四水平计算机试验设计的构造

覃红¹, 欧祖军^{2*}, CHATTERJEE Kashinath³

1. 华中师范大学数学与统计学学院, 武汉 430079;

2. 吉首大学数学与统计学院, 吉首 416000;

3. Department of Statistics, Visva-Bharati University, Santiniketan 731235, India

E-mail: qinhong@mail.ccnu.edu.cn, ozj9325@mail.ccnu.edu.cn, kashinathchatterjee@gmail.com

收稿日期: 2016-12-21; 接受日期: 2017-03-15; 网络出版日期: 2017-05-03; * 通信作者

国家自然科学基金 (批准号: 11271147 和 11561025) 和湖南省自然科学基金 (批准号: 2017JJ2218) 资助项目

摘要 编码理论在最优设计的构造中已被广泛应用. 本文通过修正 Gray 映射编码变换来构造用于计算机试验的四水平均匀设计, 其效率通过离散偏差进行刻画; 同时也给出了离散偏差的下界, 这个下界可作为获得最优设计的一个基准.

关键词 效率 下界 修正 Gray 映射 均匀设计

MSC (2010) 主题分类 62K15, 62K05, 62K99

1 引言

在高新技术和工程领域中, 基于计算机技术的模拟已被广泛应用. 最近的研究文献表明计算机试验的设计和建模受到高度的重视 (参见文献 [1–3]). 均匀设计是一种重要的计算机试验设计, 它要求试验点均匀分布在实验域里. 如果在模拟研究中对建模的函数了解甚少, 那么可以使用均匀设计来获取数据并建模, 因为均匀设计具有稳健性. 同时由于其经济且灵活的试验次数可以同时来研究多个高水平因子, 因而均匀设计在实际试验中得到了成功应用.

均匀性测度在均匀设计的构造中具有重要作用. 一个 p ($p \geq 2$) 水平 U-型设计 U 是一个 $N \times m$ 的阵列, 其元素来自集合 $\{1/2p, 3/2p, \dots, (2p-1)/2p\}$, 使得该集合中的每个元素在阵列中的每一列出现相同的次数. 所有的 p 水平 U-型设计 U 的集合记为 U-型设计类 $\mathcal{U}(N; p^m)$. U 的每一行 (或试验) 可以作为 $\Omega^m = [0, 1]^m$ 上的一个点 $y = (y_1, y_2, \dots, y_m)$. 若 U-型设计 $U \in \mathcal{U}(N; p^m)$ 在 $\mathcal{U}(N; p^m)$ 上具有最小的均匀性测度值, 则称 U 在此均匀性测度下是最优 (或均匀) 的.

衡量设计的均匀性有很多均匀性测度, 其中离散偏差是人们常采用的一种均匀性测度. 因为离散偏差对行变换、列变换和坐标平移具有不变性^[3, 4], 本文考虑由文献 [4] 中的离散偏差来刻画四水平设计的效率测度.

文献 [5] 首先开创了四元编码设计的研究并建立了相关的理论与计算方面的结果; 文献 [6] 建立了 $1/4$ 部分四元编码设计的解析结果, 随后该结果被推广到更高的部分设置^[7], 并提出了一些关于设

英文引用格式: Qin H, Ou Z J, Chatterjee K. Construction of four-level designs for computer experiments (in Chinese). Sci Sin Math, 2017, 47: 1089–1100, doi: 10.1360/SCM-2016-0788

计结构和性质的基本定理^[8]. 由于其构造方法简单且便于应用, 四元编码设计在构造非正规设计时非常方便, 更为重要的是, 许多通过四元编码构造的设计具有非常好的统计性质^[9].

本文的主要目的是, 利用修正 Gray 映射编码变换来构造用于计算机试验的四水平均匀设计, 其设计效率通过离散偏差来度量. 本文余下部分结构如下: 第 2 节主要介绍一些基本的记号和基础知识; 第 3 节讨论基于修正 Gray 映射编码变换规则的四水平设计的构造; 第 4 节研究所构造的设计在离散偏差意义下的均匀性; 第 5 节给出一些数值例子; 第 6 节对全文进行总结.

2 基本记号和基础知识

考虑一个含 m 个四水平因子的试验. 用 $x = (x_1, x_2, \dots, x_m)$ 表示这 m 个因子的一个水平组合, 其中 $x_j = 0, 1, 2, 3, 1 \leq j \leq m$. 所有 $4^m (= v)$ 个水平组合 (或试验) 构成的集合记为 V , 其元素按字典顺序排列. 另用 $\mathcal{D}(N; 4^m)$ 表示一类具有 N 个来自集合 V 中的水平组合的设计, 使得对任意的设计 $d \in \mathcal{D}(N; 4^m)$, 每个因子的水平在 d 中出现相同的次数. 不难发现, 在映射 $f: (x_1, x_2, \dots, x_m) \rightarrow (y_1, y_2, \dots, y_m)$ 下, $\mathcal{D}(N; 4^m)$ 与 $\mathcal{U}(N; 4^m)$ 中的元素是一一对应的, 其中 $y_j = (2x_j + 1)/8, x_j = 0, 1, 2, 3$.

为方便起见, 记 $\mathcal{U}^*(N; 2^m)$ 为两水平 U- 型设计类, $\mathcal{D}^*(N; 2^m)$ 为每个因子各水平出现相同次数的两水平设计类. 记 V^* 为设计类 $\mathcal{D}^*(N; 2^m)$ 中的所有 $2^m (= v^*)$ 可能水平组合. 为了区分前面所定义的设计类, 用 d (或 d^*) 表示 $\mathcal{D}(N; 4^m)$ (或 $\mathcal{D}^*(N; 2^m)$) 中的设计. 对任意的 $x \in V$ 和 $d \in \mathcal{D}(N; 4^m)$, 记 $n_d(x)$ 为水平组合 x 在 d 中出现的次数, n_d 为 $n_d(x)$ 按字典顺序排列的 $v \times 1$ 向量. 类似地, 对任意的 $x \in V^*$ 和 $d^* \in \mathcal{D}^*(N; 2^m)$, $n_{d^*}(x)$ 为水平组合 x 在 d^* 中出现的次数, n_{d^*} 为 $n_{d^*}(x)$ 按字典顺序排列的 $v^* \times 1$ 向量. 此外, 记 c_{ij} (或 c_{ij}^*) 为 d (或 d^*) 的第 i 行与 j 列的水平重合数, 因此,

$$c_{ii} = c_{ii}^* = m, \quad 1 \leq i \leq m.$$

利用 Hilbert 空间的可再生核, 文献 [4] 介绍了一种基于离散域上的离散偏差; 文献 [10] 建立了离散偏差与广义最小混杂及最小矩混杂间的联系. 为后文的叙述需要, 下面将对离散偏差及其下界作基本的介绍.

对任意的设计 $d \in \mathcal{D}(N; 4^m)$, 其离散偏差值 $\text{DD}(d, a, b)$ 可由以下公式获得:

$$[\text{DD}(d, a, b)]^2 = -\left(\frac{a+3b}{4}\right)^m + \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \prod_{l=1}^m K(y_{il}, y_{jl}), \quad (2.1)$$

其中, $y_{il}, y_{jl} \in \{1/8, 3/8, 5/8, 7/8\}, a > b > 0$,

$$K(y_{il}, y_{jl}) = \begin{cases} a, & \text{若 } y_{il} = y_{jl}, \\ b, & \text{若 } y_{il} \neq y_{jl}. \end{cases}$$

下面的定义将有助于我们在 U- 型设计类中基于离散偏差如何选择最优设计.

定义 1 设计 $\tilde{d} \in \mathcal{D}(N; 4^m)$ 在离散偏差下是最优的当且仅当对任意的设计 $d \in \mathcal{D}(N; 4^m)$ 有

$$[\text{DD}(\tilde{d}, a, b)]^2 \leq [\text{DD}(d, a, b)]^2.$$

下面将对任意的设计 $d \in \mathcal{D}(N; 4^m)$ 给出其离散偏差的一个下界. 令

$$L_1 = \begin{pmatrix} a & b & b & b \\ b & a & b & b \\ b & b & a & b \\ b & b & b & a \end{pmatrix}, \quad M_1 = \bigotimes_{j=1}^m L_1, \quad \Gamma_1(0) = 1_4^T, \quad \Gamma_1(1) = I_4,$$

其中 1_4 是一个全为 1 的四元向量, I_4 为四阶的单位阵.

注意到

$$L_1 = \beta(\Gamma_1(0))^T \Gamma_1(0) + \alpha(\Gamma_1(1))^T \Gamma_1(1),$$

其中 $\alpha = (a - b)$, $\beta = b$.

令 $\Omega_1 = \{u = u_1 \cdots u_m \mid u_j = 0, 1; 1 \leq j \leq m\}$. 对 $0 \leq r \leq m$, 令 $\Omega_{1r} = \{u \in \Omega_1 \mid \sum u_j = r\}$. 对任意的 $u \in \Omega_1$, 定义矩阵

$$\Delta_1(u) = \bigotimes_{j=1}^m \Gamma_1(u_j).$$

基于上述记号, 矩阵 M_1 可表示为

$$M_1 = \beta^m \sum_{r=0}^m \left(\frac{\alpha}{\beta}\right)^r \sum_{u \in \Omega_{1r}} (\Delta_1(u))^T \Delta_1(u). \quad (2.2)$$

对任意的设计 $d \in \mathcal{D}(N; 4^m)$, 由 (2.2) 可知, (2.1) 可改写为

$$\begin{aligned} [\text{DD}(d, a, b)]^2 &= -\left(\frac{a+3b}{4}\right)^m + \frac{1}{N^2} n_d^T M_1 n_d \\ &= -\left(\frac{a+3b}{4}\right)^m + \frac{\beta^m}{N^2} \sum_{r=0}^m \left(\frac{\alpha}{\beta}\right)^r \sum_{u \in \Omega_{1r}} n_d^T (\Delta_1(u))^T \Delta_1(u) n_d. \end{aligned} \quad (2.3)$$

另一方面, 基于设计 d 的 $c_{i_1 i_2}$, (2.1) 也可改写为

$$[\text{DD}(d, a, b)]^2 = -\left(\frac{a+3b}{4}\right)^m + \frac{a^m}{N} + \frac{b^m}{N^2} \sum_{i_1=1}^N \sum_{i_2(\neq i_1)=1}^N \left(\frac{a}{b}\right)^{c_{i_1 i_2}}. \quad (2.4)$$

为获得 (2.3) 和 (2.4) 中 $[\text{DD}(d, a, b)]^2$ 的下界, 我们需要利用如下的引理.

引理 1 ^[11] 设 z_i ($i = 1, \dots, q$) 为非负整数, 且 $\sum_{i=1}^q z_i = c$, 则有

$$\sum_{i=1}^q z_i^g \geq \alpha_1 w^g + \alpha_2 (w+1)^g,$$

其中 w 为不超过 c/q 的最大整数, α_1 和 α_2 是使得 $\alpha_1 + \alpha_2 = q$ 和 $\alpha_1 w + \alpha_2 (w+1) = c$ 成立的整数.

引理 2 ^[11] 设 z_i ($i = 1, \dots, q$) 为非负整数, 且 $\sum_{i=1}^q z_i = c$, 则有

$$\sum_{i=1}^q \beta^{z_i} \geq \beta^w (\alpha_1 + \beta \alpha_2),$$

其中 w 、 α_1 和 α_2 如引理 1 所示.

引理 3 若设计 $d \in \mathcal{D}(N; 4^m)$, 则由 (2.3) 可得如下不等式:

$$[\text{DD}(d, a, b)]^2 \geq \text{LB}_1[\text{DD}(d, a, b)],$$

其中

$$\text{LB}_1[\text{DD}(d, a, b)] = -\left(\frac{a+3b}{4}\right)^m + \frac{\beta^m}{N^2} \sum_{r=0}^m \left(\frac{\alpha}{\beta}\right)^r \binom{m}{r} [a_{1r} w_{1r}^2 + b_{1r} (w_{1r} + 1)^2],$$

w_{1r} 是不超过 $N/4^r$ 的最大整数, a_{1r} 和 b_{1r} 是使得 $a_{1r} + b_{1r} = 4^r$ 和 $a_{1r} w_{1r} + b_{1r} (w_{1r} + 1) = N$ 成立的整数.

证明 对 $u \in \Omega_{1r}$, 令 $h(u) = \Delta_1(u) n_d$. 从 $\Delta_1(u)$ 的定义不难发现, $h(u)$ 是一个 $4^r \times 1$ 的向量. 而且, $1_{4^r}^T h(u) = n$, 其中 1_{4^r} 是一个 $4^r \times 1$ 的全为 1 的向量. 因此, 由引理 2, 有

$$h(u)^T h(u) \geq a_{1r} w_{1r}^2 + b_{1r} (w_{1r} + 1)^2.$$

由 (2.3) 可知引理 3 成立. □

引理 4 若设计 $d \in \mathcal{D}(N; 4^m)$, 则由 (2.4) 可得到如下不等式:

$$[\text{DD}(d, a, b)]^2 \geq \text{LB}_2[\text{DD}(d, a, b)],$$

其中

$$\text{LB}_2[\text{DD}(d, a, b)] = -\left(\frac{a+3b}{4}\right)^m + \frac{a^m}{N} + \frac{b^m}{N^2} \left(\frac{a}{b}\right)^{w_1} \left[a_1 + \left(\frac{a}{b}\right) b_1\right],$$

w_1 是不超过 $m(N-4)/4(N-1)$ 的最大整数, a_1 和 b_1 是使得

$$a_1 + b_1 = N(N-1) \quad \text{和} \quad a_1 w_1 + b_1 (1 + w_1) = \frac{mN(N-4)}{4}$$

成立的整数.

证明 由引理 2 和 (2.4) 可知引理 4 成立. □

下面的定理给出了设计 $d \in \mathcal{D}(N; 4^m)$ 在 (2.1) 下的一个下界.

定理 1 对任意的设计 $d \in \mathcal{D}(N; 4^m)$, 有

$$[\text{DD}(d, a, b)]^2 \geq \text{LB}[\text{DD}(d, a, b)],$$

其中

$$\text{LB}[\text{DD}(d, a, b)] = \max\{\text{LB}_1[\text{DD}(d, a, b)], \text{LB}_2[\text{DD}(d, a, b)]\}.$$

为比较设计 $d \in \mathcal{D}(N; 4^m)$ 的效率, 定义其效率如下:

$$\text{Eff}[\text{DD}(d, a, b)] = \frac{\text{LB}[\text{DD}(d, a, b)]}{[\text{DD}(d, a, b)]^2}. \quad (2.5)$$

当 $\text{Eff}[\text{DD}(d, a, b)] = 1$ 时, 设计 d 为均匀设计.

3 构造方法

本文的主要目的是, 基于修正 Gray 映射编码变换四水平均匀设计的构造方法. 下面的定义给出了本文所用到的变换方式 (替换规则).

定义 2 一个强度为 g 的正交表 $OA(N, m, p, g)$ 是一个由 p 个符号构成的 N 行 m 列矩阵, 使得在每一个 $N \times g$ 子矩阵中 p 个符号的所有可能行组合出现相同的次数.

下面的替换规则 R 将用于构造四水平均匀的 U -型设计.

定义 3 替换规则 R 是把两个二元列通过映射替换为一个四元列:

$$00 \rightarrow 0, \quad 01 \rightarrow 1, \quad 10 \rightarrow 2, \quad 11 \rightarrow 3.$$

注 1 在替换规则 R 下, 设计 $d^* \in \mathcal{D}^*(N; 2^{2m})$ 可替换为设计 $d \in \mathcal{D}(N; 4^m)$, 反之亦然, 即设计 $d \in \mathcal{D}(N; 4^m)$ 在替换规则 R 下亦可替换为设计 $d^* \in \mathcal{D}^*(N; 2^{2m})$. 在这个替换过程中, d^* 的列不必各不相同, 某些列可以重复用到. 进一步, 这些列可以分成 m 个组, 使得每个组是一个 $OA(N, 2, 2, 2)$.

例 1 考虑如下设计 $d_1^* \in \mathcal{D}^*(4; 2^6)$:

$$d_1^* = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 \end{pmatrix}.$$

上述矩阵是对 $OA(4, 3, 2, 2)$ 中的每个列重复一次后获得的. 基于该矩阵, 6 个列构成了三个组, 每组两个列: 第 1 和 2 列为第一组, 第 3 和 4 列为第二组, 第 5 和 6 列为第三组. 因此, 基于上面所定义的替换规则 R 可获得如下设计 $d_1 \in \mathcal{D}(4; 4^3)$:

$$d_1 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 3 \\ 2 & 3 & 1 \\ 3 & 2 & 2 \end{pmatrix}.$$

例 2 考虑设计 $d_2^* \in \mathcal{D}^*(8; 2^6)$, 该设计是删除 $OA(8, 7, 2, 2)$ 的一个列后所获得的,

$$d_2^* = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

对设计 d_2 的 6 个列作如下分组: 第 1 和 2 列为第一组, 第 3 和 4 列为第二组, 第 5 和 6 列为第三组. 因此, 基于上面所定义的替换规则 R 可获得如下设计 $d_2 \in \mathcal{D}(8; 4^3)$:

$$d_2 = \begin{pmatrix} 0 & 0 & 1 & 1 & 2 & 2 & 3 & 3 \\ 0 & 2 & 1 & 3 & 1 & 3 & 0 & 2 \\ 0 & 3 & 1 & 2 & 2 & 1 & 3 & 0 \end{pmatrix}^T.$$

4 构造设计的均匀性

本节讨论通过替换规则 R 所构造的设计的均匀性及其偏差的下界. 对于两水平设计也用离散偏差来度量设计的均匀性.

假设 d 是通过替换规则 R 所构造的属于 $\mathcal{D}(N; 4^m)$ 中的设计, 因为四水平设计 d 构造的复杂性, 所以, 偏差测度的下界作为基准在四水平最优设计的构造中非常重要. 本节的主要目标是获得通过替换规则 R 构造的设计的下界, 以作为最优设计构造的基准.

通过替换规则 R, 设计 d 的每个列替换为两个两水平列, 因为 d 是 U- 型设计, d 的每个列的每个水平出现相同的次数, 所以, 替换规则 R 把设计 $d \in \mathcal{D}(N; 4^m)$ 替换成了设计 $d^* \in \mathcal{D}^*(N; 2^{2m})$. 因此, d^* 的列可分成 m 个组使得每个组中的水平组合 00、01、10 和 11 出现相同的次数. 不失一般性, 假设第 i 个组由第 $2i-1$ 和 $2i$ 列构成, $i = 1, 2, \dots, m$.

如第 2 节中的定义那样, 令 $n_{d^*}(x)$ 为水平组合 x 在 d^* 中出现的次数, n_{d^*} 为 $n_{d^*}(x)$ 按字典顺序排列的 $v^* \times 1$ 向量. 此外, 令 c_{ij}^* 为 d^* 第 i 行与第 j 行的重合数.

对任意的设计 $d^* \in \mathcal{D}^*(N; 2^{2m})$, 其离散偏差值 $DD(d^*, a, b)$ 可由以下公式获得:

$$[DD(d^*, a, b)]^2 = -\left(\frac{a+b}{2}\right)^{2m} + \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \prod_{l=1}^{2m} K(y_{il}, y_{jl}), \quad (4.1)$$

其中, $y_{il}, y_{jl} \in \{1/4, 3/4\}$, $a > b > 0$,

$$K(y_{il}, y_{jl}) = \begin{cases} a, & \text{若 } y_{il} = y_{jl}, \\ b, & \text{若 } y_{il} \neq y_{jl}. \end{cases}$$

下面的定义将有助于我们在 U- 型设计类 $\mathcal{D}^*(N; 2^{2m})$ 中基于离散偏差如何选择最优设计.

定义 4 设计 $\tilde{d}^* \in \mathcal{D}^*(N; 2^{2m})$ 关于离散偏差是最优的当且仅当对任意的设计 $d^* \in \mathcal{D}^*(N; 2^{2m})$ 有

$$[DD(\tilde{d}^*, a, b)]^2 \leq [DD(d^*, a, b)]^2.$$

下面将讨论设计 $d^* \in \mathcal{D}^*(N; 2^{2m})$ 的离散偏差的下界. 令

$$L_2 = \begin{pmatrix} a & b \\ b & a \end{pmatrix}, \quad M_2 = \bigotimes_{j=1}^{2m} L_2, \quad \Gamma_2(0) = 1_2^T, \quad \Gamma_2(1) = I_2,$$

其中 1_2 是全为 1 的二元向量, I_2 是二阶单位阵.

不难发现, $L_2 = \beta(\Gamma_2(0))\Gamma_2(0) + \alpha(\Gamma_2(1))\Gamma_2(1)$, 其中 $\alpha = (a-b)$, $\beta = b$.

令 $\Omega_2 = \{u = u_1 \cdots u_{2m} \mid u_j = 0, 1; 1 \leq j \leq 2m\}$. 对 $0 \leq r \leq 2m$, 令 $\Omega_{2r} = \{u \in \Omega_2 \mid \sum u_j = r\}$. 对任意的 $u \in \Omega_2$, 定义矩阵

$$\Delta_2(u) = \bigotimes_{j=1}^{2m} \Gamma_2(u_j).$$

基于上述记号, 矩阵 M_2 可表示为

$$M_2 = \beta^{2m} \sum_{r=0}^{2m} \left(\frac{\alpha}{\beta}\right)^r \sum_{u \in \Omega_{2r}} (\Delta_2(u))^T \Delta_2(u). \quad (4.2)$$

由 (4.1) 和 (4.2) 可知, 对任意的设计 $d^* \in \mathcal{D}^*(N; 2^{2m})$, $[\text{DD}(d^*, a, b)]^2$ 可表示为

$$\begin{aligned} [\text{DD}(d^*, a, b)]^2 &= -\left(\frac{a+b}{2}\right)^{2m} + \frac{1}{N^2} n_{d^*}^T M_2 n_{d^*} \\ &= -\left(\frac{a+b}{2}\right)^{2m} + \frac{\beta^{2m}}{N^2} \sum_{r=0}^{2m} \left(\frac{\alpha}{\beta}\right)^r \sum_{u \in \Omega_{2r}} n_{d^*}^T (\Delta_2(u))^T \Delta_2(u) n_{d^*}. \end{aligned} \quad (4.3)$$

另一方面, 基于设计 d^* 的 $c_{i_1 i_2}^*$, (4.1) 可改写为

$$[\text{DD}(d^*, a, b)]^2 = -\left(\frac{a+b}{2}\right)^{2m} + \frac{a^{2m}}{N} + \frac{b^{2m}}{N^2} \sum_{i_1=1}^N \sum_{i_2(\neq i_1)=1}^N \left(\frac{a}{b}\right)^{c_{i_1 i_2}^*}. \quad (4.4)$$

下面的引理有助于我们在 (4.3) 和 (4.4) 下分别得到 $[\text{DD}(d^*, a, b)]^2$ 的下界.

引理 5 若设计 $d^* \in \mathcal{D}^*(N; 2^{2m})$, 则在 (4.3) 下, 有以下不等式成立:

$$[\text{DD}(d^*, a, b)]^2 \geq \text{LB}_1^*[\text{DD}(d^*, a, b)],$$

其中

$$\text{LB}_1^*[\text{DD}(d, a, b)] = -\left(\frac{a+b}{2}\right)^{2m} + \frac{\beta^{2m}}{N^2} \sum_{r=0}^{2m} \left(\frac{\alpha}{\beta}\right)^r \binom{2m}{r} [a_{2r} w_{2r}^2 + b_{2r} (w_{2r} + 1)^2],$$

w_{2r} 是不超过 $N/2^r$ 的最大整数, a_{2r} 和 b_{2r} 是使得 $a_{2r} + b_{2r} = 2^r$ 和 $a_{2r} w_{2r} + b_{2r} (w_{2r} + 1) = N$ 成立的整数.

证明 类似引理 2 的证明及 (4.3) 可知引理 5 成立. $\square$

引理 6 若设计 $d^* \in \mathcal{D}^*(N; 2^{2m})$, 则在 (4.4) 下, 有以下不等式成立:

$$[\text{DD}(d^*, a, b)]^2 \geq \text{LB}_2^*[\text{DD}(d^*, a, b)],$$

其中

$$\text{LB}_2^*[\text{DD}(d^*, a, b)] = -\left(\frac{a+b}{2}\right)^{2m} + \frac{a^{2m}}{N} + \frac{b^{2m}}{N^2} \left(\frac{a}{b}\right)^{w_2} \left[a_2 + \left(\frac{a}{b}\right) b_2\right],$$

w_2 是不超过 $m(N-2)/(N-1)$ 的最大整数, a_2 和 b_2 是使得 $a_2 + b_2 = N(N-1)$ 和 $a_2 w_2 + b_2 (1 + w_2) = mN(N-2)$ 成立的整数.

证明 由引理 2 和 (4.4) 可知, 引理 6 成立. $\square$

下面的定理给出了设计 $d^* \in \mathcal{D}^*(N; 2^{2m})$ 在 (4.1) 下均匀性测度的下界.

定理 2 对任意的设计 $d^* \in \mathcal{D}^*(N; 2^{2m})$, 有

$$[\text{DD}(d^*, a, b)]^2 \geq \text{LB}^*[\text{DD}(d^*, a, b)],$$

其中

$$\text{LB}^*[\text{DD}(d^*, a, b)] = \max\{\text{LB}_1^*[\text{DD}(d^*, a, b)], \text{LB}_2^*[\text{DD}(d^*, a, b)]\}.$$

为比较设计 $d^* \in \mathcal{D}^*(N; 2^{2m})$ 的效率, 定义其效率如下:

$$\text{Eff}^*[\text{DD}(d^*, a, b)] = \frac{\text{LB}^*[\text{DD}(d^*, a, b)]}{[\text{DD}(d^*, a, b)]^2}. \quad (4.5)$$

5 数值例子

本节通过一些数值例子来说明本文理论结果的合理性, 所有的例子都是通过替换规则 R 获得的.

例 3 考虑如下设计 $d_3^* \in \mathcal{D}^*(8; 2^{14})$, 该设计是通过对正交表 $\text{OA}(8, 7, 2, 2)$ 的每列重复两次获得的:

$$d_3^* = \begin{pmatrix} 00 & 00 & 00 & 00 & 00 & 00 & 00 \\ 00 & 01 & 10 & 01 & 11 & 11 & 10 \\ 01 & 10 & 01 & 10 & 01 & 11 & 10 \\ 01 & 11 & 11 & 11 & 10 & 00 & 00 \\ 10 & 00 & 01 & 11 & 10 & 01 & 11 \\ 10 & 01 & 11 & 10 & 01 & 10 & 01 \\ 11 & 10 & 00 & 01 & 11 & 10 & 01 \\ 11 & 11 & 10 & 00 & 00 & 01 & 11 \end{pmatrix}.$$

下面把 14 个列按上表中的顺序依次分成 7 组, 每组包含两个列, 通过替换规则 R 可获得如下设计 $d_3 \in \mathcal{D}(8; 4^7)$:

$$d_3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 3 & 3 & 2 \\ 1 & 2 & 1 & 2 & 1 & 3 & 2 \\ 1 & 3 & 3 & 3 & 2 & 0 & 0 \\ 2 & 0 & 1 & 3 & 2 & 1 & 3 \\ 2 & 1 & 3 & 2 & 1 & 2 & 1 \\ 3 & 2 & 0 & 1 & 3 & 2 & 1 \\ 3 & 3 & 2 & 0 & 0 & 1 & 3 \end{pmatrix}.$$

例 4 考虑如下设计 $d_4^* \in \mathcal{D}^*(8; 2^{42})$, 该设计是通过对正交表 $OA(8, 7, 2, 2)$ 的每个列重复 6 次获得的:

$$d_4^* = \begin{pmatrix} 00 & 00 \\ 00 & 01 & 00 & 01 & 01 & 01 & 01 & 00 & 01 & 01 & 01 & 10 & 11 & 11 & 11 & 01 & 01 & 01 & 11 & 11 & 11 \\ 01 & 00 & 01 & 00 & 01 & 01 & 10 & 11 & 10 & 11 & 11 & 01 & 00 & 01 & 01 & 10 & 11 & 11 & 01 & 01 & 11 \\ 01 & 01 & 01 & 01 & 00 & 00 & 11 & 11 & 11 & 10 & 10 & 11 & 11 & 10 & 10 & 11 & 10 & 10 & 10 & 10 & 00 \\ 10 & 10 & 11 & 11 & 10 & 11 & 00 & 01 & 01 & 00 & 01 & 01 & 01 & 00 & 01 & 11 & 10 & 11 & 10 & 11 & 01 \\ 10 & 11 & 11 & 10 & 11 & 10 & 01 & 01 & 00 & 01 & 00 & 11 & 10 & 11 & 10 & 10 & 11 & 10 & 01 & 00 & 10 \\ 11 & 10 & 10 & 11 & 11 & 10 & 10 & 10 & 11 & 11 & 10 & 00 & 01 & 01 & 00 & 01 & 01 & 00 & 11 & 10 & 10 \\ 11 & 11 & 10 & 10 & 10 & 11 & 11 & 10 & 10 & 10 & 11 & 10 & 10 & 10 & 11 & 00 & 00 & 01 & 00 & 01 & 01 \end{pmatrix}.$$

下面把 42 个列按上表中的顺序依次分成 21 组, 每组包含两个列, 通过替换规则 R 可获得如下设计 $d_4 \in \mathcal{D}(8; 4^{21})$:

$$d_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 2 & 3 & 3 & 3 & 1 & 1 & 1 & 3 & 3 & 3 \\ 1 & 0 & 1 & 0 & 1 & 1 & 2 & 3 & 2 & 3 & 3 & 1 & 0 & 1 & 1 & 2 & 3 & 3 & 1 & 1 & 3 \\ 1 & 1 & 1 & 1 & 0 & 0 & 3 & 3 & 3 & 2 & 2 & 3 & 3 & 2 & 2 & 3 & 2 & 2 & 2 & 2 & 0 \\ 2 & 2 & 3 & 3 & 2 & 3 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 3 & 2 & 3 & 2 & 3 & 1 \\ 2 & 3 & 3 & 2 & 3 & 2 & 1 & 1 & 0 & 1 & 0 & 3 & 2 & 3 & 2 & 2 & 3 & 2 & 1 & 0 & 2 \\ 3 & 2 & 2 & 3 & 3 & 2 & 2 & 2 & 3 & 3 & 2 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 3 & 2 & 2 \\ 3 & 3 & 2 & 2 & 2 & 3 & 3 & 2 & 2 & 2 & 3 & 2 & 2 & 2 & 3 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

例 5 考虑如下设计 $d_5^* \in \mathcal{D}^*(12; 2^{10})$, 该设计是通过对正交表 $OA(12, 11, 2, 2)$ 删除一列获得的:

$$d_5^* = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \end{pmatrix}.$$

下面把上表中的 10 个列分成 5 个组, 每组含两个列, 第 i 个组由列 $(2i-1, 2i)$ ($i = 1, 2, \dots, 5$) 构成,

根据替换规则 R 获得如下设计 $d_5 \in \mathcal{D}(12; 4^5)$:

$$d_5 = \begin{pmatrix} 3 & 0 & 2 & 1 & 2 & 3 & 3 & 1 & 0 & 0 & 2 & 1 \\ 3 & 2 & 1 & 0 & 2 & 1 & 2 & 3 & 3 & 1 & 0 & 0 \\ 3 & 0 & 0 & 2 & 1 & 0 & 2 & 1 & 2 & 3 & 3 & 1 \\ 3 & 3 & 1 & 0 & 0 & 2 & 1 & 0 & 2 & 1 & 2 & 3 \\ 3 & 2 & 3 & 3 & 1 & 0 & 0 & 2 & 1 & 0 & 2 & 1 \end{pmatrix}^T.$$

例 6 考虑如下设计 $d_6^* \in \mathcal{D}^*(12; 2^{22})$, 该设计是通过对正交表 $OA(12, 11, 2, 2)$ 的每列重复两次获得的:

$$d_6^* = \begin{pmatrix} 11 & 11 & 11 & 11 & 11 & 11 & 11 & 11 & 11 & 11 & 11 & 11 \\ 00 & 01 & 10 & 00 & 00 & 01 & 11 & 11 & 10 & 01 & 01 & 01 \\ 10 & 00 & 01 & 10 & 00 & 00 & 01 & 11 & 11 & 10 & 10 & 10 \\ 01 & 10 & 00 & 01 & 10 & 00 & 00 & 01 & 11 & 11 & 01 & 01 \\ 10 & 01 & 10 & 00 & 01 & 10 & 00 & 00 & 01 & 11 & 11 & 11 \\ 11 & 10 & 01 & 10 & 00 & 01 & 10 & 00 & 00 & 01 & 11 & 11 \\ 11 & 11 & 10 & 01 & 10 & 00 & 01 & 10 & 00 & 00 & 10 & 10 \\ 01 & 11 & 11 & 10 & 01 & 10 & 00 & 01 & 10 & 00 & 00 & 00 \\ 00 & 01 & 11 & 11 & 10 & 01 & 10 & 00 & 01 & 10 & 00 & 00 \\ 00 & 00 & 01 & 11 & 11 & 10 & 01 & 10 & 00 & 01 & 01 & 01 \\ 10 & 00 & 00 & 01 & 11 & 11 & 10 & 01 & 10 & 00 & 10 & 10 \\ 01 & 10 & 00 & 00 & 01 & 11 & 11 & 10 & 01 & 10 & 00 & 00 \end{pmatrix}.$$

把 22 个列按上表中的顺序依次分成 11 组, 每组包含两个列, 通过替换规则 R 可获得如下设计 $d_6 \in \mathcal{D}(12; 4^{11})$:

$$d_6 = \begin{pmatrix} 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 \\ 0 & 1 & 2 & 0 & 0 & 1 & 3 & 3 & 2 & 1 & 1 \\ 2 & 0 & 1 & 2 & 0 & 0 & 1 & 3 & 3 & 2 & 2 \\ 1 & 2 & 0 & 1 & 2 & 0 & 0 & 1 & 3 & 3 & 1 \\ 2 & 1 & 2 & 0 & 1 & 2 & 0 & 0 & 1 & 3 & 3 \\ 3 & 2 & 1 & 2 & 0 & 1 & 2 & 0 & 0 & 1 & 3 \\ 3 & 3 & 2 & 1 & 2 & 0 & 1 & 2 & 0 & 0 & 2 \\ 1 & 3 & 3 & 2 & 1 & 2 & 0 & 1 & 2 & 0 & 0 \\ 0 & 1 & 3 & 3 & 2 & 1 & 2 & 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 3 & 3 & 2 & 1 & 2 & 0 & 1 & 1 \\ 2 & 0 & 0 & 1 & 3 & 3 & 2 & 1 & 2 & 0 & 2 \\ 1 & 2 & 0 & 0 & 1 & 3 & 3 & 2 & 1 & 1 & 0 \end{pmatrix}.$$

表 1 设计 $d^* \in \mathcal{D}^*(N; 2^{2m})$ 的均匀性和效率

例子	d^*	N	m	$[\text{DD}(d^*, a, b)]^2$	$\text{LB}^*[\text{DD}(d^*, a, b)]$	$\text{Eff}^*[\text{DD}(d^*, a, b)]$
例 1	d_1^*	4	3	0.1189	0.1189	1.0000
例 2	d_2^*	8	3	0.0330	0.0330	1.0000
例 3	d_3^*	8	7	0.1106	0.1106	1.0000
例 4	d_4^*	8	21	0.1250	0.1250	1.0000
例 5	d_5^*	12	5	0.0492	0.0492	1.0000
例 6	d_6^*	12	11	0.0818	0.0818	1.0000

表 2 构造的四水平设计 $d \in \mathcal{D}(N; 4^m)$ 的均匀性和效率

例子	d	N	m	$[\text{DD}(d, a, b)]^2$	$\text{LB}[\text{DD}(d, a, b)]$	$\text{Eff}[\text{DD}(d, a, b)]$
例 1	d_1	4	3	0.0996	0.0996	1.0000
例 2	d_2	8	3	0.0371	0.0371	1.0000
例 3	d_3	8	7	0.1034	0.1014	0.9807
例 4	d_4	8	21	0.1250	0.1250	1.0000
例 5	d_5	12	5	0.0470	0.0427	0.9085
例 6	d_6	12	11	0.0798	0.0794	0.9950

表 1 给出了基于替换规则 R 来构造四水平设计的两水平设计 $d^* \in \mathcal{D}^*(N; 2^{2m})$ 的效率, 其效率是通过离散偏差来衡量的, 其中 $a = 1, b = 0.5$. 根据表 1 不难发现, 基于替换规则 R 来构造四水平设计的两水平设计 $d^* \in \mathcal{D}^*(N; 2^{2m})$ 在离散偏差下都是均匀设计.

表 2 给出了所构造的四水平设计的效率. 根据表 2 不难发现, 基于替换规则 R 利用两水平均匀设计所构造的四水平设计是均匀的或者近似均匀的 ($\text{Eff} \geq 90\%$).

6 结论

本文基于修正 Gray 映射编码提出了替换规则 R, 其主要目的是基于两水平设计并利用替换规则 R 来构造四水平设计, 其均匀性是在离散偏差意义下而言的, 并获得了相关设计的离散偏差的下界. 从例 1–6 的设计表及其表 1 和 2 中的数值计算结果发现, 由替换规则 R 所构造的四水平设计具有非常高的效率. 此外, 替换规则 R 还能有效地用于四水平设计的折叠反转设计中, 以及二、四混水平的最优设计.

本文所考虑的替换规则 R 为构造高效率的四水平均匀设计提供了一个新的且值得进一步研究的方向. 沿着此方向, 有关高效率均匀设计的构造将是一个非常有意义的问题, 在未来将进一步进行研究.

参考文献

- 1 Bates R A, Buck R J, Riccomagno E, et al. Experimental design and observation for large systems. J R Stat Soc Ser B Stat Methodol, 1996, 58: 77–94

- 2 Koehler J R, Owen A B. Computer experiments. In: Handbook of Statistics, vol. 13. Amsterdam: Elsevier, 1996, 261–308
- 3 Fang K T, Li R, Sudjianto A. Design and Modeling for Computer Experiments. London: Chapman and Hall, 2005
- 4 Hickernell F J, Liu M Q. Uniform designs limit aliasing. Biometrika, 2002, 89: 893–904
- 5 Xu H, Wong A. Two-level nonregular designs from quaternary linear codes. Statist Sinica, 2007, 17: 1191–1213
- 6 Phoa F K H, Xu H. Quarter-fraction factorial designs constructed via quaternary codes. Ann Statist, 2009, 37: 2561–2581
- 7 Zhang R C, Phoa F K H, Mukerjee R, et al. A trigonometric approach to quaternary code designs with application to one-eighth and one-sixteenth fractions. Ann Statist, 2011, 39: 931–955
- 8 Phoa F K H. A code arithmetic approach for quaternary code designs and its application to $(1/64)$ th-fraction. Ann Statist, 2012, 40: 3161–3175
- 9 Xu H, Phoa F K H, Wong W K. Recent developments in nonregular fractional factorial designs. Statist Surv, 2009, 3: 18–46
- 10 Qin H, Fang K T. Discrete discrepancy in factorial designs. Metrika, 2004, 60: 59–72
- 11 Chatterjee K, Li Z H, Qin H. Some new lower bounds to centered and wrap-round L_2 -discrepancies. Statist Probab Lett, 2012, 82: 1367–1373

Construction of four-level designs for computer experiments

QIN Hong, OU ZuJun & CHATTERJEE Kashinath

Abstract Application of coding theory is quite familiar in the context of constructing optimal designs. The present paper aims at constructing uniform designs for computer experiments via a modified Gray map code transformation. The efficiency of these constructed designs is measured through the popular discrete discrepancy measure introduced by Hickernell and Liu (2002). As a benchmark, lower bound of the uniformity measure under consideration is also derived for obtaining the best design.

Keywords efficiency, lower bound, modified Gray map, uniform design

MSC(2010) 62K15, 62K05, 62K99

doi: 10.1360/SCM-2016-0788