高等学校教学用书

正交及回归试验设计方法

高等学校教学用书

正交及回归试验设计方法

东北工学院 高允彦 编

冶金工业出版社

高等学校教学用书

正交及回归试验设计方法

东北工学院　高允彦　编

冶金 工业 出版社 出版
（北京北河沿大街嵩祝院北巷39号）
新华书店北京发行所 发行
冶金工业出版社印刷厂印刷

*

850×1168 1/32　印张 8 字数 210 千字
1988年 5 月第一版　　1988年 5 月第一次印刷
印数00,001～3,800册
ISBN 7-5024-0186-5

TF·67（课）　定价1.60元

前　言

在生产及科学试验中，为了达到预期的目标就需要找出使该目标达到最优的一些条件（因素或变量）的数值；或者建立目标与变量之间的关系式，以便确切地了解它们之间的内在联系。如何才能够又快、又好地完成这项工作呢？这类问题在数学上称为最优化问题。

最优设计的理论及应用得到了迅速的发展，在工农业生产及科学试验过程中，在寻求新工艺参数、新配方、提高产品质量、增加产量及降低成本方面、制定产品标准以及在进行数据处理和寻求经验公式等方面都取得了很大的成绩。

解决最优化问题的方法大致分为两类：一类是间接最优化（解析最优化）方法；另一类是直接最优化（试验最优化）方法。所谓间接最优化方法就是要求把所研究的对象（如物理或化学过程）用数学方法描述出来，然后再用数学解析的方法求出最优解。但是在很多情况下，事先无法用数学方法描述，对于这种情况有一种可供采用的办法，就是通过大量试验构造一类函数来逼近这些试验数据，从函数类中求最优解，并通过试验进行验证。然而也有很多实际问题可直接通过少量试验，根据试验结果的比较、分析求得最优解，这就是直接最优化方法。

书中所介绍的均属与直接最优化有关的方法问题，并未涉及到有关的理论。本书共分五部分，包括优选方法，正交试验设计方法，回归分析方法及其试验设计方法，混料回归试验设计方法等内容。除了第一部分以较少的篇幅介绍了优选法外，其它几部分内容均有内在联系。初学者应按顺序阅读，在掌握书中所介绍的安排和组织试验的方法以及对试验结果的分析方法后，便可以在生产和试验过程中应用这些科学的方法。

本书曾编成讲义，供大学高年级学生作为选修课程使用。这次出版前又进行了修改，可供生产、科研部门的工程技术人员及具有高中文化程度的在工农业生产第一线工作的人员参考。

本书在编写过程中得到编者所在教研室各位同志的支持，东北工学院数学系计算优化教研室王作昌副教授对该书进行了审查和校对，在此一并向他们表示衷心的感谢。

由于水平有限，实践经验不足，错误和不当之处敬请读者批评指正。

1987年2月

目　录

0.618法，亦称黄金分割法。0.618是：

$$\omega=\frac{\sqrt{5}-1}{2}=0.6180339887\cdots\cdots$$

的四位有效近似值。黄金分割法也就是先在：

$$x_1=A+(B-A)\omega$$

处做第一次试验，再在x_1的对称点：

$$x_2=B-(B-A)\omega$$

处做第二次试验，如果函数$f(x)$为单峰函数，可在第二次试验的基础上比较试验结果$y_1=f(x_1)$及$y_2=f(x_2)$哪个大。如$f(x_1)$大，就去掉（A，x_2）部分，在留下的范围，又形成一个新的含优区间（x_2，B）继续试验，因该范围中已有了一个试验点x_1，然后再用上述求对称点的方法求出x_1的对称点做第三次试验，如此一直做到所需的满意结果为止。

上例中熔炼的是特种钢，含优区间$[A,B]$为$[1000,2000]$，为提高其强度而加入某种元素，这是一个单极值的单峰函数问题。为了寻求最大值点，分别在x_1点和x_2点做二次试验，其中

$$\begin{aligned}x_1&=A+(B-A)\omega\\&=1000+(2000-1000)\times0.618\\&=1618\text{克}\\x_2&=B-(B-A)\omega\\&=2000-(2000-1000)\times0.618\\&=1382\text{克}\end{aligned}$$

把加入量为1618克和1382克的试验结果进行比较，如x_2点的强度大于x_1点的强度，则把$x_1\sim B$部份去掉，在余下的区间$[A,x_1]$又成为一个缩小了的含优区间，该区间内已有了一个试验点x_2，继续试验时再求x_2的对称点x_3，试验点的安排见图1-1。

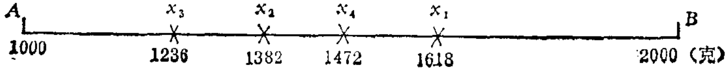

图 1-1

$$
\begin{aligned}
x_3 &= x_1 - (x_1 - A)\omega \\
&= 1618 - (1618 - 1000) \times 0.618 \\
&= 1236\text{克}
\end{aligned}
$$

再比较x_3点和x_2点的强度，若x_2点的强度仍大于x_3的强度，则去掉$A \sim x_3$部分，在余下的[x_3，x_1]区间求出x_2的对称点$x_4 =$ 1472克处试验，再对x_2、x_4进行比较，如此不断缩小含优区间，不断在新区间内求已知试验点的对称点，并对二点加以比较，直至获得满意结果为止。

例 1-1 沸腾炉在运行过程中，飞灰中的含炭量高达40%以上，为降低飞灰中的含炭量，在炉内设二次风装置，使煤粉充分燃烧，对于提高锅炉效率，节约煤炭具有十分重要的意义。二次风使炉内形成风幕，风速太大，有过剩空气产生，效果变坏；风速小，燃料不能充分燃烧，二次风究竟选用多大风速合适呢？试验结果表明，风速在14米/秒以下时，飞炭量占飞灰量的40%以上。风速在25米/秒以上时，飞炭量有上升的趋势。因此确定优选范围在14～25米/秒之间。考查指标：飞灰量（公斤/小时）越少越好。因此这是在14米/秒(A)～25米/秒（B）含优区间内，求谷值点的问题。

首先，确定第一个试验点x_1：

$$
\begin{aligned}
x_1 &= A + (B - A)\omega \\
&= 14 + (25 - 14) \times 0.618 \\
&= 20.8(\text{米/秒})
\end{aligned}
$$

第二个试验点x_2与点x_1对称：

$$
\begin{aligned}
x_2 &= B - (B - A)\omega \\
&= 25 - (25 - 14) \times 0.618 \\
&= 18.2(\text{米/秒})
\end{aligned}
$$

见图1-2。

经测试，x_1点的飞灰量为290公斤/小时；x_2点飞灰量为370公斤/小时。因点x_1的试验结果比点x_2好，便去掉$A \sim x_2$部份，在余下的区间[x_2，B]成为一个新的含优区间，该区间已含有一个试

图 1-2

验点x_1，则第三个试验点x_3与点x_1对称：

$$\begin{aligned}x_3 &= x_2+(B-x_2)\omega\\ &=18.2+(25-18.2)\times 0.618\\ &=22.4(米/秒)\end{aligned}$$

x_3点的飞灰量为280公斤/小时，同时飞灰量中含炭量下降到20%，而B点的飞灰量为281公斤/小时。只经过三次试验便得到满意的结果，并选定二次风风速22.4米/秒为较好点。

1.1.2　0.618法的初步解释

这里回答两个问题：为什么取0.618；又为什么都是对称取点。

优选法不仅能找到最优点，而且还要以最少的试验次数找到，因此，需正确选定试验点。如果第一个试验点取在含优区间〔A，B〕内的x_1点处是合适的，则x_1在〔A，B〕内应有合适的比例位置：

$$\frac{x_1-A}{B-A}=\omega \tag{1-1}$$

目的是求最优点，但是最优点在什么位置，好又好到什么程度都不知道，即使得到x_1点的试验效果y_1，也无法辨认x_1的好坏。只有再找一点进行比较，于是在〔A，B〕中再找一点x_2，取得x_2点的试验效果y_2，比较y_1和y_2可鉴别x_1和x_2的好坏，对任意的单极值函数，无法预先估计x_1和x_2谁好、谁差，因此，根据试验结果去掉$A\sim x_2$或$x_1\sim B$的可能性是同等的，这便要求$A\sim x_2$和$x_1\sim B$两部分的长度相等，即：

$$x_2-A=B-x_1 \tag{1-2}$$

这就是选取x_2和x_1点在含优区间中点对称的原因。

如果式（1-1）中的比例常数ω取得适当，无论去掉哪一部

分，则余下的新的含优区间$[x_2, B]$或$[A, x_1]$内仍有一个已试点x_1或x_2，基于前述观点，新含优区间内已试点应在该区间内占有同样合适的比例位置。假定x_2比x_1好，余下的含优区间为$[A, x_1]$，则应满足：

$$\frac{x_2-A}{x_1-A}=\frac{x_1-A}{B-A} \tag{1-3}$$

由此可见x_1和x_2应同时满足（1-2）式和（1-3）式，即满足联立方程：

$$\begin{cases} x_2-A=B-x_1 \\ (x_1-A)^2=(B-A)(x_2-A) \end{cases} \tag{1-4}$$

解（1-4）式，有：

$$x_1=A+(B-A)\frac{\pm\sqrt{5}-1}{2}$$

$$x_2=A+(x_1-A)\frac{\pm\sqrt{5}-1}{2}$$

很明显x_1是$[A, B]$的内分点，x_2是$[A, x_1]$的内分点，由（1-1）式知$\omega>0$，可见：

$$\omega=\frac{\sqrt{5}-1}{2}$$

所以

$$\frac{x_2-A}{x_1-A}=\frac{x_1-A}{B-A}=\omega=\frac{\sqrt{5}-1}{2}=0.618\cdots\cdots$$

1.1.3 分数法

有一些试验的试验点只能取整数，不可能是0.618的倍数，也有的预先规定了试验总次数，这时应用分数法比0.618法更为方便，试验结果的精度更高。

任何一个小数都可以用分数表达或用分数近似表达，如圆周率π用小数可表示为3.1415926……，用分数可近似表示为$\frac{355}{113}$。0.6180339887……也可以用一批渐近分数：$\frac{3}{5}$；$\frac{5}{8}$；

$\frac{8}{13}$；$\frac{13}{21}$；$\frac{21}{34}$；$\frac{34}{55}$；$\frac{55}{89}$……近似表示。这批渐近分数的构成规律是斐波那奇数列：

1，1，2，3，5，8，13，21，34，55，89，……相邻两数的商。这个数列从第三项起每一项都是前面两项的和。

斐波那奇数列一般表达式是：

$$F_n=\frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^{n+1}-\left(\frac{1-\sqrt{5}}{2}\right)^{n+1}\right]\quad (n=0,1,2,\cdots\cdots)$$

不难得出：

$$\lim_{n\to\infty}\frac{F_n}{F_{n+1}}=\frac{\sqrt{5}-1}{2}=0.618\cdots\cdots$$

由渐近性质可以看到，0.618法和分数法的差异不大。

分数法的具体使用方法如下：

例 1-2 卡那霉素发酵液生物测定，当培养温度为37±1°C时，培养时间在16小时以上，为缩短培养时间、决定优选培养温度；试验范围定为29～50°C，精确度要求±1°C。由给出条件可知，测量温度应以整数温度为宜，因此用分数法安排试验。中间试验点共有20个，见图1-3。

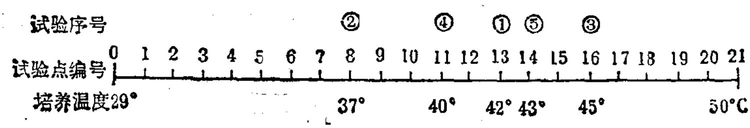

图 1-3

第一个试验点选在整个试验区间接近0.618的$\frac{13}{21}$处，即第13号试验点，42°C处；第二个试验点选在与第一个试验点的对称点上，即第8号试验点（37°C），经比较①点培养时间比②点短，故去掉8号以下的区间，在8号至21号新含优区间内找①点的对称点为16号（③点），该点仍处在新含优区间内接近0.618的

$\frac{8}{13}$处。经比较①点培养时间比③点短，去掉16号以上的区间，在余下的新区间找①点的对称点为④点，经比较试验结果，①点好于④点，又把含优区间缩小到11号至16号，并找到①点的对称点为⑤点（14号、43°C），经过5次试验，证明在42°～43°C培养较好，只需8～9小时。

例 1-3 选择一个电阻，调试电器设备的线路。调试者手里只有几种阻值不等的电阻，阻值分别为0.5，1.0，1.3，2.0，3.0,5.0,5.5（千欧）等七种，要求优选一个合适的电阻。

首先把这些电阻由小到大顺序排列并编号如下：

阻值（千欧）		0.5	1.0	1.3	2.0	3.0	5.0	5.5	
排 列	(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)

为使试验点适于分数法的某一分母值，我们在该排列两端增加虚点（0）和(8)，也就是假设我们手中有对应（0）和（8）的两种电阻，然后在0～8之间运用分数法。第一个试验点用分数5/8在（5）点，即阻值选在3（千欧）处，第二个试验点选在(5)点的对称点（3）点上，阻值为1.3（千欧），如此做下去就可较快找到较好点。

通过该例可看出，有时试验范围中的份数不够分数法中的分母数，若试验范围的份数为10份时，可用两种方法解决，一种是能否缩短试验范围，如能缩短二份，则可应用5/8确定第一个试验点，如不便缩短，还可用第二种办法，即虚设几个份数凑到13份应用8/13确定第一个试验点。

1.1.4 均分法、对分法、0.618法、分数法的试验精确度

均分法是把含优区间均匀分成若干等份，并在每一分点上进行试验，同时对各分点的试验结果进行比较，选出极值点。

对分法，在含优区间内，如果做一次试验便可决定下次试验的试验点的方向，则可用对分法，而不必先做两个试验进行比较

后找出好点再确定第三次试验点的位置。对分法总是在含优区间的中点安排试验，根据中点的试验结果确定取舍哪一部份，而后在新含优区间的中点再进行试验，如此循环直至获得满意的结果为止。因此对分法也叫做平分法。

试验的“精确度”是指试验得到的最优点到真正的最优点的最大可能的距离。设含优区间$[A, B]$，长为$l=B-A$，对于均分法来说经n次试验后精确度是$l/n+1$。也可推出0.618法的精确度是$(0.618)^n l$，分数法的精确度是l/F_{n+1}，平分法的精确度是$l/2^n$。

我们把$(n+1)$；F_{n+1}；2^n；$1/(0.618)^n$分别叫做均分法，分数法，平分法和0.618法的n次试验的“搜索范围”，并用L_n表

表 1-1 四种方法的部份搜索范围（L_n）值

n	均分法	0.618法	分数法	平分法
	$L_n=n+1$	$L_n=\frac{1}{(0.618)^n}$	$L_n=F_n+1$	$L_n=2^n$
1	2	1.618	2	2
2	3	2.619	3	4
3	4	4.237	5	8
4	5	6.855	8	16
5	6	11.091	13	32
6	7	17.945	21	64
7	8	29.0	34	128
8	9	47.0	55	256
9	10	76.0	89	512
10	11	123	144	1024
11	12	199	233	2048
12	13	322	377	4096
13	14	521	610	8192
14	15	843	987	16384
15	16	1364	1597	32768
16	17	2207	2584	65536
17	18	2571	4181	131072
18	19	5778	6765	262144
19	20	9349	10946	524288
20	21	15127	17711	1048576

示。各种方法的搜索范围都可预先算出来，知道了试验范围长度l和试验次数n，试验的精确度可按l/L_n进行计算。表1-1给出了四种方法的L_n的部分数值，当试验次数增加时，搜索范围增加的速度称为这些方法的“搜索效率”，可见均分法的“搜索效率”最低，平分法最高。

根据试验次数，可对各种不同的方法计算出试验结果的精确度，其数值越小，说明该试验结果的自变量越接近最优点。

1.2 分批试验法

前面介绍的0.618法，分数法，平分法等都是根据前面的试验结果再安排后面的试验，这种安排试验的方法叫做序贯试验法，它的优点是总试验次数很少，但往往有时做完一个试验要较长的时间才能得到试验结果，若等到上次结果出来后才安排下次试验，要花费较多的时间，为加速进行试验，常采用一批同时做几个试验的方法，它也适用于试验结果必须在同一条件下才能得到鉴别的情况。根据不同的条件和要求，分批试验的方法也不同。这里只介绍每批试验数目都相同的分批试验法。

1.2.1 预给要求法

如能预先确定总的试验个数（知道了试验范围和所要求的精确度也可计算出试验个数），或预先限定了试验的批数和每批试

表 1-2 每批做两个试验

预给试验批数	含优区间等分份数	第一批试验点	图示	精确度
一	3	1，2	△△	$l/3$
二	7	3，4	○○△△○○	$l/7$
三	15	7，8	6△△6	$l/15$
四	31	15，16	14△△14	$l/31$
五	63	31，32	30△△30	$l/63$
n	$L_n^{(2)}=2^{n+1}-1$	$L_{n-1}^{(2)}$，$L_{n-1}^{(2)}+1$	$L_{n-1}^{(2)}-1$△△$L_{n-1}^{(2)}-1$	$l/L_n^{(2)}$

验的个数，可使用预给要求法。

根据每批试验的个数和预先规定的试验批数，已分别把较好的试验方案列成表格备用。

（1）每批做偶数个试验时，见表1-2，表1-3，表1-4。

表 1-3　每批做四个试验

预给试验批数	含优区间等分份数	第一批试验点	图　示	精确度
一	5	1，2，3，4	△△△△	$l/5$
二	17	5，6，11，12	○○○○△△○○○○△△○○○○	$l/17$
三	53	17，18，35，36	16△△16△△16	$l/53$
四	161	53，54，107，108	52△△52△△52	$l/161$
五	485	161，162，323，324	160△△160△△160	$l/485$
n	$L_n^{(4)}=2\times 3^n-1$	$L_{n-1}^{(4)}$，$L_{n-1}^{(4)}+1$ $2L_{n-1}^{(4)}+1$，$2L_{n-1}^{(4)}+2$	$L_{n-1}^{(4)}-1$△△$L_{n-1}^{(4)}-1$△△$L_{n-1}^{(4)}-1$	$l/L_n^{(4)}$

表 1-4　每批做六个试验

预给试验批数	含优区间等分份数	第一批试验点	图　示	精确度
一	7	1，2，3，4，5，6	△△△△△△	$l/7$
二	31	7，8，15，16，23，24	6△△6△△6△△6	$l/31$
三	127	31，32，63，64，95，96	30△△30△△30△△30	$l/127$
四	511	127，128，255，256，383，384	126△△126△△126△△126	$l/511$
n	$L_n^{(6)}=2\times 4^n-1$	$L_{n-1}^{(6)}$，$L_{n-1}^{(6)}+1$ $2L_{n-1}^{(6)}+1$，$2L_{n-1}^{(6)}+2$ $3L_{n-1}^{(6)}+2$，$3L_{n-1}^{(6)}+3$	$L_{n-1}^{(6)}-1$△△$L_{n-1}^{(6)}-1$△△$L_{n-1}^{(6)}-1$△△$L_{n-1}^{(6)}-1$	$l/L_n^{(6)}$

以表1-2每批做两个试验为例，说明该表的使用方法。若只做一批试验，把含优区间均分三等份，分别在两个分点1，2上试

验，如图1-4所示。

图 1-4

若做两批试验，分含优区间为七等份，在3，4两点上做第一批试验，如图1-5所示。如果4点的结果比3点好，则第二批试验做5，6两点，反之第二批试验做1，2两点。

图 1-5

若做三批试验，含优区间分为15等份，第一批试验在7，8两点做，如第7点好，则把8点以上的区间去掉，如第8点好，则去掉第7点以下的部分。在余下的区间再做两批试验，即如上述两批试验的情况，以此类推可做更多批数的试验。表1-2中图示的△表示第一批试验点的位置，△左右两侧的数字表示下几批待做试验的试验点数。

容易推出，若每批做$2k$个试验，共做n批，则应将试验范围等分，$L_n^{(2k)}=2(k+1)^n-1$份，第一批试验点是$L_{n-1}^{(2k)}$，$L_{n-1}^{(2k)}+1$，$2L_{n-1}^{(2k)}+1$，$2L_{n-1}^{(2k)}+2$，……$kL_{n-1}^{(2k)}+(k-1)$，$kL_{n-1}^{(2k)}+k$。试验结果的精确度是$l/L_n^{(2k)}$，$l=B-A$是含优区间的长度。

（2）每批做奇数个试验时，见表1-5，表1-6。

以表1-5每批做三个试验为例，说明该表的使用方法。若只

表 1-5 每批做三个试验

预给试验批数	含优区间等分份数	第一批试验点	图　示	精确度
一	4	1，2，3	△△△	$l/4$
二	10	4，5，9	○○○△△○○○△	$l/10$
三	28	10，14，24	9△3△9△3	$l/28$
四	76	28，38，66	27△9△27△9	$l/76$
n	$L_n^{(3)}$	$L_{n-1}^{(3)}, L_{n-1}^{(3)}+L_{n-2}^{(3)}$ $2L_{n-1}^{(3)}+L_{n-2}^{(3)}$	$L_{n-1}^{(3)}-1\triangle L_{n-2}^{(3)}-1$ $\triangle L_{n-1}^{(3)}-1\triangle L_{n-2}^{(3)}-1$	$l/L_n^{(3)}$

注：$L_n^{(3)}=\frac{1}{2}\left\{(1+\sqrt{3})^{n+1}+(1-\sqrt{3})^{n+1}\right\}$

表 1-6 每批做五个试验

预给试验批数	含优区间等分份数	第一批试验点	图　示	精确度
一	6	1，2，3，4，5	△△△△△	$l/6$
二	21	6，7，13，14，20	○○○○○△△○○○○○△△○○○○○△	$l/21$
三	81	21，27，48，54，75	20△5△20△5△20△5	$l/81$
四	306	81，102，183，204，285	80△20△80△20△80△20	$l/306$
n	$L_n^{(5)}$	$L_{n-1}^{(5)}$，$L_{n-1}^{(5)}+L_{n-2}^{(5)}$，$2L_{n-1}^{(5)}+L_{n-2}^{(5)}$，$2L_{n-1}^{(5)}+2L_{n-2}^{(5)}$，$3L_{n-1}^{(5)}+2L_{n-2}^{(5)}$	$L_{n-1}^{(5)}-1\triangle L_{n-2}^{(5)}-1\triangle L_{n-1}^{(5)}-1\triangle L_{n-2}^{(5)}-1\triangle L_{n-1}^{(5)}-1\triangle L_{n-2}^{(5)}-1$	$l/L_n^{(5)}$

注：$L_n^{(5)}=\frac{1}{2\sqrt{21}}\left\{(9+\sqrt{21})\left(\frac{3+\sqrt{21}}{2}\right)^n-(9-\sqrt{21})\left(\frac{3-\sqrt{21}}{2}\right)^n\right\}$

做一批试验，把含优区间均分为四等份，分别在三个分点1，2，3上试验，如图1-6所示。

若只做两批试验，分含优区间为10等份，第一批在4，5，9