正交设计的数学理论

马 希 文 编

人 民 教 育 出 版 社

正交设计的数学理论

马希文 编

人民教育出版社

内 容 提 要

本书主要讨论正交设计的统计分析问题，前四章介绍最基本的情况——多因子可加正交设计的统计分析问题，可以作为专门化课教学用书；后六章用划分来处理多因子正交设计、拟水平设计、正规正交表等问题，内容较抽象.

本书可供从事正交设计理论研究的高等学校教师、科技人员、研究生、大学高年级学生参考.

正交设计的数学理论

马希文 编

人民教育出版社出版

新华书店北京发行所发行

尚志印刷厂印刷

*

开本 787×1092 1/16 印张 12.25 字数 270,000

1981年1月第1版 1981年10月第1次印刷

印数 00,001—9,500

书号 13012·0578 定价 0.92 元

前　言

这是一本数学书，它讨论的是正交设计的数学理论，主要是正交设计的统计分析问题.

正交设计是一种安排多因素试验的方法. 经验证明，这个方法简单易行，灵活多样，效果良好. 因此它的应用日益广泛，取得了许多可喜的成绩.

随着正交设计的推广，许多同志迫切需要学习和研究正交设计的理论. 为了适应这种需要，作者编写了这本书.

正交设计的方法是丰富多采的，任何一本详细介绍正交设计的书几乎都讲到“并列法”、“拟水平法”、“拟因子法”、“直和法”、“部分追加法”、“分割法”等内容. 其中有一些是近似方法，例如部分追加法. 但是即使撇开这些不管，要想对各种方法都加以详尽的讨论并得出尽可能一般的结论，也会弄得繁杂不堪，使人得不到要领. 所以本书着重介绍正交设计的基本概念与方法，其他内容则以习题形式出现. 这些习题并不难，且可以使读者逐步掌握如何应用本书的理论去解决各种应用问题的方法.

本书共有十章. 前四章是一个单元，在这四章中讨论了最基本的情况——多因子可加正交设计的统计分析问题，并对它做了一些理论上的说明；还讨论了两个因子的交互作用. 这些章节讲得很细，附有大量的习题. 读者遇到的多数问题都可以从这四章中得到解答.

对于有志于更深入了解正交设计的读者，我们又准备了后六章的材料. 我们向读者推荐了一个工具——划分（第五章），并用它解决了有任意多个因子的交互作用的正交设计的统计分析问题（第六章、第七章）. 然后，我们就去讨论正规正交表的构造与基本性质（第八章）及如何应用正交表来分析正交设计（第九章）. 最后，我们讨论了分割法的一般情况（第十章）. 这一部分比较抽象，初读本书的读者可能感到困难。因此，建议读者在充分掌握了前四章之后再来读后六章.

我们假定读者有数理统计的基础知识并对正交设计的应用有所了解. 所以本书不再对数理统计的基本内容加以说明，也不详细讨论如何把本书的理论与实际应用问题结合起来. 但是，本书中有些地方用到一些数学上稍偏僻一些的内容，如向量空间的投影、有限域等. 这些内容我们也以习题的形式加以整理. 这些习题并不一定要做，有兴趣的读者可以查阅其他书籍.

象本书这样处理正交设计的数学理论，目前恐怕还只是一种探索，缺点、错误在所难免，欢迎读者批评指正.

最后，对于陈家鼎、汪仁官、董自励以及其他许多同志对本书的关心、鼓励与协助，作者谨在此表示深深的谢意，没有他们巨大的支持，本书的完成是不可想象的.

作者

1978 年 1 月

目　　录

第一章　引　言

§1　直观的讨论

本书讨论正交设计的数学理论，所以首先要搞清楚正交设计的基本概念如因子、水平等，从数学上是怎么说的.

如果要用正交设计进行某个实验，比如小麦增产实验，我们首先碰到的是品种、播种时间、播种方式、中耕方法、施肥方式、施肥量等，因为这些都是可能影响实验的结果的因素，这些因素就是农艺条件. 把这些农艺条件确定下来，就可以进行试验，并取得亩产的数据. 在小麦增产实验中，一般要进行许多个这样的试验，分别得到数据，然后加以综合的分析，才能得出必要的结论，最后达到实验的目的.

然而，在一个实验中影响各试验数据的因素很多，其中有一些因素，由于已有足够了解，或限于实验条件，或鉴于计划的安排，是我们在实验中不准备加以研究的. 这些因素，我们通常在整个实验过程中把它们固定下来. 另一些因素，则是我们要加以研究的，通常，我们为它们选择不同的条件分别进行试验以资比较. 这后一类的因素，就叫做实验的因子. 例如，上述的小麦增产实验中，如果取定了小麦品种，那么品种就不叫因子；如果播种方式我们打算取"原有方式"和"改进方式"加以对比，那么播种方式就是一个因子，而"原有方式"和"改进方式"就是这个因子的两个水平.

因此我们可以说：凡对试验的数据有影响的，而且在实验中确定了若干个条件加以对比的因素，就叫做因子；这些条件就分别叫做因子的水平. 这句话阐明了实验的因子、水平、数据这些概念的关系，具体说来，它包括两个内容：

(1) 在一个实验中，每一个因子都有一些确定的、准备加以对比的条件，即水平；

(2) 经过试验，因子对数据可能发生影响.

这样，很容易导致如下看法：把因子看成变量，因子的水平看作它取的值，数据看成因变量，而实验就成了取得函数值的一种手段.

设 A，B 是两个因子，$A_1, \cdots, A_s$ 与 $B_1, \cdots, B_t$ 分别是他们的水平. 当因子 A 取了水平 A_i，因子 B 取了水平 B_j 时，试验结果可以写成 $x(A_i, B_j)$ 或 $x(i, j)$ 或 x_{ij}，这叫条件 A_iB_j 下的数据. 一般说来，设 $F^1, \cdots, F^m$ 是 m 个因子，则在 F^1 取水平 $F^1_{\lambda_1}, \cdots, F^m$ 取水平 $F^m_{\lambda_m}$ 的条件下，或简称在 $F^1_{\lambda_1}\cdots F^m_{\lambda_m}$ 的条件下，数据就是 $x(F^1_{\lambda_1}, \cdots, F^m_{\lambda_m})$ 或简写为 $x(\lambda_1, \cdots, \lambda_m)$ 或 $x_{\lambda_1\cdots\lambda_m}$.

综上所述，我们把实验看成是因子与数据之间的某种函数关系. 但是，这个看法过于概括，过于空泛，因为它只说明实验中的数据依赖于因子的水平如何选取，却丝毫没有说明这种依赖关系是什么关系. 从这里出发，我们很难进行任何有价值的讨论. 因此，这个看法有待充实，有待于从各种实用问题中找到更具体的数量关系来充实它.

以上述的小麦增产实验为例，如果取"播种方式"为因子 A，"施肥方法"为因子 B，对每个因

子都取两个水平如下:

A: 播种方式

A_1: 原有方式

A_2: 改进方式

B: 施肥方法

B_1: 原有方法

B_2: 改进方法

那么,对于条件 $A_1B_1, A_1B_2, A_2B_1, A_2B_2$ 分别进行试验,可以把试验数据排列成如下的二元表:

数据(亩产,斤/亩)		B(施肥方法)	
		B_1(原有方法)	B_2(改进方法)
A(播种方式)	A_1(原有方式)	560	600
	A_2(改进方式)	580	640

这个表的含义是很明显的.

粗略说来,播种方式的改进与施肥方法的改进这两者中的每一个都可以得到增产,而两方面的同时改进,还可以"好上加好". 但从具体数量关系来看,播种方式的改进可以增产20斤/亩;施肥方法的改进可以增产 40 斤/亩;两者同时改进,可以增产 80 斤/亩. 这个 80 斤/亩,并不等于 20 斤/亩加上 40 斤/亩,所以"好上加好"的"加",不能简单地理解为数学上的"加".

这是因为当我们说"播种方式的改进可以增产 20 斤/亩"时,我们是就原有的施肥方法来说的: $580-560=20$. 如果就改进方法来说,这个增产的数值将成为 $640-600=40$. 为了不使"播种方式的改进可以增产"的数值过分依赖于某一种施肥方式,一种自然的想法是用这两者的平均数,即 $(20+40)/2=30$(斤/亩)来表示它. 对于任何一种具体的施肥方法而言,播种方式的改进的效果虽然不能用这个数值精确反映出来,但是这个数却粗略地反映出任意取定了施肥方法之后,播种方式的改进可以产生什么效果. 同理,施肥方法的改进的效果,也应该用 $(40+60)/2=50$(斤/亩)这个数据来表示.

这个观点可以用于更加一般的情况. 设在一个实验中,有 A, B 这两个因子,它们分别有 $A_1, \cdots, A_s$ 与 $B_1, \cdots, B_t$ 这些水平. 用 x_{ij} 表示在条件 A_iB_j 下的数据,那么我们可以列成如下的表:

x	B_1	$\cdots$	B_j	$\cdots$	B_t
A_1	x_{11}	$\cdots$	x_{1j}	$\cdots$	x_{1t}
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
A_i	x_{i1}	$\cdots$	x_{ij}	$\cdots$	x_{it}
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
A_s	x_{s1}	$\cdots$	x_{sj}	$\cdots$	x_{st}

把 A 的水平从 A_1 改为 A_i,产生的效果可以用

$$a_i=\frac{(x_{i1}-x_{11})+\cdots+(x_{it}-x_{1t})}{t}=\frac{1}{t}\sum_{j=1}^{t}x_{ij}-\frac{1}{t}\sum_{j=1}^{t}x_{1j}$$

这个数来表示．通常，用

$$x_{i\cdot}=\frac{1}{t}\sum_{j=1}^{t}x_{ij}$$

表示 A_i 条件下的各数值的平均值，则又可以写出

$$a_i=x_{i\cdot}-x_{1\cdot}$$

这里，计算 A 的某一水平的效果，是用 A_1 即 A 的 1 水平作为基准的．这个作法反映了前例中 1 号水平是“原有方式(方法)”这种情况．一般说来，1 号水平并不占有这种特殊地位，所以我们宁可用所有水平的平均值

$$\bar{x}=\frac{1}{s}\sum_{i=1}^{s}x_{i\cdot}=\frac{1}{st}\sum_{i=1}^{s}\sum_{j=1}^{t}x_{ij}$$

来作为比较的基准，也就是用

$$a_i=x_{i\cdot}-\bar{x}\qquad(i=1,\cdots,s)$$

作为因子 A 的 i 水平 A_i 的效果．同理，对因子 B，我们令

$$x_{\cdot j}=\frac{1}{s}\sum_{i=1}^{s}x_{ij}$$

那么

$$b_j=x_{\cdot j}-\bar{x}\qquad(j=1,\cdots,t)$$

可以用来作为 B 的 j 水平 B_j 的效果．

这时很容易想到，一个因子的各水平的效果的总和是 0．例如

$$\sum_{i=1}^{s}a_i=\sum_{i=1}^{s}x_{i\cdot}-s\bar{x}=0$$

换句话说，一个因子的各水平的效果不是一组无关的数，他们之间存在着约束条件．关于这个问题，下文将有专门的章节讨论．

回到前面的数字例，可以求出

$$\bar{x}=595(\text{斤/亩})$$
$$x_{1\cdot}=580$$
$$x_{2\cdot}=610$$
$$x_{\cdot 1}=570$$
$$x_{\cdot 2}=620$$

以及

$$a_1=-15$$
$$a_2=15$$

$$b_1=-25$$
$$b_2=25$$

如果用原有的播种方式和施肥方法，则得到的效果是 $a_1+b_1=-40$ (斤/亩)，这是对于总平均来说的效果. 从这个观点看来，原有的播种方式和施肥方法应使亩产从平均的 595 斤/亩跌落到 595－40＝555 (斤/亩). 这个数与实际数之差不过是 5 斤/亩. 同样，改进的播种方式与施肥方法应得到 595＋15＋25＝635(斤/亩)，这个数与实际数据之差也不过是 5 斤/亩.

从这个例子来看，各因子的效果差不多是可以叠加的，也就是说:

各因素对数据的总效果＝各因素对数据效果的总和

数据＝总平均＋各因子对数据的效果之和

经验证明，这个公式在许多情况都是大体正确的. 这就是所谓多因子可加实验模型，但是还应在其中补充一个误差项，因为实验总有误差，而且有时误差还相当大，以至我们只有在对实验误差进行了适当处理之后，才能得到正确的结论. 实验误差一般来说可以看成随机变量.

这样一来，上面的讨论可以概括为如下的定义:

定义 设 $F^j=\{F^j_1, \cdots, F^j_{s_j}\}$ $(j=1, \cdots, m)$，是 m 个集合，$\theta^j_k(k=1, 2, \cdots, s_j; j=1, 2, \cdots, m)$，是一组实数，$\theta^0$ 是一个实数，ε 是一个随机变量，则由等式

$$y(\lambda_1, \cdots, \lambda_m)=\theta^0+\sum_{j=1}^{m}\theta^j_{\lambda_j}+\varepsilon$$

确定的关系叫做一个多因子可加实验. 其中 $F^1, \cdots, F^m$ 叫这个实验的因子，$F^j_1, \cdots, F^j_{s_j}$ 叫做 F^j 的水平 $(j=1, \cdots, m)$，$s_1, \cdots, s_m$ 分别叫各因子的水平数，向量

$$\theta^j=\begin{pmatrix}\theta^j_1\\ \vdots\\ \theta^j_{s_j}\end{pmatrix}$$

叫 F^j 的效应值(向量) $(j=1, \cdots, m)$，其各分量分别叫各水平的效应值，θ^0 叫总平均，$\theta^0, \theta^1, \cdots, \theta^m$ 统称为实验的参数，ε 叫实验的误差，$y(\lambda_1, \cdots, \lambda_m)$ 叫做水平搭配 $F^1_{\lambda_1}\cdots F^m_{\lambda_m}$ 或 $(\lambda_1, \cdots, \lambda_m)$ 的试验数据.

读者应注意到，这个定义中没有提到约束条件. 我们将在适当的时候谈这个问题.

最后声明一点，为了避免冗长的定义、定理的叙述，今后凡是由上下文自明的说明，我们一律略去. 这样做，可以使重点更加突出.

习题与补充

(1) 如果在多因子可加实验中所有的因子都是二水平的，那么可以把因子的水平写成 F^j_{-1}, F^j_{+1}，把效应值写成 $\pm a_j$，这时实验本身就可以写成

$$y(\lambda_1, \cdots, \lambda_m)=\theta^0+\sum_{j=1}^{m}(-1)^{\lambda_j}a_j+\varepsilon$$

其中 $\lambda_1, \cdots, \lambda_m$ 都取 ±1 为值. 这叫“称重”实验. 它可以用这样的模型加以说明: 有 m 件物品，重量分别是 a_1,

$\cdots, \alpha_m$, 把它们组合起来放在天平上称重, θ^0 是天平的系统误差, $\lambda_j=1$ 或 -1 分别表示第 j 件物品放在左盘或右盘, $y(\lambda_1, \cdots, \lambda_m)$ 表示在右盘上应放多少砝码.

(2) 在上题中, 我们利用 8 次称重来估计 4 件物品的重量. 取正交表 $L_8(2^7)$ 的 1, 2, 4, 7 各列, 每一列表示一件物品, 每一行表示一次称重(即一个试验), 1 水平理解为"放在右盘", 2 水平理解为"放在左盘". 用 $y_1, \cdots, y_8$ 表示 8 次称重的数据. 试根据(1)写出 $y_1, \cdots, y_8$ 与 $\theta^0, \alpha_1, \alpha_2, \alpha_3, \alpha_4$ 的关系式, 说明以上写出的各式中的误差部分 $\varepsilon_1, \cdots, \varepsilon_8$ 可以看成独立同分布的正态随机变量, 而且期望值是 θ, 方差 σ^2 未知. 按照一般线性统计的方法解出

$$\hat{\theta}^0=W_0/8,\quad W_0=y_1+y_2+y_3+y_4+y_5+y_6+y_7+y_8$$
$$\hat{\alpha}_1=W_1/8,\quad W_1=y_1+y_2+y_3+y_4-y_5-y_6-y_7-y_8$$
$$\hat{\alpha}_2=W_2/8,\quad W_2=y_1+y_2-y_3-y_4+y_5+y_6-y_7-y_8$$
$$\hat{\alpha}_3=W_3/8,\quad W_3=y_1-y_2+y_3-y_4+y_5-y_6+y_7-y_8$$
$$\hat{\alpha}_4=W_4/8,\quad W_4=y_1-y_2-y_3+y_4-y_5+y_6+y_7-y_8$$
$$\hat{\sigma}^2=\left(\sum_{i=1}^{n} y_i^2-(W_0^2+W_1^2+W_2^2+W_3^2+W_4^2)/8\right)/3$$

而且

$$\operatorname{var}\alpha_j=\frac{\sigma^2}{8}$$

由此说明利用正交设计进行称重的应用价值.

(3) 设在多因子可加实验中 $E(\varepsilon)=0$, 则

$$x(\lambda_1, \cdots, \lambda_m)=Ey(\lambda_1, \cdots, \lambda_m)$$

叫做条件 $F^1_{\lambda_1}\cdots F^m_{\lambda_m}$ 的工程平均值. 在因子效应值有约束条件

$$\mathbf{1}'_{s_j}\theta^j=\sum_{j=1}^{s_j}\theta^j_{s_j}=0$$

的前提下证明

$$\theta^0=\frac{1}{s_1s_2\cdots s_m}\sum_{\lambda_1=1}^{s_1}\sum_{\lambda_2=1}^{s_2}\cdots\sum_{\lambda_m=1}^{s_m}x(\lambda_1, \cdots, \lambda_m)$$

$$\theta^j_{\lambda_j}=\frac{s_j}{s_1s_2\cdots s_m}\sum_{\lambda_1=1}^{s_1}\cdots\sum_{\lambda_{j-1}=1}^{s_{j-1}}\ \sum_{\lambda_{j+1}=1}^{s_{j+1}}\cdots\sum_{\lambda_m=1}^{s_m}x(\lambda_1, \cdots, \lambda_m)-\theta^0$$

并写出 $m=2$ 的情况, 与本节中的例题相比较.

§2 设计阵与展开阵

上节末尾的定义只谈了实验, 它是基于对客观数据的分析而对数据的构造提出的一个模型. 它不涉及在一个实验中到底安排怎样的一批试验, 但它又可以用于安排的每一个试验. 至于在一个实验中要安排多大的一批试验以及各号试验中各因子取什么水平, 要用设计矩阵来描述.

例 1 上节考查的小麦增产实验中, 一共进行了四个试验, 列表如下:

试验号	A (播种方式)	B (施肥方法)
1	A_1(原有方式)	B_1(原有方法)
2	A_1(原有方式)	B_2(改进方法)
3	A_2(改进方式)	B_1(原有方法)
4	A_2(改进方式)	B_2(改进方法)

这个表的实质部分是一个矩阵.

$$A=\begin{pmatrix}1&1\\1&2\\2&1\\2&2\end{pmatrix}$$

从上例可以看出,对于一个给定的实验,要描述一批试验是怎样安排的,可以给出一个$n\times m$的矩阵,n是试验的大小(即共要进行几个试验),m是因子数,矩阵的i行j列的元素λ_{ij}表示第i号试验中因子F^j取的水平号.这个矩阵叫做设计阵.这样就不难理解如下的定义:

定义 设给定了一个多因子可加实验.又设

$$A=\begin{pmatrix}\lambda_{11} & \cdots & \lambda_{1m}\\ \cdots\cdots & \cdots\cdots & \cdots\cdots\\ \lambda_{n1} & \cdots & \lambda_{nm}\end{pmatrix}$$

是一个矩阵,其中$\lambda_{ij}\in\{1, 2, \cdots, s_j\}$ $(i=1, \cdots, n; j=1, \cdots, m)$. 则$A$叫做一个设计矩阵.再设$\varepsilon_1, \cdots, \varepsilon_n$是一组随机变量,则等式

$$y_i=\theta^0+\sum_{j=1}^{m}\theta^j_{\lambda_{ij}}+\varepsilon_i \qquad (i=1, \cdots, m)$$

叫做相应于A的设计式.n叫总试验数或试验大小,y_i及ε_i分别叫做i号试验的数据与误差,$i=1, \cdots, n$.

例2 在例1中,相应的设计式是以下四个等式:

$$y_1=\theta^0+\theta^1_1+\theta^2_1+\varepsilon_1$$
$$y_2=\theta^0+\theta^1_1+\theta^2_2+\varepsilon_2$$
$$y_3=\theta^0+\theta^1_2+\theta^2_1+\varepsilon_3$$
$$y_4=\theta^0+\theta^1_2+\theta^2_2+\varepsilon_4$$

其中$\theta^1_1=a_1, \theta^1_2=a_2, \theta^2_1=b_1, \theta^2_2=b_2$.

设计阵简单明了,实用方便.但在理论研究时,由于用它表达的设计式中出现了$\theta^j_{\lambda_{ij}}$这样复杂的项,所以很不方便.为了克服这个困难,我们再研究一种辅助的方法,以便简化设计式.

例3 在例2中,设计式可以用矩阵形式写成

$$\begin{pmatrix}y_1\\y_2\\y_3\\y_4\end{pmatrix}=\begin{pmatrix}1&1&0&1&0\\1&1&0&0&1\\1&0&1&1&0\\1&0&1&0&1\end{pmatrix}\begin{pmatrix}\theta^0\\\theta^1_1\\\theta^1_2\\\theta^2_1\\\theta^2_2\end{pmatrix}+\begin{pmatrix}\varepsilon_1\\\varepsilon_2\\\varepsilon_3\\\varepsilon_4\end{pmatrix}$$

或者写成

$$\begin{pmatrix}y_1\\y_2\\y_3\\y_4\end{pmatrix}=\begin{pmatrix}1\\1\\1\\1\end{pmatrix}\theta^0+\begin{pmatrix}1&0\\1&0\\0&1\\0&1\end{pmatrix}\begin{pmatrix}\theta^1_1\\\theta^1_2\end{pmatrix}+\begin{pmatrix}1&0\\0&1\\1&0\\0&1\end{pmatrix}\begin{pmatrix}\theta^2_1\\\theta^2_2\end{pmatrix}+\begin{pmatrix}\varepsilon_1\\\varepsilon_2\\\varepsilon_3\\\varepsilon_4\end{pmatrix}$$

仔细观察右端的第二、第三项中的两个矩阵，就会发现它们各对应于一个因子，它们的每一列对应于相应的因子的一个水平. 如果在某一号试验中用到了某一个水平，就在矩阵的相应位置上出现一个 1，否则就出现 0.

定义 如果 $A=(\lambda_{ij})$ 是设计阵. 令

$$a^j_{i\lambda}=\begin{cases}1 & \text{如果 } \lambda_{ij}=\lambda \\ 0 & \text{其他}\end{cases}\qquad \begin{pmatrix} i=1,\cdots,n \\ j=1,\cdots,m \\ \lambda=1,\cdots,s_j\end{pmatrix}$$

对任何 $j=1,\cdots,m$，矩阵

$$A_j=\begin{pmatrix} a^j_{11} & \cdots & a^j_{1s_j} \\ \cdots & \cdots & \cdots \\ a^j_{n1} & \cdots & a^j_{ns_j}\end{pmatrix}$$

叫做 A 的关于因子 F^j（或第 j 列）的展开阵. 此外，为了形式上的方便，我们把

$$A_0=1_n$$

叫做总平均的展开阵，把分块矩阵

$$\boldsymbol{A}=(A_0\ A_1\ \cdots\ A_m)$$

叫做 A 的展开阵.

例 4 在例 3 中，

$$A_0=\begin{pmatrix}1\\1\\1\\1\end{pmatrix},\quad A_1=\begin{pmatrix}1&0\\1&0\\0&1\\0&1\end{pmatrix},\quad A_2=\begin{pmatrix}1&0\\0&1\\1&0\\0&1\end{pmatrix}$$

$$\boldsymbol{A}=\begin{pmatrix}1&1&0&1&0\\1&1&0&0&1\\1&0&1&1&0\\1&0&1&0&1\end{pmatrix}$$

如令

$$\boldsymbol{y}=\begin{pmatrix}y_1\\y_2\\y_3\\y_4\end{pmatrix},\quad \varepsilon=\begin{pmatrix}\varepsilon_1\\\varepsilon_2\\\varepsilon_3\\\varepsilon_4\end{pmatrix},\quad \theta^1=\begin{pmatrix}\theta^1_1\\\theta^1_2\end{pmatrix},\quad \theta^2=\begin{pmatrix}\theta^2_1\\\theta^2_2\end{pmatrix},\ \theta=\begin{pmatrix}\theta^0\\\theta^1\\\theta^2\end{pmatrix}$$

则可以把设计式写成

$$y=A_0\theta^0+A_1\theta^1+A_2\theta^2+\varepsilon=\boldsymbol{A}\theta+\varepsilon$$

由展开阵的定义容易证明，对任何 $i=1,\cdots,n$ 及 $j=1,\cdots,m$，$a^j_{i1},\ \cdots,\ a^j_{is_j}$ 中恰好有一个 1，其余都是 0，所以有

$$\sum_\lambda a^j_{i\lambda}=1\qquad (i=1,\cdots,n;\ j=1,\cdots,m)$$

把这个关系用于展开阵，就可以得到如下的引理：

引理1 在上面的定义中，

$$A_j\mathbf{1}_{s_j}=\mathbf{1}_n \qquad (j=0, 1, \cdots, m)$$

注意这里 $j=0$ 的情况，我们把 s_0 理解为 1，以求得形式上的统一。

用展开阵来表达设计，正如前面例子所表明的那样，是十分简明的。我们今后记

$$y=\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad \varepsilon=\begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}, \quad \theta=\begin{pmatrix} \theta^0 \\ \theta^1 \\ \vdots \\ \theta^m \end{pmatrix}$$

则设计式可以写成

$$y=A\theta+\varepsilon$$

或

$$y=\sum_{j=0}^{m}A_j\theta^j+\varepsilon$$

这里请读者注意 $\theta^0, \theta^1, \cdots, \theta^m$ 分别是 $s_0=1, s_1, \cdots, s_m$ 维的列向量，A 则应看成一个分块矩阵。

今后，展开阵将是我们进行理论研究的主要工具。

习题与补充

(1) 对于在 $L_9(3^4)$ 的前三列上安排了三水平因子的实验，写出设计阵、展开阵及设计。

(2) 设在一个实验中 $m=5$，$s_1=s_2=s_3=s_4=2$，$s_5=3$。把 F^1，F^2，F^3，F^4 分别安排在 $L_8(2^7)$ 的 1, 2, 4, 7 各列上。对于这样得到的每一种应进行试验的水平搭配 $F^1_{\lambda_1}F^2_{\lambda_2}F^3_{\lambda_3}F^4_{\lambda_4}$，都与 F^5 的三个水平搭配起来进行试验。如此得到 $L_8(2^7)\times 3$ 的乘积设计，并且对其中的每一种都重复做两次。试写出设计阵、展开阵及设计。(提示：$n=8\times 3\times 2=48$)

(3) 求证：$\mathrm{tr}A_jA_j'=n$。(提示：tr 表示矩阵的迹，即主对角线元素之和)

(4) 求证：$A_j'A_j$ 是一个对角矩阵，其第 λ 个对角线元素是 F^j_λ 在实验中出现的次数。

(5) 设 $n\geqslant s_j$。求证：A_j 非退化的充要条件是 F^j 的任何水平在 A 的第 j 列中都出现。

(6) 设 $s_1, \cdots, s_m\geqslant 2$。一个 $n\times(1+s_1+\cdots+s_m)$ 阶的 0-1 矩阵 $A=(A_0A_1\cdots A_m)$ 如满足

$$A_j\mathbf{1}_{s_j}=\mathbf{1}_n \qquad (j=0, 1, \cdots, m)$$

则它一定是某个设计阵的展开阵。

(7) 设 F^j_λ 是因子 F^j 的水平。展开阵 A_j 的第 λ 列叫做这个水平的特征向量，记做 $a_j(\lambda)$。求证：$\sum\limits_{\lambda=1}^{s_j}a_j(\lambda)=\mathbf{1}_n$，$\|a_j(\lambda)\|^2=\mathbf{1}_n'a_j(\lambda)=F^j_\lambda$ 在实验中出现的次数，此外 $a_j(1), \cdots, a_j(s_j)$ 是一组两两正交的向量，并研究 $a_j(1), \cdots, a_j(s_j)$ 线性无关的充要条件。

§3 对　　比

在 §1 中我们已经注意到，一个多因子可加实验中，任何因子的各水平的效应值都应满足

$$\mathbf{1}_{s_j}'\theta^j=\theta^j_1+\cdots+\theta^j_{s_j}=0 \qquad (j=1, \cdots, m)$$

我们通常把这些关系式叫做各因子的约束条件.

我们没有把约束条件作为实验的定义的一部分,这是因为今后我们将会看到,对于正交设计的理论来说,它并不是必要的,而仅仅提供了讨论的便利.但是,我们今后总是假定约束条件成立,因为这是自然的,而且可以大大简化我们的讨论.虽然在有些情况下,我们还要稍稍改变一下这个条件,那时再另做声明.

本节的目的在于从代数的角度研究约束条件的含义,以后还要从统计的角度研究它.

回顾§1,我们曾把一个因子的效果的数量描述归结为一组数之间的比较,就是这个因子的各水平的数据的平均值之间的比较.在那里,我们曾取过某一个水平作为比较的基准,直接把各水平与这个水平加以比较,后来又取总平均作为比较的基准.

一般说来,比较一组数 $a_1, \cdots, a_n$ 都可以用它们的平均数 $\bar{a}=\sum_{i=1}^{n} a_i$ 做为基准.这时我们令

$$a_i' = a_i - \bar{a},$$

a_i' 的符号+,-表示 a_i 在这一组数中是偏大的还是偏小的,a_i' 的绝对值的大小则表示 a_i 在这一组数中是与众不同的还是比较平常的.我们把 a_i' 叫做 a_i 在这一组数中的变动.

当然我们也经常要直接地对这一组数做两两比较,就是说,考查 $a_i - a_j$,其中 $i \neq j$.这叫做 $a_1, \cdots, a_n$ 的一个简单对比.在代数研究中,我们还会经常地遇到各简单对比的种种线性组合,并把它叫做数组的对比.

例1 $a_2 - a_1$,$a_3 - a_1$ 是简单对比,所以

$$a_2 + a_3 - 2a_1 = (a_2 - a_1) + (a_3 - a_1)$$

是 a_1, a_2, a_3 的对比.同理

$$-na_1' = \sum_{i=1}^{n} a_i - na_1 = \sum_{i=2}^{n} (a_i - a_1)$$

及

$$a_1' = -\frac{1}{n}(a_2 - a_1) - \frac{1}{n}(a_3 - a_1) - \cdots - \frac{1}{n}(a_n - a_1)$$

都是 $a_1, \cdots, a_n$ 的对比.

一组数的对比归根到底是这一组数的线性组合.下面的引理告诉我们,一组数的线性组合什么时候是这组数的对比:

引理1 设 $u = c_1 a_1 + \cdots + c_n a_n$ 是 $a_1, \cdots, a_n$ 的线性组合,则以下三个命题等价:

(1) u 是 $a_1, \cdots, a_n$ 的对比.

(2) 当 $a_1 = \cdots = a_n = 1$ 时 $u = 0$.

(3) $c_1 + \cdots + c_n = 0$.

证明:(1)推出(2),(2)推出(3)都是明显的.(3)推出(1)可以由下式看出:

$$\begin{aligned} c_1 a_1 + \cdots + c_n a_n &= c_1(a_1 - a_1) + \cdots + c_n(a_n - a_1) + (c_1 + \cdots + c_n) a_1 \\ &= c_2(a_2 - a_1) + \cdots + c_n(a_n - a_1) \end{aligned}$$

请读者完成这个证明.

从这个引理可以看出,一组数的某个线性组合是否是一个对比,完全归结于线性组合的系数的性质,而这些系数又可以看成是向量,所以我们可以给出如下的定义:

定义 n 维空间$\mathscr{R}^n$中的向量 c 如满足$c'\mathbf{1}_n=0$,就叫做 $\mathscr{R}^n$ 的一个对比向量或简称 $\mathscr{R}^n$ 中的一个对比.

这样一来,如果把数组 $a_1,\cdots,a_n$ 写成一个向量a,则当而且仅当 c 是$\mathscr{R}^n$中的一个对比时,$c'a$ 是 a 的一个对比.

例 2 从例 1 可知$\begin{pmatrix}-1\\+1\\0\end{pmatrix}$, $\begin{pmatrix}-1\\0\\+1\end{pmatrix}$, $\begin{pmatrix}-2\\+1\\+1\end{pmatrix}$都是 $\mathscr{R}^3$ 中的对比,而

$$\begin{pmatrix}1-\frac{1}{n}\\-\frac{1}{n}\\\vdots\\-\frac{1}{n}\end{pmatrix}$$

是 $\mathscr{R}^n$中的一个对比. 此外,如果 $a_1,\cdots,a_n$ 是一组数,则它们的变动 $a_1',\cdots,a_n'$也组成 $\mathscr{R}^n$ 中的一个对比.

$$\begin{pmatrix}a_1'\\\vdots\\a_n'\end{pmatrix}=a-\bar{a}\mathbf{1}_n=a-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n'a=\left(I_n-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\right)a$$

这是因为

$$\mathbf{1}_n'\begin{pmatrix}a_1'\\\vdots\\a_n'\end{pmatrix}=\mathbf{1}_n'\left(I_n-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\right)a=\left(\mathbf{1}_n'-\frac{1}{n}\mathbf{1}_n'\mathbf{1}_n\mathbf{1}_n'\right)a=0$$

$\mathscr{R}^n$的对比都满足 $c'\mathbf{1}_n=0$, 也就是都与 $\mathbf{1}_n$ 正交. 可见全体对比恰好组成 $\mathscr{R}^n$ 中$[\mathbf{1}_n]$的余子空间(这里$[\mathbf{1}_n]$表示 $\mathbf{1}_n$ 生成的子空间). 用 D^n 表示这个子空间,则有

$$\mathscr{R}^n=[\mathbf{1}_n]+D^n$$

这里的"+"表示子空间的正交和,所以上式是 $\mathscr{R}^n$ 的正交分解式.

设 a 是任何一组数组成的向量, 我们把 a 投影到$[\mathbf{1}_n]$和 D^n 上, 由于 $\mathrm{Pr}_{[\mathbf{1}_n]}=\mathbf{1}_n(\mathbf{1}_n'\mathbf{1}_n)^{-1}\mathbf{1}_n'=\frac{1}{n}\mathbf{1}_n\mathbf{1}_n'$, 所以有

$$\mathrm{Pr}_{[\mathbf{1}_n]}a=\frac{1}{n}\mathbf{1}_n\mathbf{1}_n'a=\bar{a}\mathbf{1}_n$$

可见

$$\mathrm{Pr}_{D^n}a=a-\mathrm{Pr}_{[\mathbf{1}_n]}a=a-\bar{a}\mathbf{1}_n$$

结合例 2 中的讨论可以知道：一个数组的变动组成的向量就是这个数组对应的向量在 D^n 中的投影.

另一方面，由正交分解式的性质可以得到：

$$\|a\|^2=\|\mathrm{Pr}_{[1_n]}a\|^2+\|\mathrm{Pr}_{D^n}a\|^2$$

就有

$$\sum_{i=1}^n a_i^2=n\bar{a}^2+\sum_{i=1}^n (a_i')^2$$

这叫数组 a 的平方和的分解式．从此又可以得到

$$\sum_{i=1}^n (a_i')^2=\sum_{i=1}^n a_i^2-n\bar{a}^2$$

即

$$\sum_{i=1}^n (a_i-\bar{a})^2=\sum_{i=1}^n a_i^2-n\bar{a}^2$$

上式左端叫做数组 a 的变动平方和或简称 a 的平方变动.

D^n 是一个 $n-1$ 维的空间，而 n 个数的变动是 n 个对比，所以这些变动就不是独立的．事实上，我们已经知道 $\sum_{i=1}^n a_i'=0$．选取 D^n 的一个标准正交基 b_1, …, b_{n-1}，我们可以把向量 $\begin{pmatrix}a_1'\\ \vdots\\ a_n'\end{pmatrix}$ 用 $n-1$ 个对比表示出来，即写成

$$\begin{pmatrix}a_1'\\ \vdots\\ a_n'\end{pmatrix}=u_1b_1+\cdots+u_{n-1}b_{n-1}$$

其中 $u_1,\cdots,u_{n-1}$ 都是 a 的对比．为了证明这一点，我们只须指出，$u_1,\cdots,u_{n-1}$ 是 $\begin{pmatrix}a_1'\\ \vdots\\ a_n'\end{pmatrix}$ 在 b_1, …, b_{n-1} 上的投影值，就是说

$$u_t=b_t'\begin{pmatrix}a_1'\\ \vdots\\ a_n'\end{pmatrix}=b_t'\left(I_n-\frac{1}{n}1_n1_n'\right)a \qquad (t=1,\cdots,n-1)$$

而 b_t 是 $\mathscr{R}^n$ 中的对比，$b_t'1_n=0$，从而

$$u_t=b_t'a \qquad (t=1,\cdots,n-1)$$

可见 u_t 是 a 的对比，$t=1,\cdots,n-1$.

另一方面，由 $b_1,\cdots,b_{n-1}$ 是两两正交的，所以

$$\sum_{i=1}^n (a_i')^2=\|u_1b_1+\cdots+u_{n-1}b_{n-1}\|^2=\sum_{i=1}^{n-1}u_i^2$$

这就把 a 的变动平方和写成了 a 的 $n-1$ 个对比的平方和.

在应用问题中，常把

$$\sigma^2(a)=\frac{1}{n-1}\sum_{i=1}^{n}(a_i-\bar{a})^2$$

叫做 a 的均方变动，这就是 $u_1^2, \cdots, u_{n-1}^2$ 的平均值.

回想本节开头讲的约束条件，我们就看到:

(1) θ^j 作为一个数组，其变动就是它自身. (其实，θ^j 本身就是另一个数组的变动组成的向量.)

(2) $(\theta^j)'\theta^j$ 是它的平方变动.

今后，我们常把 $p_j=s_j-1$ 叫做因子 F^j 的自由度，把

$$\sigma_j^2=\sigma^2(\theta^j)=\frac{1}{p_j}\sum_{\lambda=1}^{s_j}(\theta_\lambda^j)^2$$

叫做 F^j 的均方效应.

习题与补充

(1) 设 $\mathscr{B}$ 是 $\mathscr{R}^n$ 的子空间，那么 $\mathscr{R}^n$ 的任何向量 v 可以写成 $\mathscr{B}$ 中的向量 v_1 和一个与 $\mathscr{B}$ 正交的向量 v_2 的和: $v=v_1+v_2$. 这里 v_1, v_2 是由 v 唯一确定的，v_1 叫 v 在 $\mathscr{B}$ 上的投影，记做[①] $\mathrm{Pr}_B v$ 或 $\mathrm{Pr}(\mathscr{B})v$. Pr_B 是 $\mathscr{R}^n$ 到 $\mathscr{B}$ 的线性变换，所以也用 Pr_B 表示这个线性变换的矩阵. Pr_B 是幂等的，就是说 $\mathrm{Pr}_B\mathrm{Pr}_B=\mathrm{Pr}_B$. 此外有: $\mathrm{Pr}_{R^n}=I_n$. 如果 $\overline{\mathscr{B}}$ 是 $\mathscr{B}$ 的子空间，那么还有 $\mathrm{Pr}_{\overline{B}}=\mathrm{Pr}_{\overline{B}}\mathrm{Pr}_B$. 当 $\mathscr{B}$ 是由非退化矩阵 B 的列向量生成的，即 $\mathscr{B}=[B]$ 时，$\mathrm{Pr}_B=B(B'B)^{-1}B'$. 特别，如果 B 的列向量构成 $\mathscr{B}$ 的标准正交基，则 $B'B=I$, $\mathrm{Pr}_B=BB'$.

(2) 取定 $\mathscr{D}^n$ 的一个标准正交基 $B=(b_1\cdots b_{n-1})$，看成 $n\times(n-1)$ 的矩阵，求证 B 可以加一列 $\frac{1}{\sqrt{n}}1_n$ 变为正交阵. 由此证明 $B'B=I_{n-1}$, $\mathrm{Pr}_{D^n}=BB'=I_n-\frac{1}{n}1_n1_n'$.

(3) 求证 $I_n-\frac{1}{n}1_n1_n'$ 是幂等阵.

(4) 设 $a\in\mathscr{R}^n$, b 是由 a 的变动组成的向量，$c'a$ 是 a 的对比，则 $c'a=c'b$.

(5) 在灵活运用正交设计时，写出 $\mathscr{D}^n$ 的标准正交基的技巧十分重要. 试写出 $\mathscr{D}^2$, $\mathscr{D}^3$ 的一个标准正交基.

(6) 证明

$$\begin{aligned}
u_1&=a_1-a_2\\
u_2&=a_1+a_2-2a_3\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
u_{n-1}&=a_1+a_2+\cdots+a_{n-1}-(n-1)a_n
\end{aligned}$$

是 a_n 的一组互相正交的对比. 把它们标准化，以得到 $\mathscr{D}^n$ 的一个标准正交基.

(7) 设 $n=n_1+n_2$. 如果 B_1, B_2 是 $\mathscr{D}^{n_1}$, $\mathscr{D}^{n_2}$ 的标准正交基，则适当选取 α_1, α_2 可以使 $\begin{pmatrix}\alpha_1 1_{s_1} & B_1 & 0\\ \alpha_2 1_{s_2} & 0 & B_2\end{pmatrix}$

---

① 这里的 B 即 $\mathscr{B}$. 以后凡是上、下角的草体字母均改为相应的大写字母，含义可从上、下文知道.