

均匀试验设计的理论、方法和应用 — 历史回顾

方开泰

(香港浸会大学 香港九龙塘)

摘要

本文回顾计算机仿真试验设计的主要两种方法：拉丁超立体抽样和均匀设计，在过去二十五年的发展，特别是均匀设计的发展，包括均匀设计的优良性研究、新的均匀性测度、均匀设计表的构造，以及均匀性在因子设计中的应用。

关键词：均匀设计，拉丁超立体抽样，因子设计，正交性，均匀性。

1 历史回顾

廿世纪七十年代，在系统工程、高科技发展的推动下，计算机仿真（仿真）试验（computer experiments）的需求十分强烈，迫切要求高质量的试验设计。于是计算机仿真试验设计（Design of computer experiments）在那时成为一个最有挑战的课题。在北美洲，三位学者（McKay, M.D., Beckman, R.J. and Conover, W.J. (1979))在“Technometrics”提出了“拉丁超立方体抽样”（Latin Hypercube Sampling）（简称 LHS）的方法，并立即得到广泛的应用，一批学者对其理论和方法作了系统地研究和发展，形成了一个独立的分枝。差不多在同一时间，在中国，我和王元院士提出了“均匀设计”（Uniform Design）（简称 UD）。文章最初在 1978 年发表在中国科学院数学研究所的内部通讯，后来中、英文稿分别发表在《应用数学学报》和《科学通报》。那时，中国正处于文化大革命刚结束，百废待兴的时代，学术上与世界几乎隔绝。有趣的是，LHS 和 UD 有异曲同工之处。表现于：

(A) 两种方法均将试验点均匀地散布于输入参数空间，故在文献中广泛使用术语“充满空间的设计”（space filling design）LHS 给出的试验点带有随机性，故称为抽样；而 UD 是通过均匀设计表来安排试验，不带有随机性。

(B) 两种方法的最初理论均来自“总均值模型”（Overall Mean Model）LHS 希望试验点对输出变量的总均值提供一个无偏估值，且方差较小，而 UD 是希望试验点能给出输出变量总均值离实际总均值的偏差最小。

(C) 两种设计均基于 U-型设计。

(D) 两种设计能应用于多种多样的模型，且对模型的变化有稳健性。

经过了廿多年的发展，两种不同思路的方法是分道扬镳，还是相互补充，相互融合呢？本文想作一些回顾和讨论。为此，我们需要介绍两种方法的思路、模型、方法和应用。

2 总均值模型

设输入变量 x_1, \dots, x_s 与输出变量有一个确定性的关系

$$y = f(x_1, \dots, x_s), \quad \mathbf{x} = (x_1, \dots, x_s) \in C^s. \quad (2.1)$$

这里假定试验区域为单位立方体 $C^s = [0,1]^s$ ，变量 y 在 C^s 上的总均值为

$$E(y) = \int_{C^s} f(x_1, \dots, x_s) dx_1, \dots, dx_s \quad (2.2)$$

若在 C^s 上取了 n 个试验点 $\mathbf{x}_1, \dots, \mathbf{x}_n$ ， y 在这 n 个试验点上的均值为

$$\bar{y}(D_n) = \frac{1}{n} \sum_{i=1}^n f(\mathbf{x}_i), \quad (2.3)$$

此处 $D_n = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ 代表这 n 个点的一个设计。

LSH 方法是用抽样的方法来选取 D_n 使相应的估计 $\bar{y}(D_n)$ 是无偏的，即

$$E(\bar{y}(D_n)) = E(y),$$

且方差 $\text{Var}(\bar{y}(D_n))$ 尽可能地小。McKay, Beckman and Conover (1979) 指出 LSH 比简单随机抽样要好，即前者所获得的总均值比后者有较小的方差。若设计点集 D_{random} 中的点 $\mathbf{x}_1, \dots, \mathbf{x}_n$ 独立同分布，遵从 C^s 上的均匀分布，相应样本均值 $\bar{y}(D_{random})$ 是 $E(y)$ 的无偏估计，其方差为 $\text{Var}(f(\mathbf{x}))/n$ ，其中 \mathbf{x} 在 C^s 上均匀分布。如果试验点同分布，但相互之间有相关性，得

$$\text{Var}(D_{LHS}) = \text{Var}(f(\mathbf{x}))/n + (n-1)\text{Cov}(f(\mathbf{x}_1), f(\mathbf{x}_2))/n, \quad (2.4)$$

右边第一项是随机抽样时样本均值的方差，故 $\text{Var}(D_{LHS}) < \text{Var}(f(\mathbf{x}))/n$ 当且仅当 $\text{Cov}(f(\mathbf{x}_1), f(\mathbf{x}_2))/n < 0$ 。LHS 就是设计试验点使 $\text{Cov}(f(\mathbf{x}_1), f(\mathbf{x}_2))/n < 0$ 。具体做法如下。为了便于理解，设 $n=8, s=2$ ，LHS 布点的步骤如下：

- (1) 将试验区域（不失一般性可设为单位正方形），每边均分 8 份，故单位正方形被分成 $64=8^2$ 个小正方形。
- (2) 随机地取 $(1,2,\dots,8)$ 的两个置换，例如 $(2,5,1,7,4,8,3,6)$ 和 $(5,8,3,6,1,4,7,8)$

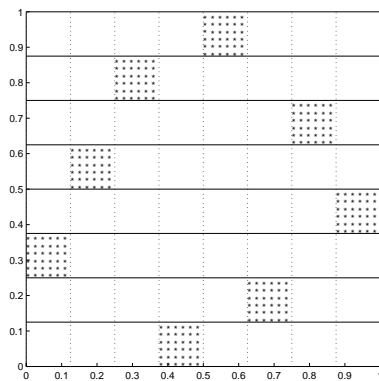
将它们排成一个矩阵，得

$$\begin{bmatrix} 2 & 5 \\ 5 & 8 \\ 1 & 3 \\ 7 & 6 \\ 4 & 1 \\ 8 & 4 \\ 3 & 7 \\ 6 & 2 \end{bmatrix}$$

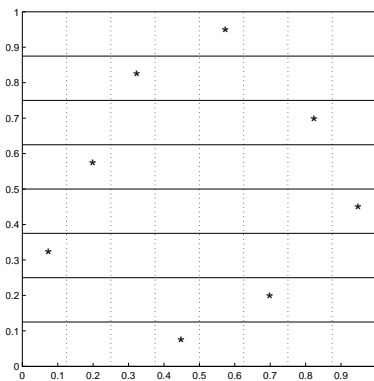
。由矩阵的每一行 (2,5)、(5,8)、...、(6,2) 决

定了 8 个小正方形，如图 1(a)，我们看到，每行、每列有一个小正方形。

- (3) 在已选中的 8 个小正方形上分别选点，使小正方形中的每个点被选中的机会相等，于是得图 1(b)的结果。这 8 个点就是一个 LHS。



(a)



(b)

图 1 LHS ($n=8, s=2$)

理论上可以证明，LHS 的样本均值的方差为

$$\text{Var}(\bar{y}(D_{LHS})) = \frac{1}{n} \text{Var}(f(x)) - \frac{c}{n} + o\left(\frac{1}{n}\right), \quad c > 0, \quad (2.5)$$

故 LHS 比随机抽样的方案好，因为其样本值有较小的方差。

均匀设计也是考虑了“总均值模型”，由数论方法中著名的 Koksma-Hlawka 不等式，可得，

$$|E(y) - \bar{y}(D_n)| \leq V(f)D(D_n), \quad (2.6)$$

式中 $V(f)$ 是函数 f 在 C^s 上的总变差，函数 f 平稳， $V(f)$ 则小，函数 f 波动大， $V(f)$ 则大； $D(D_n)$ 为点集在 D_n 在 C^s 上的偏差，偏差是度量点集均匀性的一种

测度。显然，我们希望 $D(D_n)$ 越小越好，即点在 C^s 上散布越均匀越好。仍以

$n=8, s=2$ 为例，UD 构造的步骤如下：

- [i] 将试验区域每边均分 8 份，共得 $8^2 = 64$ 个小正方形。

- [ii] 考虑在 64 个小正方形中选取 8 个小正方形,使得每行每列恰有一个小正方形。这种取法共有 $N = 8! \times 8! = 40320^2 \approx 1,625,702,400$ 种,对每一种取法,将试验放在小正方形之中心。然后比较这么多设计的偏差(即均匀性),使偏差达到最小的一个即为均匀设计。

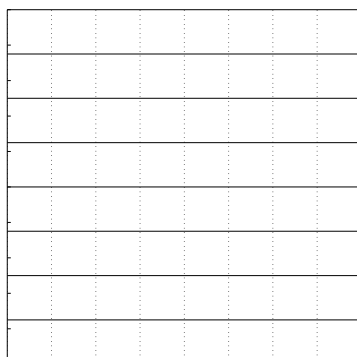


图 2 均匀设计 ($n=8, s=2$)

LHS 及 UD 的构造均需要 U-型设计的概念。

定义一: 有 n 次试验, s 个输入变量的任一个 U-型设计可表成一个 $n \times s$ 的矩阵,第 j ($j=1, \dots, s$) 列的元素由 $1, 2, \dots, q_j$ 组成,且这 q_j 个数出现的次数相同。该设计记为 $U(n; q_1 \times \dots \times q_s)$ 。若某些 q_j 相等时,该设计为 $U(n; q_1^{r_1} \times \dots \times q_m^{r_m})$, r_1, \dots, r_m 为正整数,且 $r_1 + \dots + r_m = s$ 。当 $q_1 = \dots = q_s = q$ 时,记为 $U(n; q^s)$ 。

在文献中,U-型设计又称为均衡设计 (Balanced design) 或格子点设计 (Lattice design)。

3 LHS 的发展

LHS 在西方十分流行,有关的文章很多,综合评述的文章不少,如 Sack, Welch, Mitchell and Wynn (1989), Betes, Buck, Riccomango and Wynn (1966), Koehler and Owen (1996)。熟悉随机抽样的人都知道,随机抽样的表现不稳定,故 LHS 提供的设计,有的很好,也有的很差。例如图 3 (a) 的设计很好,图 3 (b) 的设计很差。于是需要进一步来改进。其改进的思路列举如下:

- (1) **中心化的 LHS** (Centered LHS or Midpoint LHS) :研究发展 LHS 中的步骤 (3) 并不重要,不如免去在小立方体上抽样,而将试验点放小立方体的中心。以后讲到的 LHS 均为中心化 LHS。
- (2) **对称 LHS** (Symmetric Latin Hypercube) :人们发现,比较好的 LHS 有某种对称性。图 3 给出了两个 LHS ($n=6, s=2$)。显示 (a) 的设计比 (b) 均匀,它们若用矩阵来表示则为

$$D_{6-1} = \begin{bmatrix} 1 & 3 \\ 2 & 6 \\ 3 & 2 \\ 4 & 5 \\ 5 & 1 \\ 6 & 4 \end{bmatrix}, \text{ and } D_{6-2} = \begin{bmatrix} 1 & 5 \\ 2 & 4 \\ 3 & 3 \\ 4 & 2 \\ 5 & 1 \\ 6 & 6 \end{bmatrix}.$$

注意， D_{6-1} 的第一行与第六行之和，第二行与第五行之和，第三行与第四行之和均为 (7,7)，而 D_{6-2} 不具有这个性质。前者称为对称性的 LHS。有关中心化和对称的 LHS 的介绍和算法，可参见 Ye, Li and Sudjianto (2003)。

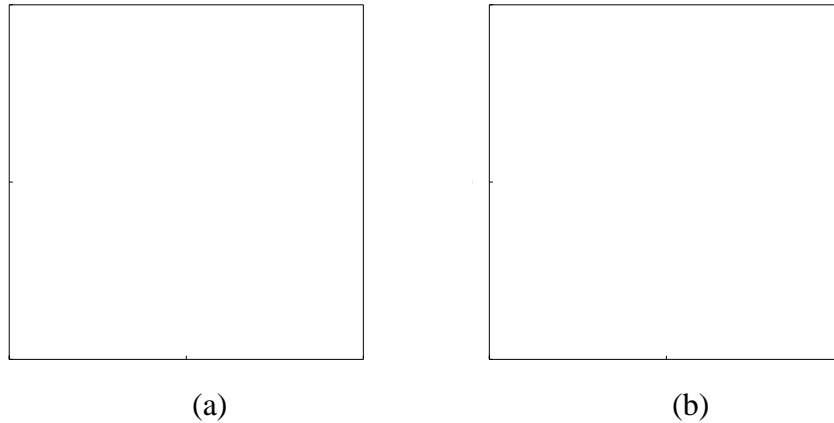


图 3 对称和非对称的 LHS

(3) **列正交的 LHS**：如果将设计表的水平安排成零对称，两个列的内积如为零，则这两个称为正交。若一个 LHS 的任两列都列正交，则有一系列好的性质。表 1 就是一个列正交的 LHS。详见 Ye (1998)。

表 1 列正交 LHS

No	1	2	3	4
1	1	-2	-4	3
2	2	1	-3	-4
3	3	-4	2	-1
4	4	3	1	2
5	0	0	0	0
6	-4	-3	-1	-2
7	-3	4	-2	1
8	-2	-1	3	4
9	-1	2	4	-3

(4) **将正交设计或均匀设计随机化**：这一想法是大大缩小 LHS 抽样的范围，即只在好的设计中抽样。所谓的 OA-based LHS 是将正交设计随机化，而 (t, m, s) 一网的随机化 (randomly permuted (t, m, s) -net) 是将均匀设计随机化。这一方

法可使均值 $\bar{y}(D_n)$ 的方差降至

$$\frac{1}{n} \text{var}(f(\mathbf{x})) - \frac{d}{n} + O\left(\frac{1}{n^{3/2}}\right), \quad d > 0.$$

- (5) **最优 LHS (Optimal LHS)**: 改进 LHS 的另一办法是外加一个准则, 用此准则来筛选 LHS, 求得在此准则下最优的设计。在文献中, 流行的准则有 IMSE (总均方差)、熵 (Entropy)、极小极大 (Minimax)、极大极小 (Maximin) 距离, 以及在此基础上定义的 ϕ_p 准则等。这一类的方法, 是选定一个准则 (如上述中的一个), 然后在所有 U-型设计中找一个使该准则达到最优者, 这就是我们追求的 Optimal Latin Hypercube Design。详细讨论可参见 Koehler and Owen (1969)。显然, 如果选均匀性作准则, 则相应的最优设计就是均匀设计。这时, LHS 和 UD 两种不同的思路达到异途同归。美国福特公司的工程经理 Agus Sudjianto 博士正是用这种方法将 LHS 和 UD 组合在一起放在他们的软件包中。而且, 他们发现, 如果均匀性准则取中心化偏差, 在上述诸准则中, 以均匀设计最好计算, 这也是他们近几年来使用均匀设计的原因之一。

4 均匀设计的发展

由于均匀性准则是个几何准则, 与统计推断似乎无关, 故均匀设计的理论发展比 LHS 的理论要艰难得多。在过去的十年中, 均匀设计的理论和方法有一个飞跃的发展, 表现在:

(1) 均匀设计的优良性

“总均值模型”虽然在发展 LHS 及 UD 起了关键性的作用, 但是一个试验的结果, 仅仅估计总均值是远远不够的。绝大部份试验的目的, 是通过试验数据来建立一个近似模型 (有解析表达、好计算)

$$y = \hat{f}(x_1, \dots, x_s) = \hat{f}(\mathbf{x}), \quad (4.1)$$

用它来逼近原模型 (2.1)。如果, 对预先给定的 $\delta > 0$, 能使

$$\left| f(x_1, \dots, x_s) - \hat{f}(x_1, \dots, x_s) \right| < \delta, \quad \mathbf{x} \in C^s.$$

该近似模型就十分理想。在文献中, 常常取

$$\hat{f}(\mathbf{x}) = \beta_1 f_1(\mathbf{x}) + \dots + \beta_m f_m(\mathbf{x}), \quad (4.2)$$

其中 f_1, \dots, f_m 为已知函数, β_1, \dots, β_m 是欲估的参数。如果能找到这样的

f_1, \dots, f_m , 什么样的设计是最好的呢? 谢民育和我 (Xie and Fang (2000)) 运用统计决策理论的思想和方法, 证明了均匀性测度是允许的, minimax 的而且在

一定的条件下是最优的。

当 (4.2) 不能很好近似真模型 (2.1) 时，稳健设计的思想是考虑

$$y = \beta_1 f_1(x) + \cdots + \beta_m f_m(x) + h(x) \quad (4.3)$$

这里 $h(x)$ 表示拟合不足的部份。稳健设计的思想是假定 $h(x)$ 来自一个函数类 \mathcal{H}

中，来找一个设计达到最好。众所周知，均方误 $MSE(D_n, h) = \int [f(x) - \hat{f}(x)]^2 dx$

是公认的用来衡量 $\hat{f}(x)$ 近似 $f(x)$ 质量的标准。由于 $h(x)$ 的任意性，在文献中

常用“平均均方误”和“最大均方误”作为准则，前者是 $MSE(D_n, h)$ 在 $h \in \mathcal{H}$ 中的

平均，后者是 $\max[MSE(D_n, h)]$ 。Hickernell (1999) 证明，均匀设计在这两个准

则下均是最优的。

Wein (1991) 也考虑模型 (4.3)，在试验有重复时，对近似模型 (4.1) 可作“lack of fit test”，并以 (4.3) 作备择假设。在一定的函数类 \mathcal{H} 下，他定义了两个准则，并证明均匀设计在两个准则下都是最优的。上述理论的建立使均匀设计建立在更加坚实的基础上，表现均匀设计对建立近似模型 $\hat{f}(x)$ 是最好的。

最近，Hickernell and Liu (2002) 指出“当用线性模型来拟合数据时，混杂现象不利于模型系数的估计以及确定回归中的项是否显著。如果模型已知，最优设计能给出回归系数的有效估计；稳健试验设计则留心模型的不确定性给回归系数的估计带来的不精确性。虽然很难要求一个设计既是有效的，又是稳健的。但我们将指出均匀设计可限制混杂现象所带来的影响，能提供一个合理的有效及稳健的估计。”

(2) 均匀性的度量

王元和我二十五年前在提出均匀设计时，用“星偏差” (star discrepancy) 作为均匀性的度量，因为星偏差广泛用于数论方法 (或伪蒙特卡罗方法) 之中，星偏差正是分布似合检验中的柯尔莫哥洛夫—斯来尔洛夫统计量。由于星偏差不便于计算，我们采用一种近似的算法。Hickernell (1998) 指出，星偏差及 L_p -偏差对坐标系的旋转缺乏不变性。于是他改用泛函分析的工具——希尔伯特再生核空间，来定义均匀性测度。

设 $F(x)$ 为点集 \mathcal{X} 上的均匀分布， $F_n(x)$ 是点集 $D_n = \{x_1, \dots, x_n\}$ 的经验分布函数。他们之间的距离可通过泛函分析中的模 $\|F(x) - F_n(x)\|$ 来定义。如果用再生核空间，设核函数为对称和非负定的 $K(x, t)$ ，则文献中各种偏差均可表为

$$D(D_n, K) = \iint_{\mathcal{X} \times \mathcal{X}} K(x, t) d[F(x) - F_n(x)] d[F(t) - F_n(t)].$$

通过取不同的核函数 $K(x,t)$, 及不同的试验点区域 \mathcal{S} , 可获得不同的偏差。当 \mathcal{S} 为单位立方体时 ,Hickernell (1998) 提出了好几种偏差 ,其中以中心化 L_2 - 偏差 (centered L_2 - discrepancy) , 可卷 L_2 - 偏差 (wrap-around L_2 - discrepancy) 具有好的性质 :

- (a) 对坐标系旋转有不变性 ;
- (b) 便于计算 ;
- (c) 与因子设计中的许多准则有紧切地联系。

当 \mathcal{S} 为 $\left\{ \frac{1}{2q}, \frac{3}{2q}, \dots, \frac{2q-1}{2q} \right\}^s$ 时 , 可获得散差偏差 (discrete discrepancy)。这三种

新偏差的提出 , 特别是离散偏差的提出 , 使均匀设计与因子设计的研究相互渗透 , 相互促进。详细讨论可见 Fang, Ge, Liu (2002b), Hickernell and Liu (2002)。

(3) 均匀设计表的构造

均匀设计表的构造 , 有如下的方法 :

- [1] 好格子点法 (good lattice point method) : 由王元院士和我首先使用 , 这是数论方法中最经典的方法 , 由 Korobov 于 1959 年提出 , 至今仍是一种实用而可行的方法。后来 , 许多学者对此法有很多改进。
- [2] 拉丁方法 : 由 Fang, Pan, Shiu (1999) 提出。
- [3] 正交表扩充法 : 由方燕 (1995) 提出。
- [4] 折迭法 (collapsing method) : 是将两个均匀设计表 , 用 Kronecker 乘积的方法折迭在一起。设 U 和 V 分别为 $n \times s$ 及 $m \times t$ 的均匀设计表 , 令

$$D_{U,V} = (I_m \otimes U : V \otimes I_n)$$

式中 I_m 为 $m \times 1$ 且元素均为 1 的向量 , \otimes 表示 Kronecker 乘法。方开泰和覃

红 (Fang and Qin (2003)) 证明用这种方法构造的设计有较好的均匀性。例如

$$U = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \quad V = \begin{pmatrix} 1 & 3 \\ 2 & 2 \\ 3 & 1 \end{pmatrix}$$

则

$$D_{U,V} = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 2 & 1 & 1 & 3 \\ 1 & 2 & 2 & 2 \\ 2 & 1 & 2 & 2 \\ 1 & 2 & 3 & 1 \\ 2 & 1 & 3 & 1 \end{bmatrix}.$$

[5] 切割法 (cutting method) : 用好格子法中的指数生成间量法来生成均匀设计表 $U_p(p^s)$ 的计算量很小, 当 p 为素数时, 相应的表 $U_p(p^s)$ 有很好的均匀性。马长兴和我 (Ma and Fang (2003)) 提出了所谓的“均割法”, 从 p 个点中取出 $n(\ll p)$ 个点, 则可构造出一个 $U_n(n^s)$ 的表。从一个 $U_p(p^s)$ 出发, 可获得多个不同 n 的 $U(n^s)$ 。

[6] RBIBD 法: 可分解的平衡不完全区组设计 (resolvable balanced incomplete block designs) 在过去的几十年中有系统的理论和丰富的设计。我和我的合作者们葛根年、刘民千、殷嘉兴、覃红以及陆璇和他的合作者建立了 RBIBD 和均匀设计之间的联系, 通过这个联系可将许多已获得的 RBIBD 转化为均匀设计表 $U_n(q^n)$ 或一些混合水平的表。这些表是在离散偏差下达到最优。

有关的研究可类似地用于构造超饱和设计, 所谓超饱和设计是指试验的数目少于欲估的参数的数目 (主效应、交互效应、误差方差等)。有关的讨论很多, 可参见: Lu and Meng (2000), Lu and Sun (2001), Fang, Ge and Liu (2002a; b), Fang, Ge and Liu (2003), Fang, Ge, Liu and Qin (2002, 2003), Fang, Lin and Liu (2003) 等。

[7] 翻转法: 翻转法 (foldover) 常用于二水平的试验, 例如一个 $L_4(2^3)$, 通过翻转可以变为 $L_8(2^3)$, 使得得后者的分辨率 (resolution) 可以提高一个等级。

均匀设计也可通过翻转法来扩大, 以及作序贯均匀设计。Ye (1998) 给出了一种方法, 但更一般地情形尚有待进一步研究。

[8] 数值优化法: 求均匀设计是一个典型的优化问题。我和 P. Winker (Winker and Fang (1998)) 用门限接受法, 以星偏差为均匀性测度求得一批均匀设计表。后来, 我们 (Fang, Ma and Winker (2002)) 用中心化偏差 (Fang and Ma (2001)) 用可卷偏差, 也获得了许多均匀设计表。最近 Fang, Lu and Winker (2003) 对二水平及三水平的均匀表其中心化 L_2 偏差的下界给了一个估计, 从而加速了计算的过程。

(4) 均匀性在因子设计中的应用

均匀性表面上是一个几何的准则, 但它在统计推断中仍有丰富的内涵。在过去的八年中, 这些丰富的内涵被逐一地揭示出来。下面将介绍其中的一部份:

[1] 均匀性与字长: 在正规的因子设计中, 字长用来刻划一个因子设计有关主效应和交互效用的混杂情形。若一个 s 因素二水平设计 D , 其试验数为 $n = 2^{s-p}$ 的正规试验, 它的 s 个字长组合一个向量 $W(D) = (A_1(D), A_2(D), \dots, A_s(D))$ 。

两个重要的准则“最大分辨率”和“最小低阶混杂”都是 $W(D)$ 的函数。它们的

定义可参见方开泰、马长兴 (2001) § 2.7。这两个准则似乎与均匀性毫不相关，有趣的是我们 (Fang and Mukerjee (2000)) 对正规的 2^{s-p} 的因子设计 D ，给出了它的中心化 L_2 - 偏差和字长间的解析关系

$$[CD(D)]^2 = \left(\frac{13}{12}\right)^s - 2\left(\frac{35}{32}\right)^s + \left(\frac{8}{9}\right)^s \left[1 + \sum_{i=1}^s \frac{A_i(D)}{9^i}\right].$$

由这个关系不难看出均匀性和最小低阶个混杂准则很相近。随后我们 (Ma and Fang (2001)) 将上述关系推广至非正规因子设计和可卷 L_2 - 偏差。在因子设计的研究中，长期以来人们停留于正规设计，因为非正规因子设计那时没有字长的定义。而均匀性既可用于正规设计，也可用于非正规设计，显示了它的优势。直至 2001 年，Ma and Fang (2001), Xu and Wu (2001) 同时将字长由正规因子设计推广至非正规因子设计，但字长的计算比均匀性 (如中心化 L_2 - 偏差) 的计算复杂。均匀性的上述简单性和普遍性显示了它的优点。但均匀性的数值大小有时不易直接给出统计推断的意义。如正交表 $L_9(3^4)$ 和均匀正交表 $UL_9(3^4)$ 的前三列的平方中心化 L_2 - 偏差分别为 0.050059 和 0.0493645，它们之间均匀性的差别说明了什么？如果仅仅是看它们的数值，只能讲均匀正交表的均匀性比通常的 $L_9(3^4)$ 好，不能得到更多的性质。我和覃红 (Fang and Qin (2002)) 提出了一个向量 $(MI_1(D), \dots, MI_s(D))$ ，称为“均匀性指针” (uniformity pattern)。该向量具有字长型 $W(D)$ 的优点，又有均匀性的通用性。用这个工具，可以方便地比较各种设计。

[2] 均匀性与正交性：常见的大多数正交表是强度为 2 的正交数组，这些表要求一维和二维的投影均匀性，而均匀设计表是要求一维和 s 维的投影均匀性。其间必有许多联系，也有许多差异。

(a) 我和 Winker 在 1998 年发现，常用的许多正交表是在中心化 L_2 - 偏差下的均匀设计，可以通过优化计算来获得，并提出一个猜想：“一切正交表均是一定均匀性测度下的均匀设计”。后来，我们发现这个猜想是不对的，但在一些特殊情形下是对的，Ma, Fang and Lin (2003) 给予了详细的讨论。

(b) Fang, Ma and Mukerjee (2001) 对一个部分因子 q^{s-p} 设计 D ，定义了一个正交性的向量 $(B_1(D), \dots, B_s(D))$ ，其中 $B_1(D)$ 表示设计 D 与强度为 1 的正交性偏差， $B_2(D)$ 表示设计 D 与强度为 2 的正交性偏差，…。我们在文中给出了中心化 L_2 - 偏差，可卷 L_2 - 偏差表为 $(B_1(D), \dots, B_s(D))$ 的解析函

数。

(c) 均匀性的正交表：如果一个设计既是均匀又是正交设计 D ，则称为均匀正交设计。我和马长兴在《正交与均匀试验设计》一书中，对 $L_9(3^4)$ 和

$UL_9(3^4)$ 进行了比较，指出， $UL_9(3^4)$ 能减少混杂。

[3] 均匀性与正交表的同构性：用 $d(n, q, s)$ 表示安排 s 个 q 水平的因素并用 n 次试验的设计。两个 $d(n, q, s)$ 的设计 d_1 和 d_2 称为同构，如果通过试验次序的交换，因素排列的置换，以及一个或多个因素水平的置换，能将一个设计变为另一个设计，则称为这两个设计同构。要识别两个 $d(n, q, s)$ 的设计是否同构，要比较 $n!(q!)^s s!$ 对 $d(n, q, s)$ 的设计，故是一个计算很复杂的问题，在文献中快速的算法不多。Ma, Fang and Lin (2001) 利用设计点的均匀性和投影均匀性给出了一个算法，可以有效地鉴别不同构的设计。将这一思想发挥，Fang and Ge (2003) 给出了一个有效算法来识别不等价的哈达马矩阵，并首次发现，36 阶的哈达马矩阵至少有 392 个不等价。

[4] 均匀性用于构造超饱和设计：若一个设计的试验点数少于欲估计的参数（主效应、交互效应等）个数，该设计称为超饱和设计。超饱和设计常用来筛选因素。为了构造超饱和设计，文献中提出了许多准则，如 $E(s^2)$ ， $Ave(\chi^2)$ ， $Ave(f^2)$ ， $E(f_{NOD})$ 等，Fang, Lin and Liu (2003) 指出，在二水平时 $E(f_{NOD})$ 与 $E(s^2)$ 等价，在三水平时 $E(f_{NOD})$ 与 $Ave(\chi^2)$ 等价。Fang, Ge, Lin (2002a) 指出，在二水平时离散偏差与 $E(s^2)$ 等价，在三水平时，离散偏差与 $Ave(\chi^2)$ 等价，也与 $E(d^2)$ （参见 Lu and Sun (2001)）等价。而离散偏差可以用于多水平和混合水平的情形，故在上述文献中，利用可分解的平衡不完全区组设计与均匀设计的关系，可以构造出一大批超饱和设计。

5 建模方法 (Modelling)

在计算机仿真试验中，欲通过试验数据寻找一个近似于真模型 (2.1) 的近似模型

$$y = \hat{f}(x_1, \dots, x_s) \quad (5.1)$$

在文献中对 $\hat{f}(x_1, \dots, x_s)$ 有好几种术语，如“Approximate model”，“metamodel”，“model of the model”。方开泰、马长兴 (2001) 在书中详细讨论了用多项式回归模型、人工神经网络、多元样条函数等方法来建模。在实践中，线性模型

$$y = \beta_1 f_1(\mathbf{x}) + \dots + \beta_m f_m(\mathbf{x}) \quad (5.2)$$

是大多数学者采用的近似模型，式中 f_1, \dots, f_m 已知， β_1, \dots, β_m 未知。由于可能组成的 f_1, \dots, f_m 数量很多，回归分析中各种筛选变量的技术十分需要，除了众所周知的“向前法”、“后退法”、“逐步回归法”及“最优子集法”外，由范钊青和李润泽 (Fan and Li (2001)) 提出的一种新的惩罚函数 (penalized function) 方法具有很好的性质。李润泽 (Li (2002)) 利用这一方法进行了两个均匀设计案例的研究。

Kriging 方法为许多学者所推崇，它是在模型 (5.2) 的基础上，加了一项平稳高斯过程 $z(x)$

$$y = \sum_{j=1}^m \beta_j f_j(x_1, \dots, x_s) + z(x_1, \dots, x_s) \quad (5.3)$$

使得在试验点上真值与估计值相等，故这一方法本质上是一个插值的方法。高斯过程 $z(x)$ 中有一些未知参数 $\theta_1, \dots, \theta_t, \sigma^2$ ，我们要同时估计 $\theta_1, \dots, \theta_t, \sigma^2, \beta_1, \dots, \beta_m$ 。

在软件包 MATLAB 中已有 Kriging 的软件 (MATLAB BOX)。Li and Sudjianto (2003) 给出了利用 Kriging 方法建模的均匀设计案例研究，这是美国特汽车公司的一个实验。

Morris, Mitchell and Ylvisaker (1993) 在 Kriging 模型的框架下提出了一个贝叶斯算法，该方法要求 y 对每个输入变量 x_j 的方向导数，故许多案例是不可行的。

考虑到计算机试验不存在随机误差，许多统计方法直接应用有困难，为了建模，建立了一套“Sensitivity Analysis”，其思想是将原模型分解

$$y = f(x_1, \dots, x_s) = f_0 + \sum_{i=1}^s f_i(x_i) + \sum_{1 \leq i < j} f_{ij}(x_i, x_j) + \dots + f_{12\dots s}(x_1, \dots, x_s) \quad (5.4)$$

希望这种分解可有助于将函数的变化能类似于方差分析中的平方和分解。不失一般性，设试验区域为单位立方体。为此，Sobol (2001) 引进了 ANOVA 分解，他要求函数已中心化，即

$$\int_0^1 f_{1\dots t_k}(x_1, \dots, x_{t_k}) dx_{t_j} = 0, \quad j = 1, \dots, k. \quad (5.5)$$

这时，(5.4) 中的诸项可以方便地求得，如

$$\begin{aligned} f_0 &= \int_{[0,1]^s} f(x_1, \dots, x_s) dx_1 \dots dx_s, \\ f_i(x_i) &= \int_{[0,1]^{s-1}} f(x_1, \dots, x_s) \prod_{k \neq i} dx_k - f_0, \\ f_{ij}(x_i, x_j) &= \int_{[0,1]^{s-2}} f(x_1, \dots, x_s) \prod_{k \neq i, j} dx_k - f_0 - f_i(x_i) - f_j(x_j). \end{aligned}$$

这里 f_0 相当于总平均， $f_i(x_i)$ 相当于变量 x_i 的主效应， $f_{ij}(x_i, x_j)$ 相当于变量 x_i

和 x_j 的交互效应，…。然后，他又引进“方差”的概念

$$D = \int_{[0,1]^s} f^2(x_1, \dots, x_s) dx_1 \cdots dx_s - f_0^2,$$

并证明方差有平方和分解

$$D = \sum_{k=1}^s \sum_{t_1 < \dots < t_k} D_{t_1 \cdots t_k}$$

其中

$$D_{t_1 \cdots t_k} = \int f_{t_1 \cdots t_k}^2 dx_{t_1} \cdots dx_{t_k},$$

考核变量的主效应，用

$$S_k = \frac{D_k}{D}, \quad k = 1, \dots, s$$

的大小，考核 x_i 和 x_j 的交互作用，用

$$S_{ij} = \frac{D_{ij}}{D}.$$

一般地，考核 x_{t_1}, \dots, x_{t_k} 的交互作用，用

$$S_{t_1 \cdots t_k} = \frac{D_{t_1 \cdots t_k}}{D}.$$

上述的 $S_k, S_{ij}, S_{t_1 \cdots t_k}$ 称为敏感指针 (sensitivity indices)，用它来分析因素主效应

和交互作用的重要程度，并删去不重要的主效应和交互效应，从而用剩余的项组成一个近似模型。在上述的分析中，要计算许多高维积分。在文献中，发展了一套算法称为 FAST (Fourier amplitude sensitivity test)。该方法，是通过一个 Fourier 变换，将高维积分变为一维积分，详细讨论可参见 Saltelli and Bolado (1998)。

由于建构的复杂性，尽管学者们发展了许多方法，但仍然远远不能满足于实际的需要，这是一个很有挑战性的研究方向。

6. 应用

均匀设计的应用日益广泛，成功的案例与日俱增，读者不难从各种文献库中发现这些案例。近年来，均匀设计走向国际，有关均匀设计和均匀性的文章在国际刊物上已发表了几十篇，包括国际上顶尖的一些杂志，如“Biometrika”、“Technometrics”、“Mathematics Computation”，SIAM 的刊物等。特别令人鼓舞的是，美国福特汽车公司正将均匀设计作为他们推行 6 Sigma，以及研制新型引擎的常规方法，福特汽车公司的工程经理 A. Sudjianto，邀请我去夏和今夏去福特汽车公司讲学，并合作课题，他在邀请信中讲道：「In the past few years, we have tremendous in using Uniform Design for computer experiments. The technique has

become a critical enabler for us to execute “Design for Six Sigma” to support new product development, in particular, automotive engine design. Today, computer experiments using uniform design have become standard practices at Ford Motor Company to support early stage of product design before hardware is available. We would like to share with you our successful real world industrial experiences in applying the methodology that you developed. Additionally, your visit will be very valuable for us to gain more insight about the methodology as well as to learn the latest development in the area. 福特汽车公司是美国最大的汽车公司，均匀设计在福特汽车公司的运用成功，将会产生巨大的回响。

参考文献

- [1] Bates, R. A., Buck, R. J., Riccomagno, E. and Wynn, H. P. (1996). Experimental design and observation for large systems. *J. R. Statist. Soc. B.*, **58**, 77--94.
- [2] Box, G. E. P. and Draper, N. R. (1959). A basis for the selection of a response surface design, *J. Amer. Statist. Assoc.*, **54**, 622-654.
- [3] Fan, J. and Li, R. (2001). Variable Selection via Nonconcave Penalized Likelihood and Its Oracle Properties, *J. Amer. Statist. Assoc.*, **96**, 1348-1360.
- [4] Fang, K. T. (1980). The uniform design: application of number-theoretic methods in experimental design, *Acta Math. Appl. Sinica*, **3**, 363-372.
- [5] Fang, K. T. (2004). Uniform designs, *Encyclopedia of Statistics*, **2nd Edition**, Wiley, New York, to appear.
- [6] Fang, K. T. and Ge, G. N. (2002). A sensitive algorithm for detecting the inequivalence of Hadamard matrices, *Math.Computation*, to appear.
- [7] Fang, K. T., Ge, G. N. and Liu, M. Q. (2002a). Uniform supersaturated design and its construction, *Science in China, Ser. A*, **45**, 1080-1088.
- [8] Fang, K. T., Ge, G. N. and Liu, M. Q. (2002b). Construction on $E(f_{NOD})$ -optimal supersaturated designs via room squares, in *Calcutta Statistical Association Bulletin* Vol 52, A. Chaudhuri and M. Ghosh Eds., 71-84.
- [9] Fang, K. T., Ge, G. N. and Liu, M. Q. (2003). Construction of optimal supersaturated designs by the packing method, *Science in China*, to appear.
- [10] Fang, K. T., Ge, G. N., Liu, M. Q. and Qin, H. (2002). Construction on minimum generalized aberration designs, *Metrika*, **57**, 37-50.
- [11] Fang, K. T., Ge, G. N., Liu, M. Q. and Qin, H. (2003). Construction of uniform designs via super-simple resolvable t -designs, *Utilitas Mathematica*, to appear.
- [12] Fang, K. T. and Lin, D. K. J. (2003). Uniform designs and their application in industry, in *Handbook on Statistics 22: Statistics in Industry*, Eds by Khattree, R. and Rao, C. R., Elsevier, North-Holland, 131-170.
- [13] Fang, K. T., Lin, D. K. J. and Liu, M. Q. (2003). Optimal mixed-level supersaturated design, *Metrika*, to appear.

- [14] Fang, K. T., Lin, D. K. J. and Qin, H. (2003). A note on optimal foldover design, *Statist. & Prob. Letters*, **62**, 245-250.
- [15] Fang, K. T., Lin, D. K. J., Winker, P. and Zhang, Y. (2000). Uniform design: Theory and Applications, *Technometrics*, **42**, 237-248.
- [16] Fang, K. T., Lu, X, and Winker, P. (2003). Lower bounds for centered and wrap-around L_2 -discrepancies and construction of uniform designs by threshold accepting, *J. Complexity*, in press.
- [17] Fang, K. T., Lu, X, Tang, Y. and Yin, J. X. (2003). Constructions of uniform designs by using resolvable packings and coverings, *Discrete Mathematics*, to appear.
- [18] Fang, K. T. and Ma, C. X. (2001). Relationships Between Uniformity, Aberration and Correlation in Regular Fractions 3^{s-1} , in *Monte Carlo and Quasi-Monte Carlo Methods 2000*, Eds by Fang, K. T., Hickernell, F. J. and Niederreiter, H., Springer, 213-231.
- [19] Fang, K. T. and Ma, C. X. (2001). Wrap-around L_2 -discrepancy of random sampling, Latin hypercube and uniform designs, *J. Complexity*, 608-624.
- [20] Fang, K. T., Ma, C. X. and Mukerjee, R. (2001). Uniformity in Fractional Factorials, in *Monte Carlo and Quasi-Monte Carlo Methods 2000*, Eds by Fang, K. T., Hickernell, F. J. and Niederreiter, H., Springer, 232-241.
- [21] Fang, K. T., Ma, C. X. and Winker, P. (2002). Centered L_2 -discrepancy of random sampling and Latin hypercube design, and construction of uniform design, *Math. Computation*, **71**, 275-296.
- [22] Fang, K. T. and Mukerjee, R. (2000). A connection between uniformity and aberration in regular fractions of two-level factorials, *Biometrika*, **87**, 193-198.
- [23] Fang, K. T. and Qin, H. (2002). A note on construction of nearly uniform designs with large number of runs, *Statist. & Prob. Letters*, **61**, 215-224.
- [24] Fang, K. T. and Qin, H. (2002). Uniformity Pattern And Related Criteria For Two-level Factorials, to be submitted.
- [25] Fang, K. T. and Wang, Y. (1994). *Number-theoretic Methods in Statistics*, Chapman and Hall, London.
- [26] Fang, K. T. and Winker, P. (1998). Uniformity and Orthogonality, Technical Report, MATH-175, Hong Kong Baptist University.
- [27] Hickernell, F. J. (1998a). A generalized discrepancy and quadrature error bound, *Math Comp.*, **67**, 299-322.
- [28] Hickernell, F. J. (1998b). Lattice rules: how well do they measure up? in *Random and Quasi-Random Point Sets*, Eds P. Hellekalek and G. Larcher, Springer-Verlag, 106-166.
- [29] Hickernell, F. J. (1999). Goodness-of-Fit Statistics, Discrepancies and Robust Designs, *Statist. Probab. Lett.*, **44**, 73--78.

- [30] Huber, J. (1975). Robustness and design, in *A Survey of Statistical Design and Linear Models*, Eds by J. N. Srivastava, North-Holland, Amsterdam, 287-303.
- [31] Johnson, M. E., L. M. Moore, and D. Ylvisaker (1990). Minimax and maxmin distance designs, *J. Statist. Plan. and Infer.*, **26**, 131-148.
- [32] Koehler, J. R. and A. B. Owen (1996). Computer experiments, in *Handbook of Statistics*, Vol. 13, Eds by Ghosh, S. and Rao, C. R., Elsevier Science B. V., Amsterdam, 261-308.
- [33] Li, R. (2002). Model selection for analysis of uniform design and computer experiment, *International Journal of reliability, Quality and Safety Engineering*, **9**, 305-315.
- [34] Li, R. and Sudjianto, A. (2003). Analysis of computer experiments using penalized likelihood Gaussian Kriging model, to be submitted.
- [35] Liang, Y. Z., Fang, K. T. and Xu, Q. S. (2001). Uniform design and its applications in chemistry and chemical engineering, *Chemometrics and Intelligent laboratory Systems*, **58**, 43-57.
- [36] Lu, X. and Meng, Y. (2000). A new method in the construction of two-level supersaturated designs, *Journal of Statistical Planning and Inference*, **86**, 229-238.
- [37] Lu, X., Hu, W. and Zheng, Y. (2003). A systematical procedure in the construction of multi-level supersaturated designs, *Journal of Statistical Planning and Inference*, to appear.
- [38] Lu, X. and Sun, Y. (2001). Supersaturated design with more than two levels, *Chinese Annals of Mathematics (Series B)*, **22**, 183-194.
- [39] Ma, C. X. and Fang, K. T. (2001). A note on generalized aberration in factorial designs, *Metrika*, **53**, 85-93.
- [40] Ma, C. X. and Fang, K. T. (2003). A new approach to construction of nearly uniform designs, *International Journal of Materials and Product Technology*, to appear.
- [41] Ma, C. X., Fang, K.T. and Lin, D.K.J. (2001). On isomorphism of factorial designs, *J. Complexity*, **17**, 86--97.
- [42] Ma, C. X., Fang, K. T. and Lin, D. K. J. (2003). A note on uniformity and orthogonality, *J. Statist. Plan. Infer.*, **113**, 323-334.
- [43] Morris, M. D., Mitchell, T. J. and Ylvisaker, D. (1993). Bayesian design and analysis of computer experiments: use of derivatives in surface prediction, *Technometrics*, **35**, 243-255.
- [44] McKay, M. D., Beckman, R. J. and Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of out put from a computer code, *Technometrics*, **21**, 239-245.
- [45] Niederreiter, H. (1992). *Random Number Generation and Quasi-Monte Carlo Methods*. SIAM CBMS-NSF Regional Conference Series in Applied Mathematics,

Philadelphia.

- [46] Owen, A. B. (1992). Orthogonal arrays for computer experiments, integration and visualization, *Statist. Sinica*, **2**, 439--452.
- [47] Owen, A. B. (1994). Lattice sampling revisited: Monte Carlo variance of means over randomized orthogonal arrays, *Ann. Statist.*, **22**, 930-945.
- [48] Owen, A. B. (1995). Randomly permuted (t,m,s) -nets and (t,s) -sequences, In *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing*, Niederreiter, H. and Shiue, P. J. D., Eds, Springer, New York, 299-317.
- [49] Sacks, J., Welch, W. J., Mitchell, T. J. and Wynn, H. P. (1989). Design and analysis of computer experiments (with discussion), *Statist. Sinica*, **4**, 409-435.
- [50] Saltell, A. and Bolado, R. (1998). An alternative way to compute Fourier amplitude sensitivity test, *Comp. Statist. & Data Analysis*, **26**, 445-460.
- [51] Shewry, M. C. and Wynn, H. P. (1987). Maximum entropy sampling, *J. Appl. Statist.*, **14**, 165-170.
- [52] Sobol, I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, *Math. Comp. Simulation*, **55**, 271-280.
- [53] Tang, B. (1993) Orthogonal array-based Latin hypercubes, *J. Amer. Statist. Asso.*, **88**, 1392-1397.
- [54] Wang, Y. and Fang, K. T. (1981). A note on uniform distribution and experimental design, *KeXue TongBao*, **26**, 485-489.
- [55] Wang, Y. and Fang, K. T. (1990). Number theoretic method in applied statistics (II) . *Chinese Annals of Math. Ser. B*, **11**, 384-394.
- [56] Wang, Y. and Fang, K. T. (1990) . Number theoretic method in applied statistics, *Chinese Annals of Math. Ser. B*, **11**, 41-55.
- [57] Wiens, D. P. (1991) . Designs for approximately linear regression: two optimality properties of uniform designs, *Statist. Probab. Lett.*, **12**, 217-221.
- [58] Winker, P. and Fang, K. T. (1998) . Optimal U-type design, in *Monte Carlo and Quasi-Monte Carlo Methods 1996*, eds. by Niederreiter, H., Zinterhof, P. and Hellekalek, P., Springer, 436-448.
- [59] Xie, M. Y. and Fang, K. T. (2000) . Admissibility and minimaxity of the uniform design in nonparametric regression model, *J. Statist. Plan. Inference*, **83**, 101-111.
- [60] Ye, Q. (1998) . Orthogonal column Latin hypercubes and their application in computer experiments, *J. Ammer. Statist. Assoc.*, **93**, 1430-1439.
- [61] Ye, Q. and Li, W. and Sudjianto, A. (2003) . Algorithmic construction of optimal symmetric Latin hypercube designs, *J. Statist. Plan. Inference*, **90**, 145-159.
- [62] Yue, R. X. and Hickernell, F. J. (1999) . Robust designs for fitting linear models with misspecification, *Statist. Sinica*, **9**, 1053-1069.
- [63] Xu, Q. S., Liang, Y. Z. and Fang, K. T. (2000) . The effects of different

experimental designs on parameter estimation in the kinetics of a reversible chemical reaction, *Chemometrics and Intelligent Laboratory Systems*, **52**, 155-166.