

混料试验的正交格点填充设计^①

李光辉¹, 李俊鹏^{1,2}, 张崇岐²

1. 凯里学院 理学院, 贵州 凯里 556011; 2. 广州大学 经济与统计学院, 广州 510006

摘要: 提出一种线性变换的方法, 可以实现将超立方体内的设计点变换到混料试验域内, 并且保持设计点之间的距离不变. 将超立方体内的正交格子点集变换到混料试验域内, 并在此基础上建立非参数回归模型. 试验表明, 这种方法构造的设计具有很好的均匀性, 建立的非参数模型 also 具有很好的稳健性.

关键词: 混料试验; 正交格点; 线性变换

中图分类号: O212.6

文献标志码: A

文章编号: 1000-5471(2022)09-0048-07

Orthogonal Lattice Filling Design of Mixture Experiment

LI Guanghui¹, LI Junpeng^{1,2}, ZHANG Chongqi²

1. School of Science, Kaili University, Kaili Guizhou 556011, China;

2. School of Economics and Statistics, Guangzhou University, Guangzhou 510006, China

Abstract: In this paper, a method has been proposed to transform the design points in the hypercube to the mixture experimental region, and the distance been kept between the design points unchanged. The orthogonal grid point set in hypercube has been transformed into the mixture experimental region, and the nonparametric regression model is established. Through the example analysis, the design constructed by this method has good uniformity, and the nonparametric model also has good robustness.

Key words: mixture experiment; orthogonal lattice; linear transformation

混料试验设计是一类特殊的设计, 它广泛应用于生物、医学、工业以及科研实践中^[1-4]. 混料试验的试验域是 $q-1$ 维正规单纯形

$$S^{q-1} = \{(x_1, x_2, \dots, x_q) : \sum_{i=1}^q x_i = 1, x_j \geq 0, j = 1, 2, \dots, q\}$$

为了使得试验点能均匀分布在试验域内, 许多统计学者尝试了各种不同的方法. 文献 [5] 介绍了混料试验设计的基本概念及应用. 关于混料均匀设计方面的研究可参见文献 [6-9].

文献 [10-11] 研究了混料试验域上构造均匀设计的方法; 文献 [12] 讨论了在具有附加约束的混料试验域内构造均匀设计的方法; 文献 [13] 给出了一种单纯形内的切片设计. 这些方法所研究的设计都是在单纯形内部分布均匀, 且在某种均匀性准则下能达到较好的水准. 由于混料试验通常使用回归模型来建模, 所

① 收稿日期: 2021-03-09

基金项目: 贵州省科学技术基金项目(黔科合基础[2020]1Y010); 国家自然科学基金青年项目(11901260); 凯里学院博士专项课题(BS201807).

作者简介: 李光辉, 博士, 教授, 主要从事试验设计的研究.

以均匀设计具有很好的稳健性.

有时需要在混料试验域内采用非参数回归模型来建模, 除了需要大量的试验点外, 还需要满足“试验点之间的距离尽可能相等”这一条件. 混料格子点集虽然能均匀分布在混料试验域的内部及边界, 但是当混料分量数超过 3 以后, 相邻的格点之间的距离是不相等的, 并且由混料格子点集剖分得到的子单纯形也不完全相等. 这使得使用混料格子点集构造的设计不利于使用非参数回归模型来建模. 参考超立方体内的建模方法, 正交格子点集不仅排布均匀, 相邻的格点之间的距离相等, 并且超立方体内的完全析因设计有诸多优良的性质.

关于超立方体上的格点设计已有一些研究成果. 文献[14-15]研究了交错网格的极大极小距离设计, 以及旋转球填充设计. 这些空间填充设计都是基于格子点集来研究的, 并且在超立方体上有很多好的性质.

鉴于此, 我们考虑将超立方体内的正交格子点集变换到混料试验域内, 并且保证变换后试验点之间的距离不变. 根据正交格子点集的性质, 研究在混料试验域上正交格点设计的性质. 首先给出一种线性变换, 能将 $\mathbb{R}^{q-1}$ 的点集变换到 S^{q-1} 中, 并且保证点之间的距离不变; 然后介绍在单纯形内部, 将试验点以旋转或添加的方式, 使得设计在混料试验域内更为均匀; 最后给出实例分析, 说明单纯形内的正交格点设计具有很好的均匀性, 并且非参数建模也具有稳健性.

1 基本记号

对于 q 分量混料系统, 若试验还受到其他约束条件的限制, 我们将试验域记作

$$\mathcal{X} = \{\mathbf{x} = (x_1, x_2, \dots, x_q)^T : \mathbf{x} \in S^{q-1}, a_j \leq \varphi_j(\mathbf{x}) \leq b_j, j = 1, 2, \dots, k\}$$

其中: $\varphi_j(\mathbf{x})$ 是关于 $\mathbf{x} \in S^{q-1}$ 的已知函数; $a_j, b_j (j = 1, 2, \dots, k)$ 为已知常数. 通常 $\mathcal{X}$ 是 S^{q-1} 内部的一个不规则凸几何体.

对于任意的 $\mathbf{x} = (x_1, x_2, \dots, x_q)^T \in S^{q-1} \subset \mathcal{X}$ 以及一个 n 点设计

$$\mathcal{P}_n = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\} \subset \mathcal{X}$$

我们给出点到点集的距离定义为

$$d^2(\mathbf{x}, \mathcal{P}_n) = \min_{1 \leq i \leq n} \{d^2(\mathbf{x}, \mathbf{x}_i)\}$$

其中 $d^2(\mathbf{x}, \mathbf{x}_i) = \|\mathbf{x} - \mathbf{x}_i\|^2$ 为欧氏距离的平方.

文献[16]提出了 3 种度量混料试验域设计的均匀性指标, 即偏差的定义. 为方便讨论, 我们主要讨论均方误差偏差和最大距离偏差.

1) 均方误差偏差

$$MSED(\mathcal{P}_n) = E(d^2(\mathbf{x}, \mathcal{P}_n)) = \int_{\mathcal{X}} \frac{1}{\text{Vol}(\mathcal{X})} d^2(\mathbf{x}, \mathcal{P}_n) d\mathbf{x}$$

2) 最大距离偏差

$$MD(\mathcal{P}_n) = \max_{\mathbf{x} \in \mathcal{X}} d^2(\mathbf{x}, \mathcal{P}_n)$$

以上两种偏差的计算较为复杂, 在实际情形下一般都是采用 Monte Carlo 方法来计算近似值. 首先生成试验域内的随机混料点集, 或使用高阶格子点集

$$\mathcal{L}(q, m) = \left\{ \left(\frac{\alpha_1}{m}, \frac{\alpha_2}{m}, \dots, \frac{\alpha_q}{m} \right) : \alpha_i \in \mathbb{Z}^+, i = 1, 2, \dots, q, \sum_{i=1}^q \alpha_i = m \right\}$$

令 $\mathcal{L}_{\mathcal{X}} = \mathcal{L}(q, m) \cap \mathcal{X} = \{\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_N\}$ 作为度量均匀性的 NT-net, 再令

$$msed(\mathcal{P}_n) = \frac{1}{N} \sum_{k=1}^N d^2(\mathbf{t}_k, \mathcal{P}_n); md(\mathcal{P}_n) = \max_{1 \leq k \leq N} \{d^2(\mathbf{t}_k, \mathcal{P}_n)\} \quad (1)$$

使用 $msed(\mathcal{P}_n)$ 和 $md(\mathcal{P}_n)$ 分别作为 $MSED(\mathcal{P}_n)$ 和 $MD(\mathcal{P}_n)$ 的近似值. 在实际计算中, 一般要求 $N > 1\,000$.

$q-1$ 维空间中的正交格子点集可以表示为

$$\mathbf{y}_j = c \sum_{i=1}^{q-1} \alpha_{ji} \mathbf{e}_{q-1}(i) \quad (2)$$

其中: $\alpha_{j1}, \alpha_{j2}, \dots, \alpha_{j, q-1} \in \mathbb{Z}$; $\mathbf{e}_{q-1}(i)$ 为 $q-1$ 维向量, 其第 i 个元素为 1, 其余元素为 0; c 是已知的正数, 它表示正交格点之间的距离. 我们记由 (2) 式生成的正交格子点集为

$$\mathcal{H}(q-1, c) = \left\{ \mathbf{y}: \mathbf{y}_j = c \sum_{i=1}^{q-1} \alpha_{ji} \mathbf{e}_{q-1}(i), \alpha_{ji} \in \mathbb{Z}, j=1, 2, \dots \right\}$$

对于任意的 $\mathbf{y} \in \mathcal{H}(q-1, c)$, 定义 $\mathbf{y}_0$ 的 voronoi 单元为

$$\text{Vor}(\mathbf{y}_0) = \{ \mathbf{z}: \|\mathbf{z} - \mathbf{y}_0\|^2 \leq \|\mathbf{z} - \mathbf{y}\|^2, \forall \mathbf{y} \in \mathcal{H}(q-1, c) \}$$

这样, 我们可以考虑将超立方体内的试验点集转化到单纯形中, 并且保持格点之间的距离不变.

2 保距独立变换

为了实现将 $q-1$ 维超立方体内的正交格子点集变换到单纯形中, 我们构造一种线性变换, 将 $\mathbb{R}^{q-1}$ 中的点变换到 S^{q-1} 中, 且点与点之间的距离保持不变. 文献[17] 的单纯形保距独立变换方法如下:

定义一个 $q-1$ 阶矩阵

$$\mathbf{W} = \{w_{ij}\}_{i,j=1}^{q-1} = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{q-1}) \quad (3)$$

其中: $\mathbf{w}_1 = (\sqrt{2}, 0, \dots, 0)^T$, $\mathbf{w}_i = \frac{1}{i} \sum_{j=1}^{i-1} \mathbf{w}_j + \mathbf{e}_{q-1}(i) \sqrt{1 + \frac{1}{i}}$, $i=2, 3, \dots, q-1$.

记 $\mathbf{I}_{q-1}$ 为 $q-1$ 阶单位矩阵, $\mathbf{1}_{q-1}$ 为元素全为 1 的 $q-1$ 维列向量, $\mathbf{J}_{q-1} = \mathbf{1}_{q-1} \mathbf{1}_{q-1}^T$, 易知, (3) 式中的矩阵 $\mathbf{W}$ 具有以下性质:

$$\mathbf{W}^T \mathbf{W} = \mathbf{I}_{q-1} + \mathbf{J}_{q-1}$$

对于 $\mathbf{W}$ 的任意两列 $\mathbf{w}_i, \mathbf{w}_j$, $i \neq j$, 有

$$\|\mathbf{w}_i - \mathbf{w}_j\| = \sqrt{2}$$

$\mathbf{W}$ 的行列式 $\det(\mathbf{W}) = \sqrt{q}$.

这种变换可以将单纯形 S^{q-1} 中的点映射到 $\mathbb{R}^{q-1}$ 中.

例如, 对于任意的 $\mathbf{x} \in S^{q-1}$, 记 $\mathbf{x}_{-i} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_q)^T$, 则线性变换后的点 $\mathbf{y} = \mathbf{W}\mathbf{x}_{-i} \in \mathbb{R}^{q-1}$. 反之, 对于任意的 $\mathbf{y} \in \mathbb{R}^{q-1}$, 令 $\mathbf{x}_{-i} = \mathbf{W}^{-1}\mathbf{y}$, 再令 $x_i = 1 - \mathbf{x}_{-i}^T \mathbf{1}_{q-1}$, 这样得到的点 $\mathbf{x} = (x_1, x_2, \dots, x_q)^T$ 是位于平面 $\sum_{i=1}^q x_i = 1$ 上的点, 经过筛选可得到落入 S^{q-1} 中的正交格子点集.

对于任意的 $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{iq})^T \in S^{q-1}$, 记 $\mathbf{x}_{-1,i} = (x_{i2}, x_{i3}, \dots, x_{iq})^T$, 即 $\mathbf{y}_i = \mathbf{W}\mathbf{x}_{-1,i}$, $i=1, 2$, 有 $\|\mathbf{y}_1 - \mathbf{y}_2\| = \|\mathbf{x}_1 - \mathbf{x}_2\|$ 恒成立.

这是因为由矩阵 $\mathbf{W}$ 的性质可知

$$\begin{aligned} \|\mathbf{x}_1 - \mathbf{x}_2\|^2 &= (\mathbf{x}_1 - \mathbf{x}_2)^T (\mathbf{x}_1 - \mathbf{x}_2) = \sum_{j=2}^q (x_{1j} - x_{2j})^2 + (x_{11} - x_{21})^2 = \\ &= \sum_{j=2}^q (x_{1j} - x_{2j})^2 + \left[\left(1 - \sum_{j=2}^q x_{1j}\right) - \left(1 - \sum_{j=q}^q x_{2j}\right) \right]^2 = \\ &= \sum_{j=2}^q (x_{1j} - x_{2j})^2 + \left[\sum_{j=2}^q (x_{1j} - x_{2j}) \right]^2 = \\ &= 2 \sum_{j=2}^q (x_{1j} - x_{2j})^2 + 2 \sum_{2 \leq j < k \leq q} (x_{1j} - x_{2j})(x_{1k} - x_{2k}) = \\ &= (\mathbf{x}_{-1,1} - \mathbf{x}_{-1,2})^T (\mathbf{1}_{q-1} + \mathbf{J}_{q-1}) (\mathbf{x}_{-1,1} - \mathbf{x}_{-1,2}) = \\ &= [\mathbf{W}(\mathbf{x}\mathbf{x}_{-1,1} - \mathbf{x}_{-1,2})]^T \mathbf{W}(\mathbf{x}_{-1,1} - \mathbf{x}_{-1,2}) = \\ &= \|\mathbf{y}_1 - \mathbf{y}_2\|^2 \end{aligned}$$

由此可见, 经过变换后任意两点在 $\mathbb{R}^{q-1}$ 中的距离与 S^{q-1} 中是一致的.

3 填充设计

我们按照以下方式生成 S^{q-1} 中的正交格点.

首先生成 $\mathbb{R}^{q-1}$ 中的正交格子点集 $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N\} \subset \mathcal{H}\{q-1, c\}$, 在混料试验域中选择一点 $\mathbf{x}_0 = (x_{01}, x_{02}, \dots, x_{0q})^\top \in \mathcal{X}$ 作为参照点, 令 $\mathbf{y}_0 = \mathbf{W}\mathbf{x}_{-1,0}$, 其中

$$\mathbf{x}_{-1,0} = (x_{02}, x_{03}, \dots, x_{0q})^\top$$

其次, 令矩阵

$$\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N)^\top + \mathbf{1}_N \mathbf{y}_0 \quad (4)$$

是以 $\mathbf{y}_0$ 作为原点而形成的正交格子点集.

最后, 进行线性变换, 令矩阵

$$\mathbf{X} = (\mathbf{1}_N - \mathbf{X}_{-1} \mathbf{1}_{q-1}, \mathbf{X}_{-1}) = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)^\top \quad (5)$$

其中 $\mathbf{X}_{-1}^\top = \mathbf{W}^{-1} \mathbf{Y}^\top$. 将 $\mathbf{X}$ 中的点进行筛选, 记 $\mathcal{H}'\{q-1, c\} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ 是由 $\mathcal{H}\{q-1, c\}$ 经过逆变换而得的试验点集, 令

$$\mathcal{H}_{\mathcal{X}} = \mathcal{H}'\{q-1, c\} \cap \mathcal{X} \quad (6)$$

表示最终确定落入试验域内的正交格子点集.

将一个超立方体 $[a_1, b_1] \times [a_2, b_2] \times \dots \times [a_q, b_q]$ 记为 $H_0 = \prod_{i=1}^q [a_i, b_i]$, 该超立方体的顶点集构成的矩阵可表示为

$$(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{2^q})^\top = ((a_1, b_1)^\top \otimes \mathbf{1}_{2^{q-1}}, \mathbf{1}_2 \otimes (a_2, b_2)^\top \otimes \mathbf{1}_{2^{q-3}}, \dots, \mathbf{1}_{2^{q-1}} \otimes (a_q, b_q)^\top)$$

则有以下定理.

定理 1 在超立方体 $H_0 = \prod_{i=1}^q [a_i, b_i]$ 上的顶点设计为 $\mathcal{P} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{2^q}\}$, $c = b_i - a_i$, $i = 1, 2, \dots, q$, 则

$$MSED(\mathcal{P}) = \frac{qc^2}{12}, MD(\mathcal{P}) = \frac{qc^2}{4}$$

证 首先证明对于任意点 $\mathbf{x}_0 = (x_{01}, x_{02}, \dots, x_{0q})^\top \in H_0$, 有

$$\begin{aligned} \int_{H_0} d^2(\mathbf{x}, \mathbf{x}_0) d\mathbf{x} &= \sum_{i=1}^q \int_{a_1}^{b_1} \dots \int_{a_q}^{b_q} x_i^2 dx_1 \dots dx_q - 2 \sum_{i=1}^q x_{0i} \int_{a_1}^{b_1} \dots \int_{a_q}^{b_q} x_i dx_1 \dots dx_q + \left(\prod_{i=1}^q (b_i - a_i) \right) \sum_{i=1}^q x_{0i}^2 = \\ &= \frac{1}{3} qc^{q+2} - 2c^{q+1} \mathbf{x}_0^\top \mathbf{1}_q + c^q \mathbf{x}_0^\top \mathbf{x}_0 \end{aligned}$$

特别地, 若 $\mathbf{x}_0 \in \mathcal{P}$, 则 $\int_{H_0} d^2(\mathbf{x}, \mathbf{x}_0) d\mathbf{x} = \frac{1}{3} qc^{q+2}$. 考虑将 H_0 每一维的区间二等分, 记

$$H_j = \prod_{i=1}^q [a_{ij}, b_{ij}], j = 1, 2, \dots, 2^q$$

其中 $[a_{ij}, b_{ij}] = \left[a_i, a_i + \frac{c}{2} \right]$ 或 $[a_{ij}, b_{ij}] = \left[a_i + \frac{c}{2}, b_i \right]$. 这样一共可以剖分为 2^q 个子立方体. 所以有

$$\begin{aligned} MSED(\mathcal{P}) &= \frac{1}{\int_{H_0} d\mathbf{x}} \int_{H_0} d^2(\mathbf{x}, \mathcal{P}) d\mathbf{x} = \frac{1}{c^q} \sum_{j=1}^{2^q} \int_{H_j} d^2(\mathbf{x}, \mathbf{x}_j) d\mathbf{x} = \\ &= \frac{1}{c^q} \frac{2^q}{3} q \left(\frac{c}{2} \right)^{q+2} = \frac{qc^2}{12} \end{aligned}$$

$\max_{\mathbf{x} \in H_0} d^2(\mathbf{x}, \mathcal{P})$ 恰好是立方体中心 $(a_1, a_2, \dots, a_q)^\top + \frac{c}{2} \mathbf{1}_q$ 到某个顶点的距离, 即有 $MD(\mathcal{P}) = \frac{qc^2}{4}$.

为了使得试验域 $\mathcal{X}$ 内的正交格子点集具有良好的均匀性, 即使得设计

$$\mathcal{P}_n = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\} = \mathcal{H}'\{q-1, c\} \cap \mathcal{X}$$

的均方误差 $MSED(\mathcal{P}_n)$ 与最大距离偏差 $MD(\mathcal{P}_n)$ 尽可能地小, 我们考虑从以下两个方面来实现.

4 旋转变换

在 $\mathcal{X}$ 中实施旋转变换, 使得落入 $\mathcal{X}$ 内的正交格点尽可能地多. 我们令 $q-1$ 阶方阵为

$$\mathbf{R}_{ij}(\theta_{ij}) = (\mathbf{e}_{q-1}(1), \dots, \mathbf{e}_{q-1}(i-1), \mathbf{E}_i, \mathbf{e}_{q-1}(i+1), \dots, \mathbf{e}_{q-1}(j-1), \mathbf{E}_j, \mathbf{e}_{q-1}(j+1), \dots, \mathbf{e}_{q-1}(q-1))$$

其中 $\mathbf{E}_i = \mathbf{e}_{q-1}(i)\cos\theta_{ij} - \mathbf{e}_{q-1}(j)\sin\theta_{ij}$, $\mathbf{E}_j = \mathbf{e}_{q-1}(j)\sin\theta_{ij} + \mathbf{e}_{q-1}(i)\cos\theta_{ij}$. 在(4)式的基础上, 再令

$$\mathbf{Y}(\Theta, c) = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N)^T \mathbf{R}_{12}(\theta_{12}) \mathbf{R}_{13}(\theta_{13}) \cdots \mathbf{R}_{(q-2)(q-1)}(\theta_{(q-2)(q-1)}) + \mathbf{1}_N \mathbf{y}_0^T$$

其中 $\Theta = (\theta_{12}, \theta_{13}, \dots, \theta_{(q-2)(q-1)})$ 为旋转参数向量. 将 $\mathbf{Y}(\Theta)$ 经过(5)式的逆变换后得

$$(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \mathbf{W}^{-1} \mathbf{Y}^T(\Theta, c) \quad (7)$$

令 $\mathcal{H}_{\mathcal{X}} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\} \cap \mathcal{X}$ 表示最终确定落入试验域内的正交格子点集. 通过调整(7)式中的参数 Θ , 使得 $\mathcal{H}_{\mathcal{X}}$ 达到满足要求的均匀性.

5 补充设计点

$\mathcal{X}$ 内生成的正交格子点集为 $\mathcal{P}_n = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$, 我们采用在 $\mathcal{X} \setminus \{\bigcup_{i=1}^n \text{Vor}(\mathbf{x}_i)\}$ 中部分添加若干个试验点的方式, 使得最后所得设计的均方误差或最大距离偏差尽可能地小.

首先, 在 $\mathcal{X}$ 内使用高阶混料格子点集

$$\mathcal{L}_{\mathcal{X}} = \mathcal{L}\{q, m\} \cap \mathcal{X} = \{\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_{N_1}\}$$

作为在 $\mathcal{X}$ 上度量一个设计所使用的 NT-net. 令

$$\mathbf{x}_{n+1}^{(0)} = \arg \min_{1 \leq i \leq N_1} \text{msed}\{\mathcal{P}_n, \mathbf{t}_i\}$$

然后使用文献[18]中使用的随机搜索算法, 令矩阵 $\mathbf{R} = \{r_{ij}\}_{i,j=1}^{N,q}$, 其中

$$r_{ij} = \frac{\log(1 - u_{ij})}{\sum_{j=1}^q \log(1 - u_{ij})}$$

这里 $u_{ij} (i=1, 2, \dots, N; j=1, 2, \dots, q)$ 相互独立, 且都服从均匀分布 $u_{ij} \sim U(0, 1)$.

使用文献[19]的拟分量变换方法, 选择适当的拟分量变换参数 $\lambda > 0$, 令

$$\mathbf{Z}(\mathbf{R}, \mathbf{x}_{n+1}^{(j)}, \lambda) = (\mathbf{z}_1^{(j)}, \mathbf{z}_2^{(j)}, \dots, \mathbf{z}_N^{(j)})^T = \frac{1}{1+\lambda} (\mathbf{1}_N \mathbf{x}_{n+1}^{(j)T} + \lambda \mathbf{R})$$

其中 $\mathbf{z}_i^{(j)} = (z_{i1}^{(j)}, z_{i2}^{(j)}, \dots, z_{iq}^{(j)})^T$, 令 $\mathbf{z}_{n+1}^{(j+1)} = \arg \min_{1 \leq i \leq N_1} \text{msed}\{\mathcal{P}_n, \mathbf{z}_i^{(j)}\}$. 经过若干轮迭代, 最终确定 $\mathbf{x}_{n+1}$ 使得

$$\text{msed}\{\mathcal{P}_n, \mathbf{x}_{n+1}\} \rightarrow \inf$$

以此类推, 按照上述方法在 $\mathcal{X} \setminus \{\bigcup_{i=1}^{n+1} \text{Vor}(\mathbf{x}_i)\}$ 中再添加点 $\mathbf{x}_{n+2}$, 最终得到一个设计

$$\mathcal{P}_{n+k} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n, \dots, \mathbf{x}_{n+k}\} \quad (8)$$

如果一个 voronoi 单元的对角线长度为 d , 按照添加点的方法构造的设计能使得 $MD(\mathcal{P}_{n+k}) = \frac{d}{2}$.

考虑在试验域 $\mathcal{X}$ 上使用正交格子点集来建模, 按(7)式或(8)式生成 $\mathcal{X}$ 上的正交格子点集记为 $\mathcal{P}_N = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$. 定义高斯核函数为

$$K_h(\mathbf{x}_i, \mathbf{x}_j) = \exp\left\{-\frac{(\mathbf{x}_i - \mathbf{x}_j)^T(\mathbf{x}_i - \mathbf{x}_j)}{2h^2}\right\}, \mathbf{x}_i, \mathbf{x}_j \in \mathcal{X}$$

其中 h 为已知的窗宽参数, 若试验在各点处的观测值为 $y_1, y_2, \dots, y_N$, 则基于正交格点的非参数回归模型为

$$\eta(\mathbf{x}) = \frac{\sum_{j=1}^N K_h(\mathbf{x}, \mathbf{x}_j) y_j}{\sum_{j=1}^N K_h(\mathbf{x}, \mathbf{x}_j)}, \mathbf{x} \in \mathcal{X}, \mathbf{x}_j \in \mathcal{P}_N$$

如果在试验域内有足够多的格点, 选择适合的窗宽, 能使得非参数回归模型具有很好的预测效果.

例 1 在 S^{3-1} 上构造一个正交格点设计, 使得各相邻点之间的距离为 0.2, 通过旋转过后设计点的分布以及试验域内各个点到该设计的距离等高线图如图 1 所示.

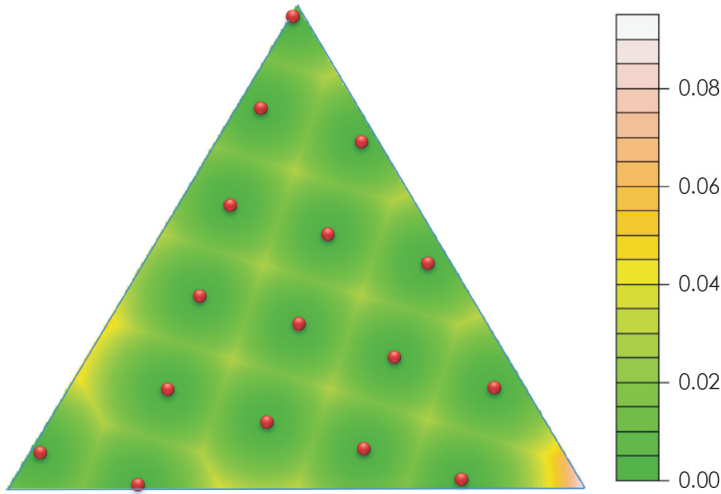
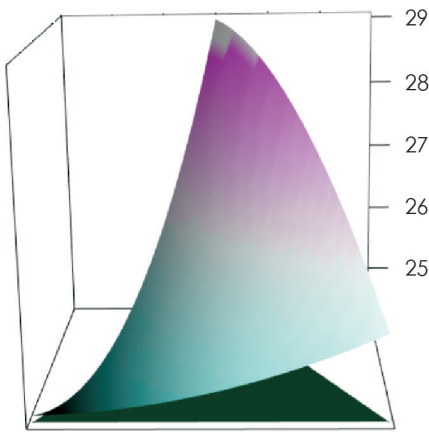


图 1 S^{3-1} 内旋转后的正交格点设计

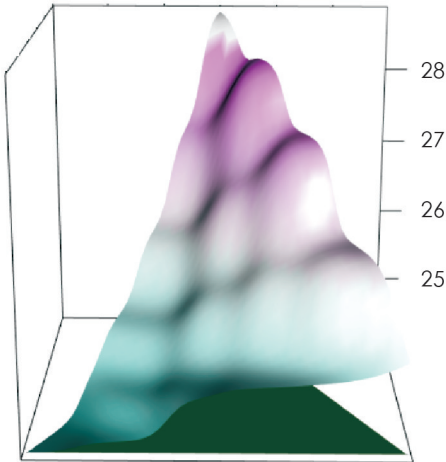
以模型

$$y = 23 + 3x_1 + 2x_2 + x_3 + x_1^2 + 5(x_2 - 0.75)^2 - 4(x_3 - 0.45)^2$$

为例, 在 S^{3-1} 上的曲面图如图 2(a) 所示, 使用正交格点建立非参数回归模型, 拟合回归曲面如图 2(b) 所示.



(a) 模型曲面图



(b) 正交格点设计建立的非参数回归模型曲面

图 2 曲面图

由图 2 可见, 使用正交格点设计建立的非参数回归模型具有很好的稳健性, 拟合效果还是不错的. 使用正交格点构造的混料试验域内的设计分布均匀, 利于非参数建模, 使用文献[20]的方法, 结合正交格子点集能有效检验设计的最优性, 关于这方面的性质有待进一步研究.

参考文献:

[1] ALTARAZI S A, ALLAF R M. Designing and Analyzing a Mixture Experiment to Optimize the Mixing Proportions of Polyvinyl Chloride Composites [J]. Journal of Applied Statistics, 2017, 44(8): 1441-1465.

[2] COETZER R, HAINES L M. The Construction of D-and I-Optimal Designs for Mixture Experiments with Linear Constraints on the Components [J]. Chemometrics and Intelligent Laboratory Systems, 2017, 171: 112-124.

[3] 赵行, 韩定强, 邹祥, 等. 基于 CFD-PBM 的 50 L 岩藻多糖发酵罐两相流模拟 [J]. 西南师范大学学报(自然科学版),

- 2021, 46(4): 114-121.
- [4] 汤绍虎, 孙敏, 周启贵, 等. 采用正交设计快速获得梨无菌外植体的研究 [J]. 西南师范大学学报(自然科学版), 2004, 29(2): 282-284.
- [5] CORNELL J A. A Primer on Experiments with Mixtures [M]. Hoboken: John Wiley & Sons, 2011.
- [6] GOMES C, CLAEYS-BRUNO M, SERGENT M. Space-Filling Designs for Mixtures [J]. Chemometrics and Intelligent Laboratory Systems, 2018, 174: 111-127.
- [7] DAS P, SINHA B K. Mixture Designs with Orthogonal Blocks [J]. Communications in Statistics-Theory and Methods, 2017, 46(16): 8155-8165.
- [8] NING J H, ZHOU Y D, FANG K T. Discrepancy for Uniform Design of Experiments with Mixtures [J]. Journal of Statistical Planning and Inference, 2011, 141(4): 1487-1496.
- [9] NING J H, CHENG P E. A Comparison Study of Nonparametric Imputation Methods [J]. Statistics and Computing, 2012, 22(1): 273-285.
- [10] NING J H, FANG K T, ZHOU Y D. Uniform Design for Experiments with Mixtures [J]. Communications in Statistics-Theory and Methods, 2011, 40(10): 1734-1742.
- [11] BORKOWSKI J J, PIEPEL G F. Uniform Designs for Highly Constrained Mixture Experiments [J]. Journal of Quality Technology, 2009, 41(1): 35-47.
- [12] LIU Y, LIU M Q. Construction of Uniform Designs for Mixture Experiments with Complex Constraints [J]. Communications in Statistics-Theory and Methods, 2016, 45(8): 2172-2180.
- [13] XIONG Z K, LIU L W, NING J H, et al. Sphere Packing Design for Experiments with Mixtures [J]. Statistics & Probability Letters, 2020, 164: 108807.
- [14] HE X. Interleaved Lattice-Based Minimax Distance Designs [J]. Biometrika, 2017, 104(3): 713-725.
- [15] HE X. Sliced Rotated Sphere Packing Designs [J]. Technometrics, 2019, 61(1): 66-76.
- [16] FANG K T, WANG Y. Number Theoretic Methods in Statistics [M]. London: Chapman and Hall, 1994.
- [17] 李光辉, 李俊鹏, 张崇岐. 单纯形体积公式的一个注记 [J]. 数学的实践与认识, 2020, 50(19): 221-227.
- [18] LI G H, ZHANG C Q. Random Search Algorithm for Optimal Mixture Experimental Design [J]. Communications in Statistics-Theory and Methods, 2018, 47(6): 1413-1422.
- [19] LI G H, ZHANG C Q. The Pseudo Component Transformation Design for Experiment with Mixture [J]. Statistics & Probability Letters, 2017, 131: 19-24.
- [20] 李光辉, 朱志彬, 燕飞, 等. 混料对称设计最优性的图检验法 [J]. 系统科学与数学, 2020, 40(2): 355-365.

责任编辑 张构