of two matrices A and B. For any positive integer q, let $V_q = \{0, 1, \ldots, q-1\}$. Denote $V^m = V_{q_1} \times \cdots \times V_{q_m}$ and $N = q_1 \cdots q_m$. Let $\lfloor x \rfloor$ be the maximum integer not exceeding x. For two different vectors $\mathbf{u} = (u_1, \ldots, u_m)$ and $\mathbf{v} = (v_1, \ldots, v_m)$ in V^m, if $u_i = v_i$ for $i = 1, \ldots, k-1$ and $u_k < v_k$ for some k, then $\mathbf{u}$ is ordered before $\mathbf{v}$. This ordering rule is called the lexicographical order.

A mixed-level (or asymmetrical) design of n runs and m factors with levels $q_1, \ldots, q_m$, denoted by $(n, q_1 \cdots q_m)$, is a set of n row vectors (or points) in V^m or an $n \times m$ matrix in which each row represents a run, each column represents a factor and the jth column takes values from a set of q_j symbols, say, V_{q_j}. In particular, an (n, q^m)-design is symmetrical. A $(n, q_1 \cdots q_m)$-design is called balanced or U-type design if all levels of each factor appear equally often and denoted by $D(n, q_1 \cdots q_m)$. In this paper, we only consider the balanced designs which are usually needed in practice. The set of all such balanced designs is denoted by $\mathcal{D}(n, q_1 \cdots q_m)$.

For a design $D = (d_{ij}) \in \mathcal{D}(n, q_1 \cdots q_m)$, the detailed computational formulas of $\mathrm{CD}^2(D)$, $\mathrm{WD}^2(D)$, $\mathrm{SD}^2(D)$ and $\mathrm{DD}^2(D)$ were derived by Hickernell [9,10] and Fang, Lin, and Liu [3], respectively. Note that all the four formulas can be expressed in the following unified form

$$\text{constant} + n^{-2} \sum_{i=1}^{N} \sum_{j=1}^{N} \prod_{k=1}^{m} f(d_{ik}, d_{jk}, q_k) - 2n^{-1} \sum_{i=1}^{N} \prod_{k=1}^{m} g(d_{ik}, q_k),$$

where $f(\cdot, \cdot, \cdot)$ and $g(\cdot, \cdot)$ are different types of functions according to different discrepancies.

For a $D(n, q_1 \cdots q_m)$ design D, let $n(i_1, \ldots, i_m)$ denote the number of times that the point $(i_1, \ldots, i_m)$ occurs in D. Then the design D can be uniquely determined by the column vector of length N given by

$$\mathbf{y}_D = (n(i_1, \ldots, i_m))_{(i_1, \ldots, i_m) \in V^m}, \tag{1}$$