112-117

1999 年 9 月 重庆大学学报 (自然科学版) Vol.22
第 22 卷第 5 期 Journal of Chongqing University(Natural Science Edition) Sep.1999

文章编号:1000-582x(1999)05-0112-06

正交设计与均匀设计的初步比较*

O224

夏之宁[1], 谌其亭[2], 穆小静[1], 李志良[1]
(1. 重庆大学 环境与化学化工学院,重庆 400044; 2. 长沙电力学院 化学系,长沙 410004)

摘 要:系统研究了试验设计与优化及其在化学中的应用。对广用的正交设计与新颖的均匀设计进行了初步比较,通过有关优化效率及试验次数的比较,指出两者各自的优缺点。作为一项方法或技术、试验设计与优化及其推广应用是项重要工作,宜取长补短,相互促进。

关键词:试验设计与优化;正交设计;均匀设计;均衡分散;齐整可比
中图分类号:O6-3 **文献标识码**:A

试验设计与优化是化学计量学的重要分支领域之一[1-4],也是科学研究中人们乐于采用的有用方法或工具。正交试验设计(OAD)是应用甚广、历时较长的试验设计方法[5-6]。均匀试验设计(HUD)则是较新开发、应用日益广泛的试验设计方法[7-8],它们均引起人们广泛关注。新近有人倡导采用"均匀设计",且声势不小。如文献[8]中专设章节(3,4)标题为"均匀设计与正交设计的比较",结语是如果采用偏差作均匀性的度量,均匀设计明显优于正交设计,并至少可节省 60% 以上的试验,文献[9]亦声言"均匀设计是一种优于正交设计的新试验设计方法"。这些提法有失偏颇。作者现道出有关观点与看法,供有关专家学者交流和广大使用人员参考。

1 主要论题

讨论的主题限于多因素实验条件最优化的范畴,而不指解决其它性质的课题。主要包括多因素实验设计和多变量计算机设计。开展优化研究,旨在提高质量,增加产量,降低成本,节省时间,改善操作条件,防止环境污染,获得优良效益。通常试验初期,由于开发潜力较大,往往改进幅度较大,效益较高;而当多次试验以后,已达到相当水平,存在潜力已相对降低,因而进展较慢,但试验单次成本不减少,故试验后期的效益常会下降。试验的基础既已建立,多数情况包括研究项目应追求总体效果或者全局性的最优化。但很多试验或因限于时间地点,如寿命试验或田间试验;或限于原材料数量,如价格较昂或不易得到;或因结果

* **收稿日期**:1998-10-05
基金项目:国家自然科学基金项目(29542010、29775035)、教育部优秀年轻教师基金与机械部资助课题
作者简介:夏之宁(1961-),男,重庆大学副教授,博士。主要从事化学计量学与毛细管电泳的研究和教学工作。

已基本满足使用者要求，只要求阶段效果好。因计算机具强计算功能，除非课题计算量特大以致超出计算能力，一般总可解出多元函数的极值（或非线性规划最优解）。本文论题拟分三部分。现作初步简要的综合比较，以后另文将补充列举较多情形和较多实例的计算机验证，并讨论分散性度量等问题，再以化学实验结果和其它材料累积作试验比较。

试验好的设计，已有一些论著涉及[1~4]。简言之，便是在较少的实验次数及批数下达到相对较好的结果。实验次数与批数的多少有时一致，而有时矛盾。处理最优化问题的指导性原则是：由稀至密，分批进行且作观察稽考，有苗头处着重加密，过稀之处适当加密，从而提高效率。

2 齐整可比性

均匀设计只具有充分均匀性，而正交设计还具有齐整可比性，故此处专门讨论。设求多元（$m>1$）连续函数 $f(x)$ 的绝对或全局极优值（极大或极小，此处设为极大值），这要求通过试验或者计算，把事实不知道的多维空间中的某点 $X^*=(X_1^*\ X_2^*\ \wedge\ X_m^*)$ 寻找出来，使得 $f(x)$ 达到极大（优）值：　　$f(x^*)=\max\limits_x \cdot f(x)$

为行文方便，假设全局最优所对应的多维 x^* 是唯一的。其中有一个选哪些多维点作试验或计算的问题，具体地说，对于第 j 维分量 x_j 要确定一个考察范围$[a_j, b_j]$，$x_j\in[a_j, b_j]$，$j=1, 2, \cdots, m$ 范围选小了以致漏掉 x_j^*，则此批实验无法找到最优点，解决最优化问题，须在下批实验中扩大$[a_j, b_j]$进行；范围扩大了，即使包括了 x_j，但还有一如何找出 x_j 或其近似值问题。区域过大势必浪费试验。衡量一个试验的好坏，关键在于它能否迅速而有效地找到或逼近全局最优点 x^* 或全体 x_j^*．不论是追求全局优化亦或是阶段提高都要面对这个根本问题。

多因素全面试验其完全组合次数太多，既难以实现又不很科学。在部分实施中正交表任意两列间具有正交性，即各种水平的搭配均衡而出现次数相等。对于不少于 3 水平的因素，则可画出趋势响应图。通过趋势响应图可粗略地估计 x^*．其理论依据是实验现象中高阶相互作用比低阶相互作用（其中主效应是零阶交互作用），在绝对值上往往倾向于较小。正交设计通常是低比例的部分实施，趋势图的估计虽然相对有效却非绝对可靠，然而若趋势图的单调性表现出规律而且极差极大，则其估计将是很有效的。例如考察 7 因素，填满正交表 $L_{18}(6^1\times3^6)$．若搭配完全则共有 4374 种组合，故很难完成此全面试验。但用上述正交设计则仅需 18 次（组合），对其中各因素作趋势图。科学家试验旨在探讨未知现象。依据大量实例，很多结果显示出有些范围$[a_j, b_j]$尚未包含 x_j^* 的趋势图（特别是对某些因素 j）就这些因素的每一种所取各个水平均不合理。若此 4374 种组合中条件的最好者都远离最优点 x^*，则这些因素水平均不尽合理。若按趋势图将各因素水平调整至较为合理，则需补充少量次数试验，便产生高效率的结果，取得较大幅度的进展。这就是经常使用的正交法。分两批 20 次试验，其结果超过不尽合理的全面试验中的最好者。多因素项目考虑范围宽广，实验者因为知识不足，亦未用过趋势图，把部分因素定量水平选偏，往往在所难免。又如考察 13 因素，填满 $L_{27}(3^{13})$ 正交设计表。若采用全面试验，则完全组合将有 $1594323\approx1.6\times10^6$ 种，这时更难实施。由于考察因素更多，试验范围更易选偏。针对范围选偏的诸因素，补充少量次数合理水平的组合很有必要。用正交设计法分两批约 30 次试验，结果将超过全面试验的完全组合。补充作第 2 批正交试验，并非单一地追求优良组合，而是用正交表设计网络去碰去试，采用多网点，加上趋

势图，逼近最优组合的机会将很大。

就多因素全面试验而言，开始阶段由于组合数量很大，故不应盲目求全，否则将意味着极大浪费。对任意可区分定量水平的因素，则先少分水平，即以 3 或 4 个先作侦察试探、优先考察范围是否选偏？分批进行试验，待选定考察范围包含了最优点 x_j^* 后（$x^* \in [a, b]$）、再逐步加密，求出各个 x_j^*，所考察因素越多、正交设计相对全面试验组合的威力越大，成效越高。

就最简单的双因素水平而言。设因素 A、B 均有两水平 1，2，即 A_1、A_2 与 B_1，B_2。若试验两种搭配组合 A_1B_1 与 A_2B_2．现假定 A_2B_2 结果较好、试问好的原因何在？是由于 A_2 还是由于 B_2 好？区别不开则分拆不清，这种情况叫效应混杂。多试验数据处理忌讳混杂在一起，鱼目混珠是最烦恼的事、正交性要求每对（组）因素的水平组合出现次数相同，正是为了避免混杂；均匀设计涉及大量混杂，牺牲了齐整可比性，客观上构成了一大损失。因此在分批试验寻找各 x_j^* 能力上，比起正交设计要逊色多了。特别是当因素数目和函数峰值较多时，尤其如此。

3 均匀分散性

文献[8] 在其中 3.2 及 3.4 提出用偏差作为度量均匀性的准则，顺便指出一点是文献[8] 中 P_{39} 定义 2 不周详。定义中 Sup 所历经过的矩形均从原点出发、这意味着偏差反映左下的分布比较灵敏而反映右上的分布比较迟钝。以第 1 表 $U_5(5^4)$ 及其 $U_5(5^2)$ 为例，其内部开域和加边闭域的正方形 $K = [0.1, 0.7]^2$ 和 $J = [0.1, 0.7]^2$，面积 $U(K) = U(J) = 0.36$，则落入点数 $n_k = 0$，$n_j = 4$．则 $n_j/n - U(j) = 0.44$，$U(K) - n_k/n = 0.36$ 均大于 $U_5(5^2)$ 的偏差 $D = 0.31$．可见偏差并没有反映出中央区域的分布。因此文[8] 定义 2 中 Sup 所经历的矩形 $[0, x]$ 应扩充为经过的矩形 $[x_c, x]$．人们普遍接受分散性带来突出性的原则。所谓突出性

表 1 第一类变换是取不同的设计先取出一张预备表

实验号	条件号									实验号	条件号								
	1	2	3	4	5	6	7	8	9		1	2	3	4	5	6	7	8	9
1	1	2	3	4	5	6	7	8	9	14	1	2	3	1	2	3	1	2	3
2	1	5	9	4	8	3	7	2	6	15	1	2	3	2	3	1	3	1	2
3	1	8	6	4	2	9	7	5	3	16	1	2	3	3	1	2	2	3	1
4	3	5	7	9	2	4	6	8	1	17	1	3	2	2	1	3	3	2	1
5	4	8	3	7	2	6	1	5	9	18	1	3	2	3	2	1	2	1	3
6	6	2	7	3	8	4	9	5	1	19	2	1	3	1	3	2	3	2	1
7	7	5	3	1	8	6	4	2	9	20	2	1	3	3	2	1	1	3	2
8	9	2	4	6	8	1	3	5	7	21	2	3	1	1	2	3	3	1	2
9	9	5	1	6	2	7	3	8	4	22	2	3	1	3	1	2	1	2	3
10	3	6	9	2	5	8	1	4	7	23	3	1	2	1	2	3	2	3	1
11	7	4	1	8	5	2	9	6	3	24	3	1	2	2	3	1	1	2	3
12	9	8	7	6	5	4	3	2	1	25	3	2	1	1	3	2	2	1	3
13	1	1	1	2	2	2	3	3	3	26	3	2	1	2	1	3	1	3	2

是指逼近绝对极值或全局最优的能力。通过均衡分散性强弱来表述两种设计的优化能力毕竟有些间接。此外将通过计算机验证直接比较两者的平均突出性。为简明起见，此处列举两个实例以表明此趋势。

例[1] $f(x,y,z)=(xyz-x^2+y^2-1.34)^2+(xy-z^2-0.09)+(e^x-e^y+z-0.41)^2$．试求其极小点 $x^{**}=(x^*,y^*,z^*)$，使 $f(x^{**})=f(x^*,y^*,z^*)$ 达到最小值(min = 0)的问题。它等价于三项分别等于零的非线性方程组求解。用正交设计法可方便地求出 $x^*=0.90221$，$y^*=1.10034$，$z^*=0.95013$(计算过程略)。至于均匀设计，则有 9 个条件作均匀设计，由文献[8] P_{70} 给出 $U_9(9^5)$ 及 $U_9(9^4)$ 两种均匀设计。本例中自变量 3 个(x, y, z)，在两表中各取 3 列。就水平选取而言，先选定三维的变化范围，如取 $0\leqslant x, y, z\leqslant 1.8$，$x_j, y_j, z_j$ 的取值均为 $0.2j-0.1$，$j=1.2\cdots9$、正交设计在 $L_9(3^4)$ 中三个水平(x_j, y_j, z_j)均取 $0.6j-0.3$，$j=1,2,3$(水平)。欲比较的是两种设计中 9 次试验的各函数值中最小者，以较小者为佳。以最小值的平均数为指标，称平均突出性。第一类变换是取不同的设计先列出一张预备表(见表 1)．使用有 3 列(本例中水平数目相同)的设计，不妨约定，把自变量 x, y, z 依顺序置放于第 1,2,3 列。对均匀设计作两种置换，表 1 是一种三列间置换，共有 31 种，另一种为逐列对水平数码作正反次置换(即 j 与 1．若干设计的预备表 $1.0-j$ 互换)，则共有 $2^3=8$ 种。$U_9(3^3)$ 的全体共 48 种，形成三套重复试验，共有 $48/3=16$ 种不同设计。按预备表其行号分组分别是 (1,2,3)，(1,2,8)，(1,3,2)，(1,3,9)，(1,4,5)，(1,4,6)，(1,5,4)，(1,5,7)，(1,6,4)，(1,6,7)，(1,7,5)，(1,7,6)(1,8,2)，(1,8,9)，(1,9,3)，(1,9,8)．若 $U_9(9^3)$ 的 48 种变换形成 4 套重复，共有 12 种不同设计，行号组为 (1,1,10)，(1,1,11)，(1,10,1)，(1,10,10)，(1,10,11)，(1,10,12)，(1,11,1)，(1,11,10)，(1,11,11)，(1,11,12)，(1,12,10)，(1,12,11) 正交设计编排原理略有不同，共有 12 种不同排法。取前两行 13，14 作为公共行(列)，则第三行即由第 15 ～ 26 行，共有 12 种不同取法。第 2 类变换是取不同范围。本文中只考虑作平移，正方体的边长和体积都保持不变，以 $J(k,l,m)$ 代表由 $0.3K-0.9\leqslant x\leqslant 0.3K+0.9$，$0.3l-0.9\leqslant y\leqslant 0.3l+0.9$，$0.3m-0.9\leqslant z\leqslant 0.3m+0.9$，组成的正方体，则共有 $6^3=216$ 种不同的范围。其共性是恰好均包括了极小值(x^*, y^*, z^*)．由前面示例布点的范围知 $J(3,3,3)$．此两类变换均相当于函数在自变量之间作了某种变换等价于采用了不同的类似函数。各种设计的最小平均数见表 2．平均数表明两种均匀设计都比不上正交设计突出。由于次数比较少，两方互有好坏，说服力不强，需在更多情况下比较，以最小值的平均数为指标即平均突出性。另外还发现，文献[8]的偏差愈小，其水平数愈差。

表 2　三种设计最小数的平均数

设计	函数数目	平均数	偏差 D
$U_9(9^3)$	16 × 216	2.183	0.310
$U_9^*(9^4)$	12 × 216	2.933	0.298
$L_9(3^3)$	12 × 216	2.083	0.421

表 3　七种设计的计算结果

设计	函数数目	平均数	偏差 D
$U_8^*(8^3)$	8 × 512	6.231	0.200
$U_8(8^2\times4^1)$	24 × 512	6.254	0.231
$U_8(8^1\times4^2)$	24 × 512	6.281	0.282
$U_8(8^1\times4^1\times2^1)$	24 × 512	7.255	0.385
$U_8(8^2\times2^1)$	24 × 512	6.188	0.341
$L_8(2^3)$	1 × 512	5.915	0.578
$L_8(4^1\times2^2)$	12 × 512	5.752	0.508

例[2] $g(x, y, z) = (x - y - 26.2)^2 + (x - yz^{10} - 30.4)^2 + (x - yz^{20} - 36.3)^2 + (x - yz^{30} - 37.8)^2 + (x - yz^{40} - 38.6)^2$,决定非线性回归的三个参数,文献[10]在附录5解出 $g(x, y, z)$ 在 $x^* = 41.707904$, $y^* = 15.838814$, $z^* = 0.9562398$ 处达到绝对极小值 $\min = 2.908947$.共取 $8^3 = 516$ 种长方体变换范围 $J(k, l, m)$,恰好都包含了极小点,就3变量8水平条件,以均匀设计,经过列置换和位级数码的正反对换,选用8种 $U_8^*(8^3)$,24种 $U_8(8^2 \times 4^1)$,24种 $U_8(8^1 \times 4^2)$,24种 $U_8(8^1 \times 4^1 \times 2^1)$,及24种 $U_8(8^2 \times 2^1)$;关于正交设计则选用了1种 $L_9(2^3)$,4种 $L_8(4^1 \times 2^2)$(详排法见另文)。$n = 8$ 个位级水平的取法是将区域以 n 等分,然后取中点,上述7类设计的计算结果见表3.

关于平均突出性,已验证了7种函数(包括上述2例),都说明上交设计较好。因为计算和验证需要按不同设计对不同函数和在不同范围进行,工作量很大,少数人做不周全。有条件和有热忱的同行,建议参加这类验证。均匀设计牺牲了齐整可比性,即使在平均突出性上有些超过,也是得不偿失。

4 其它问题

1) 离散因素 前面讨论只涉及到连续定量因素。在实际工作中常遇到非连续定量因素,如材料和作物品种等,它们有时称作离散型因素(变量)。一次均匀设计大致会是一个品种结合一项栽培条件,则各个品种的栽培条件互不相同,对材料性能大抵相似。对品种的栽培不能算单因素试验,一批试验常需对有关因素的用量与状态作些比较分析,才能有效前进。均匀设计涉及很多混杂,除直接结果比较外,对这种分析的作为不大,故同离散因素的差距较远。节省次数是手段,解决问题是目的。手段应从属目的。过分强调节省次数而采用不优秀的设计则反而降低了效益,从全局看,不是节省而是浪费。

2) 多批试验 对于多分批试验,需逐批搜索。对于少分批试验,常要求它提供系统的看法。又如数据中处理试验误差,均匀设计构造欠齐整,不如正交设计,因而使用上不够清晰有效。[日]田口玄一有一卓越贡献便是用正交表作误差表引进波动,估计误差均方,极小化误差均方,实现稳定性最优化。这对提高产品质量和降低成本起到深远影响。若以均匀表作误差表来估计误差均方,便会感到困难或欠准确。

3) 少次试验 少次试验可使用均匀设计不失为其一优点。当然也可用其它方法,如 $L_4(2^2)$ 加中心点等。超过12次试验,分批用正交设计虽非绝对连接却也基本连接。

4) 设计倾向 对不同的方法可取长补短,在不同阶段适当结合。均匀设计的分批是按试验结果直接逼近最优点。经计算机验证可知,均匀设计的分散性程度很强,突出性效果亦很好,因此特别在试验前期出一些好结果并不偶然。但是对出好成果的均匀设计若改用正交设计,其平均效果可能也会很好。用均匀设计和正交设计从事开发工作,远比不采用设计好得多。

参 考 文 献

[1] YU RUQIN. Introduction to Chemometrics[M]. Changsha: Human Education Publishing House, 1991. 227~223.
[2] DOMING S N, MORGAN S L. Experimental Design: A Chemometric approadch[M]. Amsterdan, Elsesvier, 1987.
[3] XIANG KEFENG, WI JIGIANG. Experimental Design and Data Processing[M]. Shanghai: Sci. & Tech. Press, 1983.

[4] MASSART D L, VANDEGONSTE B G M, DEMING S N et al. Chemometrics[M]: A Textbook. Amsterdam: Elserier, 1988.

[5] GROUP OF PROBABILITY STATISTIC PEKING UNIVERSITY. Orthogonal Design[M]. Beijing: People's Education Press, 1976.

[6] ZHU WEIYONG. Applications of Orthogonal and Regression Orthogonal Experimenal Design[M]. Shenyang: Liaoning People's Press, 1978.

[7] FANG KAITAI. Uniform Design and UD Table[M]. Beijing: Science Press, 1994.

[8] ZHANG CHENGFEN, LI JING. Uniform Design and Statistic Optimization[J]. quality Manag, 1994, (8): 41～42.

[9] CHINESE ASSOCIATION OF PRACTICAL STATISTICS. Orthogonal Method and Third Design[M]. Beijing, Science Press, 1985.

[10] JIAN JISHENG. Some viewpoing on Orthogonal Optimization Method[J]. Acta Appl Math Sin, 1979, 2(4): 376～390.

[11] RATHOWSKY D A. Nonlinear Resgression Modelling[M]. New York: Marcel Dekker Int, 1983.

[12] GROUP FOR THIRD DESIGN OF CHINESE PRACTICAL STATISTICAL ASSOCIATION. Third Design for Computable Project[M]. Beijing: Peking University Press, 1985.

Primary Comparison of Orthogonal and Homogenous Designs on Experimental Design and Optimization

XIA *Zhi-ning*[1], CHEN *Qi-ting*[2], MU *Xiao-jiang*[1], LI *Zhi-liang*[1]

(1. College of Environment, Chemistry and Chemical Engineering, Chongqing University, Chongqing 400044, China; 2. Department of Chemistry, Changsha University of Electrical Power, Changsha 410004, China)

ABSTRACT: Systematic studies are made on experimental design and optimization. The orthogonal array design(OAD) is equal to or better than the homogeneous/uniform design (HUD) in some principlal respects and that the homogeneous design (HUD) has its advantages, especially on the case with few experiments done.

KEYWORDS: chemometrics; experimental design; system optimization; orthogonal array design (OAD); homogeneous / uniform design(HUD)

(责任编辑 刘尚坤)