实例分析

为提高山楂原料的利用率，
研究酶法液化工艺制造山楂原汁，
拟通过正交试验来寻找酶法液化的最佳工艺条件。

- 液化率为指标
- **4**因素**3**水平

表**1** 山楂原料试验因素水平表

水平	试验因素			
	加水量（mL/100g）A	加酶量（mL/100g）B	酶解温度（℃）C	酶解时间（h）D
1	**10**	**1**	**20**	**1.5**
2	**50**	**4**	**35**	**2.5**
3	**90**	**7**	**50**	**3.5**

表2 山楂原料试验方案及结果分析

试验号	因素			
	A(果肉加水量)	B(加酶量)	C(温度)	D(时间)
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3	1	3	2
8	3	2	1	3
9	3	3	2	1

4因素3水平

$$L_9(3^4)$$

- 极差分析法
- 方差分析法

表2 山楂原料试验方案及结果分析

试验号	因素				液化率%
	A（果肉加水量）	B（加酶量）	C（温度）	D（时间）	
1	1	1	1	1	0
2	1	2	2	2	17
3	1	3	3	3	24
4	2	1	2	3	12
5	2	2	3	1	47
6	2	3	1	2	28
7	3	1	3	2	1
8	3	2	1	3	18
9	3	3	2	1	42
K_1	41	13	46	89	
K_2	87	82	71	46	
K_3	61	94	72	54	
k_1	13.7	4.3	15.3	29.7	
k_2	29.0	27.3	23.7	15.3	
k_3	20.3	31.3	24.0	18.0	
极差R	15.3	27.0	8.7	14.3	
主次顺序	B>A>D>C				
优水平	A_2	B_3	C_3	D_1	
优组合	$A_2B_3C_3D_1$				

方差分析

表2 山楂原料试验方案及结果分析

➢能够利用较少的处理安排较多的试验因素，以获得较佳的试验结果，并对试验结果进行科学分析从而得到最佳试验条件。

➢优方案只能限制在已定的水平上，而不一定是试验范围内的最优方案。

➢不能在一定的试验范围内根据数据去确定变量间的相关关系及相应的回归方程，即试验结果无法用一个经验公式来表达，因此不便于考察试验条件改变后，试验指标将作如何变化。

试验号	因素				液化率%
3	1	3	3	3	24
4	2	1	2	3	12
7	3	1	3	2	1
K_3	61	94	72	54	
k_1	13.7	4.3	15.3	29.7	
k_2	29.0	27.3	23.7	15.3	
k_3	20.3	31.3	24.0	18.0	
极差R	15.3	27.0	8.7	14.3	
主次顺序	B>A>D>C				
优水平	A_2	B_3	C_3	D_1	
优组合	$A_2B_3C_3D_1$				

表2 山楂原料试验方案及结果分析

试验号	因素				液化率%
	A(果肉加水量)	B(加酶量)	C(温度)	D(时间)	
1	1	1	1	1	0
2	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
3	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
4	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
5	2	2	3	1	47
6	2	3	1	2	28
7	[illegible]	[illegible]	[illegible]	2	1
8	[illegible]	[illegible]	[illegible]	3	18
9	[illegible]	[illegible]	[illegible]	1	42
K_1	[illegible]	[illegible]	[illegible]	89	
K_2	87	82	71	46	
K_3	[illegible]	94	[illegible]	[illegible]	
k_1	[illegible]	[illegible]	[illegible]	[illegible]	
k_2	29.0	27.3	23.7	15.3	
k_3	20.3	31.3	24.0	[illegible]	
极差R	15.3	27.0	8.7	[illegible]	
主次顺序	B>A>D>C				
优水平	A_2	B_3	C_3	D_1	
优组合	$A_2B_3C_3D_1$				

***Y*与*A*、*B*、*C*、*D*各因素之间的相互关系？**

$$Y \to Z_A \quad Z_B \quad Z_C \quad Z_D$$

回归分析

液化率

果肉加水量
（10，50，90mL/100mg）

$$\hat{Y} = a + bx$$

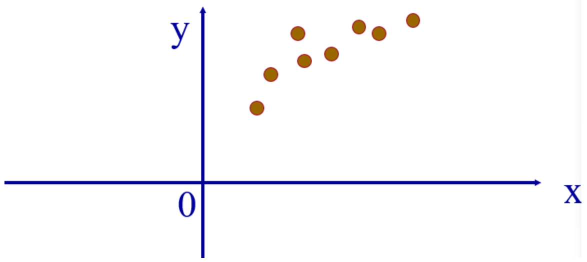
两个变量间存在协变关系，

对应于一个x, 就有一个相应的y值。

$$(x_1, y_1) \quad (x_2, y_2) \quad (x_i, y_i) \quad (x_n, y_n)$$

散点图

(scatter diagram)



- 两个变量间关系的**性质**（正向、负向协同变化）和**程度**（关系是否密切）；
- 两个变量间关系的**类型**（直线型、曲线型）；
- 是否有**异常观测值**的干扰。

定量关系？

【摘要】随着电子商务的蓬勃发展，淘宝、天猫的双十一创造了一个又一个的奇迹，为了找到淘宝双十一销售额数据背后隐含的意义，本文根据“华商情报网”上搜集的数据，梳理近十年淘宝“双十一”的交易额，建立数据分析模型，利用STATA软件，拟合年份与交易额之间的关系，并对2019年双十一销售额进行预测。从而发现数字背后存在的问题及引发的思考，为商家和消费者提供建议。
【关键词】双十一　淘宝　线性回归模型　假设与检验　消费者

1 背景

随着互联网、物联网等技术的迅猛发展，数据越来越为现代企业尤其是互联网企业所重视，大数据时代每一个企业都面临着机遇与挑战；在互联网经济的发展进程中，以淘宝、京东为代表B2C电子商务模式展现出了惊人的活力，电子商务的发展有力的推动了国家经济发展，同时电子商务本身具有交易过程过程便捷、商品品类齐全，商品售价相对低廉等特点为人们的日常生活带来了许多的便利。

淘宝的出现改变了传统的交易模式及交易习惯，淘宝双十一购物狂欢节，指的是淘宝自2009年11月11日起，每年11月11日都会例行举办的平台全品类商品促销日，并逐步发展为整个电商行业的购物狂欢节，双十一的第一年参与活动的商家较少活动力度有限，但是当次活动的营业额远远超过预期，因此，每年的11月11日，成为了淘宝天猫商城大规模促销活动的固定日期。“双十一”由此双十一成为了整个行业的狂欢盛典。自2009年淘宝推出第一次“双十一”购物大潮，截至2018年第十个“双十一”落下帷幕，淘宝创造了一个又一个销售奇迹，“双十一”已逐渐成为中国互联网甚至全球电子商务行业领域内最大规模的狂欢活动。截至2018年11月12日，第十个天猫双十一全球狂欢节落下帷幕，全天交易额达2135亿元。就交易规模来看，淘宝天猫商城“双十一”的交易数据呈爆发式增长，交易记录逐年被刷新。巨大的交易规模反映出了这场由淘宝、天猫所引导的全民购物狂欢引领了新的商业格局，推动了中国零售业的发展。

因此，本文从“双十一”活动交易总额出发，在“华商情报网”上收集近十年的交易数据[1]，梳理了2009－2018年淘宝“双十一”的销售数据，建立统计回归模型，借助STATA软件的数据分析功能，拟合出了2009年－2018年累计十年中淘宝“双十一”交易额与年份关系，通过得出的结论分析数据背后所隐含的问题，同时，对2019年“双十一”交易额进行了预测。本文分析了“双十一”交易额与年份的关系和数据中隐含的问题和猜想，给电商及消费者提出了相关的建议。

2 数据假设与检验

2009年11月11－2018年11月11日阿里“双十一”成交额数据分析

表1 2009-2018年双十一数据

年份	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
交易总额（亿元）	0.52	9.36	52	191	350	571	912	1207	1682	2135

2.1 分析与假设

本研究旨在通过对淘宝双十一的交易数据与年份的关系，发现淘宝天猫双十一交易额存在的规律，从而对2019年淘宝天猫双十一交易数据进行预测，因此，选取了从2009年到2018年天猫“双十一”成交额进行了统计分析，建立了统计回归模型，记交易额为Y，年份为X。

因变量（被解释变量）：Y：2009—2018年每年的交易额

自变量（解释变量）：X：每年的年份

建立假设条件：$\begin{cases} H_0 = 0{:}\beta = 0 \\ H_0 \neq 0{:}\beta = 0 \end{cases}$ 2.2 建立模型

为了描绘出交易额Y与年份X的关系，我们利用表格1的数据集分别作出Y对X的散点图

scatter y x:

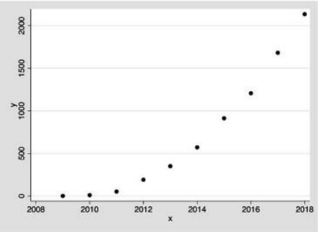


图1 交易额与年份的散点图

从以上的散点图中我们可以看到，随着时间的不断向后推移，销售额随着年份的增长呈现明显的增长趋势，并且销售额随年份的增长呈类似二次函数的趋势。因此，用二次函数来描述模型：$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon$

等式右端x为回归变量，$\beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon$是给定年份x时，交易额y的值，参数$\beta_0$，$\beta_1$，$\beta_2$为回归系数，根据表1的数据估计，影响y的其他因素作用都包含在随机误差ε中，这里的ε大致服从均值为0的正态分布。

2.3 模型求解

利用STATA统计工具的命令regress求解，定义时间变量x，输出的结果如下表所示：

表2 模型计算结果

参数	参数估计估计值	参数置信区间	标准差
β_0	1.22	[1.13, 1.30]	1.30
β_1	30.09	[27.90, 32.27]	3719.483
β_2	-120936.8	[-129732, 112141.6]	3744584
$R^2 = 0.9994$　$F = 5667.32$　$p = 0.0000$			

2.4 结果分析

根据表2可得，拟合优度$R^2 = 0.9994$　指交易额的99.94%可由模型确定，F值远远超过检验的临界值，P<0.0005，模型整体通过检验。

表2中得出了的β_0,β_1,β_2估计值，即$\widehat{\beta}_0 = 1.22$，$\widehat{\beta}_1 = -120936.8$，$\widehat{\beta}_2 = 30.09$，即$x^2$的系数为-120936.8，x的系数为30.09，常数项为1.22，检查它们的置信区间，回归变量X^2（对因变量y的影响）是显著的。

2.5成交额预测

将回归系数的值代入模型内，即可预测2019年以后的“双十一”交易额，预测值记为$\hat{y}$，得到预测模型的方程为$y = \widehat{\beta}_0 + \widehat{\beta}1x^2 + \widehat{\beta}_2 x$。只需知道年份，就可以计算预测值。据此预测出2019年阿里“双十一”的成交额为2675.568亿元。

3 狂欢背后的引发的思考

$$\hat{y} = a + bx$$

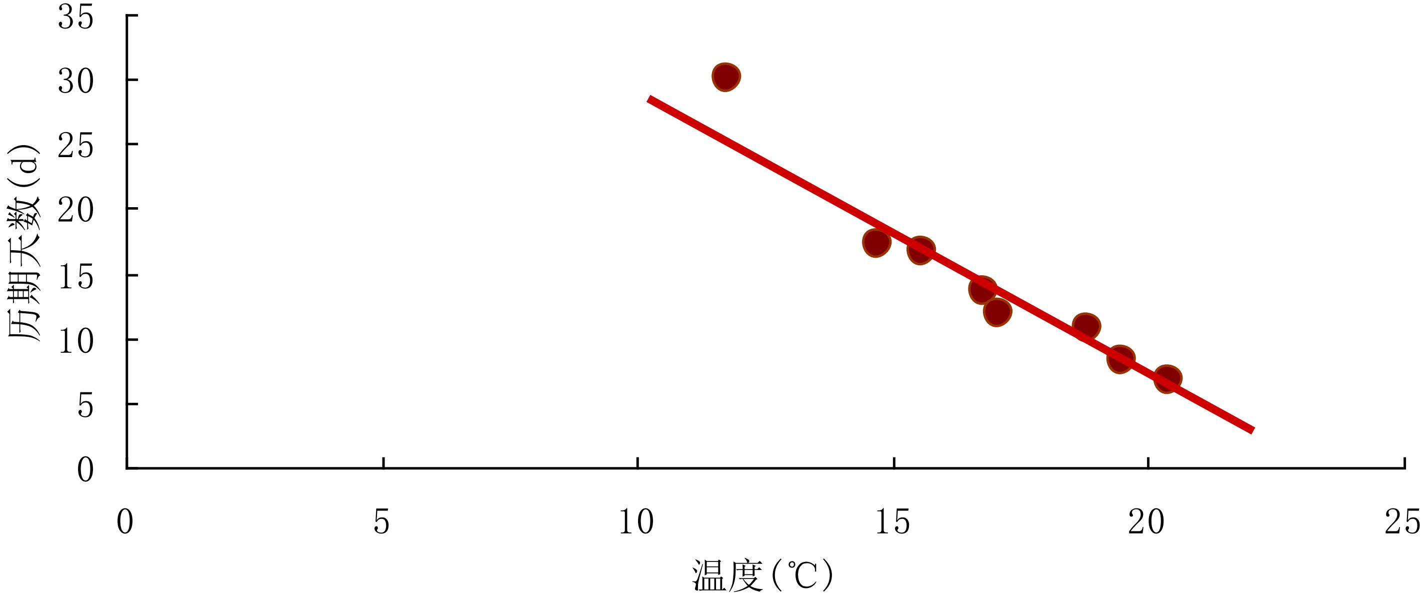
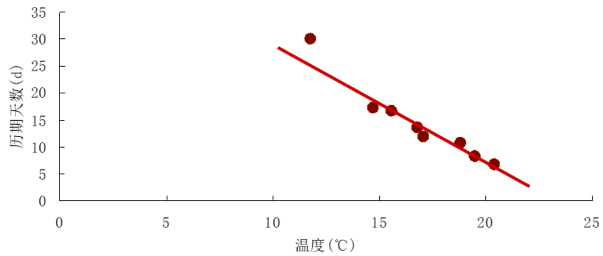


图1 黏虫孵化历期平均温度与历期天数关系图



自变量

$$\hat{y}=a+bx$$

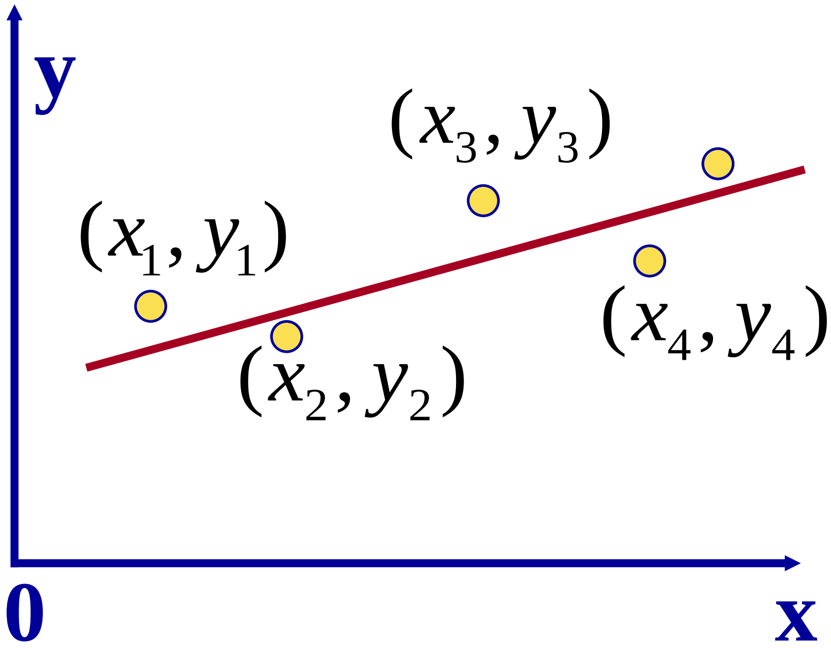
斜率(slope)

回归系数(regression coefficient)

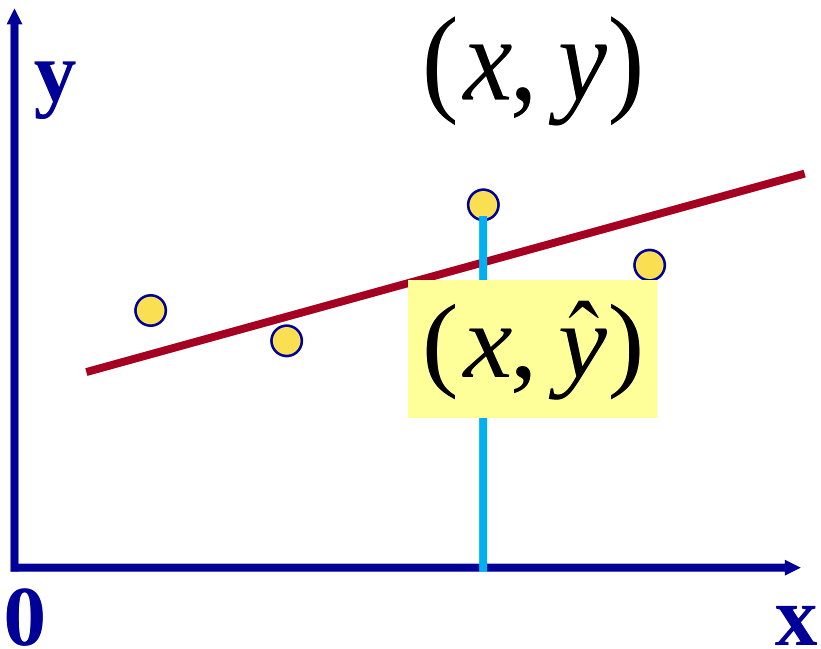
回归截距(regression intercept)

与x值相对应的y的平均值的点估计值

$$\hat{y} = a + bx \Longrightarrow \boxed{y}$$



$$\hat{y} = a + bx \Longrightarrow \boxed{\mathbf{y}}$$



$$y - \hat{y}$$

$$\sum (y - \hat{y}) = 0$$

$$\sum (y - \hat{y})^2$$

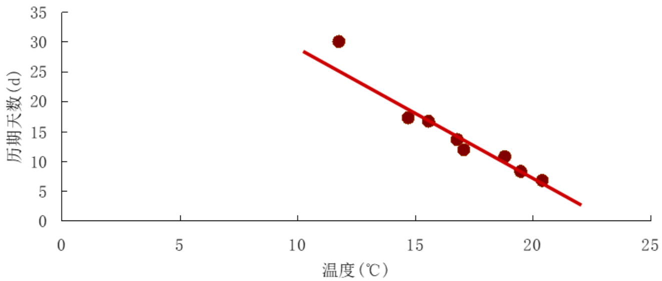
最小

$$\hat{y} = a + bx \Longrightarrow \boxed{y}$$

$$\sum_{1}^{n}(y-\hat{y})^{2} = 最小$$

$$\sum_{1}^{n}(y-a-bx)^{2}$$

Q

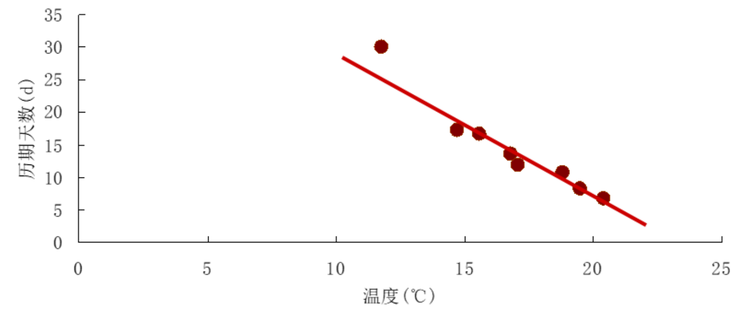


$$Q=\sum_{1}^{n}(y-\hat{y})^2=\sum_{1}^{n}(y-a-bx)^2$$

最小

$$\begin{cases}\boxed{\dfrac{\partial Q}{\partial a}=0}\\[2ex]\boxed{\dfrac{\partial Q}{\partial b}=0}\end{cases}$$

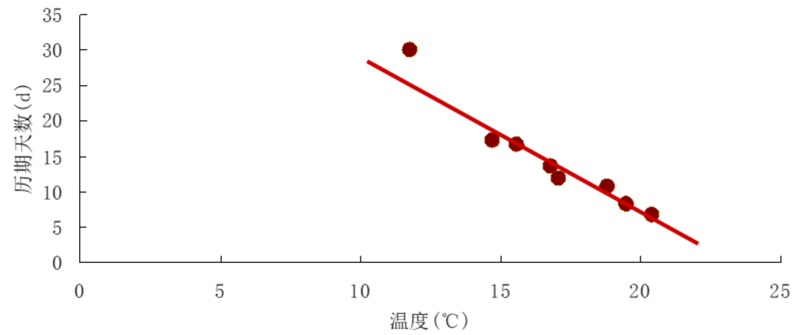
$$a = \bar{y} - b\bar{x}$$



$$b = \frac{\sum xy - (\sum x)(\sum y)/n}{\sum x^2 - (\sum x)^2/n}$$

$$= \frac{\sum (x-\bar{x})(y-\bar{y})}{\sum (x-\bar{x})^2} = \frac{SP}{SS_x}$$

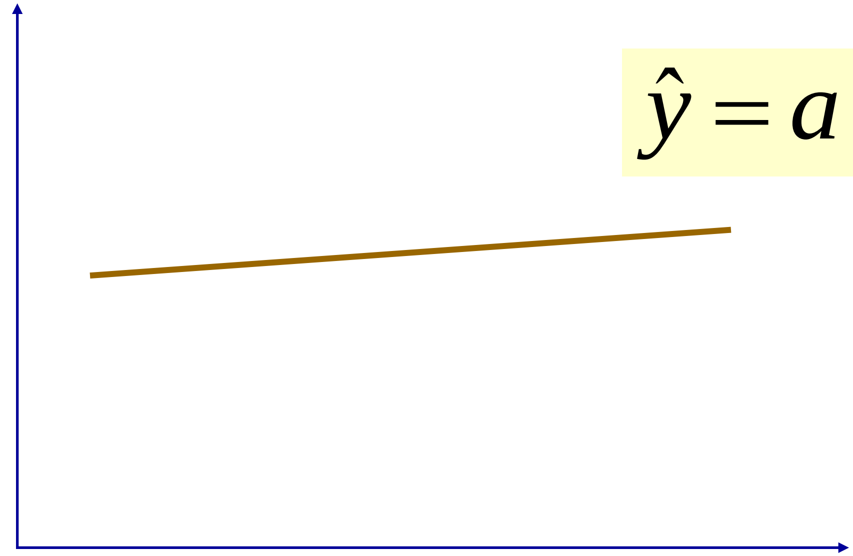
直线回归方程的一般形式：



$$\hat{y} = a + bx$$

$$a = \overline{y} - b\overline{x}$$

$$\hat{y} = \overline{y} + b(x - \overline{x})$$



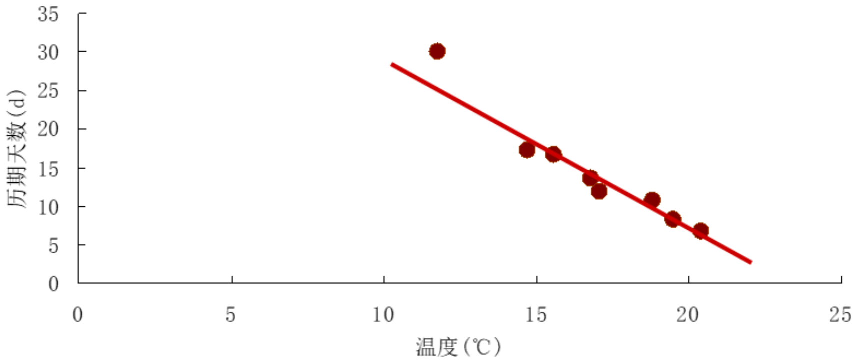
$$\hat{y} = a + bx$$

若x 和y变量间的回归系数b非常小，也能列出回归方程。

但该方程是否有意义？

假设检验

或者说，x 和y间的回归关系是否显著？



$$\hat{y} = a + bx$$

$$y = \alpha + \beta x + \varepsilon$$

$$\hat{y} = a + bx$$

$$y = \alpha + \beta x + \varepsilon$$

$$b = \frac{SP}{SS_x}$$

$$\beta = 0$$

(x_1,y_1) (x_2,y_2) (x_3,y_3) (x_n,y_n)

假设检验

回归方程的显著性

回归系数的显著性

F-检验

t-检验

表 试验结果分析

试验号	因素				液化率%
	A(果肉加水量)	B(加酶量)	C(温度)	D(时间)	
1	1	1	1	1	0
2	1	2	2	2	17
3	1	3	3	3	24
4	2	1	2	3	12
5	2	2	3	1	47
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	28
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1
8	3	2	1	3	18
[illegible]	[illegible]	[illegible]	2	1	42
[illegible]	[illegible]	[illegible]	46	89	
K_2	87	82	71	46	
K_3	61	94	72	54	
k_1	13.7	4.3	15.3	29.7	
k_2	[illegible]	[illegible]	23.7	15.3	
k_3	20.3	31.3	24.0	18.0	
极差R	15.3	27.0	8.7	14.3	
主次顺序	B>A>D>C				
优水平	A_2	B_3	C_3	D_1	
优组合	$A_2B_3C_3D_1$				

液化率→果肉加水量，加酶量，酶解温度， 酶解时间

$$\hat{Y}=b_0+b_1Z_A+b_2Z_B+b_3Z_C+b_4Z_D$$

$$1+4+C_4^2=11$$

互作?

回归分析

$$\hat{Y} = b_0 + b_1 Z_A + b_2 Z_B + b_3 Z_C + b_4 Z_D$$

由试验数据建立回归方程，描述自变量与因变量之间的函数关系，并对试验结果进行预测和优化。

- 不关心试验数据如何取得，只能被动地去处理由试验所得到的数据，而对试验的设计安排几乎不能提出任何要求。
- 盲目增加了试验次数，而且由数据所分析出的结果有时不能提供充分的信息。
- 回归系数之间如果存在相关关系，若从回归方程中剔除某个不显著因素时，需要重新计算回归系数。

正交设计 + 回归分析

回归正交设计

◆利用较少的处理安排较多的试验因素，获得较佳的试验结果；

◆不能在一定的试验范围内根据数据样本去确定变量间的相关关系及回归方程。

正交设计，回归分析

将设计方案与数据的回归分析结合起来考虑，产生于20世纪50年代初。

◆根据试验目的和数据分析来选择试验点，使得在每个试验点上获得的数据含有最大的信息，从而减少试验次数；

◆各自变量（因素）向量间满足正交性以便于回归分析；

◆用回归分析处理试验数据，将试验指标与被考察的各因素间的关系以回归方程表示出来。

回归设计

- 也称为响应曲面设计，就是在因子空间选择适当的试验点，以较少的试验处理，建立一个有效的回归方程。
- 它是在多元线性回归的基础上用主动收集数据的方法获得具有较好性质的回归方程的一种试验设计方法。

回归正交设计(Orthogonal regression design)

回归设计中最简单、最基本、最常用和最有代表性的设计方法，是回归分析与正交设计有机结合而形成的一种新的试验设计方法。

回归正交设计

- ◆利用正交试验设计中“正交性”的特点，有计划、有目的、科学合理地在因素的试验范围内选择适当的试验点，
- ◆用较少的试验实施，
- ◆建立回归方程，
- ◆试验结果用一个明确的函数表达式（即回归方程）表示，
- ◆既减少试验次数，又能快速建立回归方程。

回归正交设计

是正交试验设计与回归分析结合而形成的一种设计方法。

实验设计与统计分析

- 概述
- 对比、随机设计
- 区组设计
- 拉丁方设计
- 裂区设计
- 正交设计
- 回归正交设计
  - 一次回归正交设计
  - 二次回归正交设计
- 均匀设计

（一）定义

一次回归正交设计（**Orthogonal design by linear regression**）**：**

当研究的因变量与自变量之间呈线性关系时，可采用一次回归正交设计。

是利用回归正交设计原理建立依变量y关于m个自变量Z_1、Z_2、…、Z_m的一次回归方程。

$$\hat{Y} = b_0 + b_1 Z_1 + b_2 Z_2 + \cdots + b_m Z_m$$

$$\hat{Y} = b_0 + \sum_{j=1}^{m} b_j Z_j + \sum_{i<j} b_{ij} Z_i Z_j$$

$$\hat{Y} = b_0 + b_1 Z_1 + b_2 Z_2 + \cdots + b_m Z_m$$

$$\hat{Y} = b_0 + \sum_{j=1}^{m} b_j Z_j + \sum_{i<j} b_{ij} Z_i Z_j$$

- 古典回归分析中的多元回归计算过程中，回归计算的复杂性在于系数矩阵的运算。
- 这是因为古典回归分析中的试验点是随意的，试验点上变量x的取值是随意的，因而所构成的系数矩阵是十分复杂的。

$$\hat{Y} = b_0 + b_1 Z_1 + b_2 Z_2 + \cdots + b_m Z_m$$

$$\hat{Y} = b_0 + \sum_{j=1}^{m} b_j Z_j + \sum_{i<j} b_{ij} Z_i Z_j$$

◆根据多元线性回归的理论，用矩阵表示，以最小二乘法，可以推导出系数矩阵为

$$b = (x^T x)^{-1} x^T y$$

$$\hat{Y} = b_0 + \sum_{j=1}^{m} b_j Z_j + \sum_{i<j} b_{ij} Z_i Z_j$$

结构矩阵

$$X = \begin{bmatrix} 1 & Z_{11} & Z_{12} & \cdots & Z_{1m} & Z_{11}Z_{12} & \cdots & Z_{1,m-1}Z_{1m} \\ 1 & Z_{21} & Z_{22} & \cdots & Z_{2m} & Z_{21}Z_{22} & \cdots & Z_{2,m-1}Z_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & Z_{N1} & Z_{N2} & \cdots & Z_{Nm} & Z_{N1}Z_{N2} & \cdots & Z_{N,m-1}Z_{Nm} \end{bmatrix}$$

◆结构矩阵中的元素除第一列外，其余都是变量x在各试验点上的取值。

正规方程组的系数矩阵A

$$
A = x^T x = \begin{bmatrix}
N & \sum Z_{i1} & \sum Z_{i2} & \cdots & \sum Z_{im} & \sum Z_{i1}Z_{i2} & \sum Z_{i1}Z_{i3} & \cdots & \sum Z_{i,m-1}Z_{im} \\
 & \sum Z_{i1}^2 & \sum Z_{i1}Z_{i2} & \cdots & \sum Z_{i1}Z_{im} & \sum Z_{i1}^2 Z_{i2} & \sum Z_{i1}^2 Z_{i3} & \cdots & \sum Z_{i1}Z_{i,m-1}Z_{im} \\
 & & \sum Z_{i2}^2 & \cdots & \sum Z_{i2}x_{im} & \sum Z_{i1}Z_{i2}^2 & \sum Z_{i1}Z_{i2}Z_{i3} & \cdots & \sum Z_{i2}Z_{i,m-1}Z_{im} \\
 & & & & \vdots & \vdots & \vdots & \cdots & \vdots \\
 & & & & \sum Z_{im}^2 & \sum Z_{i1}Z_{i2}Z_{im} & \sum Z_{i1}Z_{i3}Z_{im} & \cdots & \sum Z_{i,m-1}Z_{im}^2 \\
 & \text{对} & & & & \sum (Z_{i1}Z_{i2})^2 & \sum Z_{i1}Z_{i1}Z_{i2}Z_{im} & \cdots & \sum Z_{i1}Z_{i2}Z_{i,m-1}Z_{im} \\
 & & & & & & \sum (Z_{i1}Z_{i3})^2 & \cdots & \sum Z_{i1}Z_{i3}Z_{i,m-1}Z_{im} \\
 & & & \text{称} & & & & & \vdots \\
 & & & & & & & & \sum (Z_{i,m-1}Z_{im})^2
\end{bmatrix}
$$

- 系数矩阵各元素的值决定于结构矩阵各元素的值。

正规方程组的系数矩阵A

$$
A = x^T x = \begin{bmatrix}
N & \sum Z_{i1} & \sum Z_{i2} & \cdots & \sum Z_{im} & \sum Z_{i1}Z_{i2} & \sum Z_{i1}Z_{i3} & \cdots & \sum Z_{i,m-1}Z_{im} \\
 & \sum Z_{i1}^2 & \sum Z_{i1}Z_{i2} & \cdots & \sum Z_{i1}Z_{im} & \sum Z_{i1}^2 Z_{i2} & \sum Z_{i1}^2 Z_{i3} & \cdots & \sum Z_{i1}Z_{i,m-1}Z_{im} \\
 & & \sum Z_{i2}^2 & \cdots & \sum Z_{i2}x_{im} & \sum Z_{i1}Z_{i2}^2 & \sum Z_{i1}Z_{i2}Z_{i3} & \cdots & \sum Z_{i2}Z_{i,m-1}Z_{im} \\
 & & & & \vdots & \vdots & \vdots & \cdots & \vdots \\
 & & & & \sum Z_{im}^2 & \sum Z_{i1}Z_{i2}Z_{im} & \sum Z_{i1}Z_{i3}Z_{im} & \cdots & \sum Z_{i,m-1}Z_{im}^2 \\
 & & & & & \sum (Z_{i1}Z_{i2})^2 & \sum Z_{i1}Z_{i1}Z_{i2}Z_{im} & \cdots & \sum Z_{i1}Z_{i2}Z_{i,m-1}Z_{im} \\
 & \text{对} & & & & & \sum (Z_{i1}Z_{i3})^2 & \cdots & \sum Z_{i1}Z_{i3}Z_{i,m-1}Z_{im} \\
 & & & \text{称} & & & & & \vdots \\
 & & & & & & & & \sum (Z_{i,m-1}Z_{im})^2
\end{bmatrix}
$$

- 线性代数知识可知，系数矩阵如为对角阵时，其逆矩阵便于计算。
- 因此，如果能经过某种安排适当的试验点，使系数矩阵A为对角矩阵，不仅能简化其逆矩阵计算，而且还使得回归系数间不存在相关性。

$$A = x^T x = \begin{bmatrix} N & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ & \sum Z_{i1}^2 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ & & \sum Z_{i2}^2 & \dots & 0 & 0 & 0 & \dots & 0 \\ & & & & \vdots & \vdots & \vdots & \dots & \vdots \\ & & & & \sum Z_{im}^2 & 0 & 0 & \dots & 0 \\ & & & & & \sum (Z_{i1} Z_{i2})^2 & 0 & \dots & 0 \\ & & & & & & \sum (Z_{i1} Z_{i3})^2 & \dots & 0 \\ & & & & & & & & \vdots \\ & & & & & & & & \sum (Z_{i,m-1} Z_{im})^2 \end{bmatrix}$$

对角矩阵

$$x = \begin{bmatrix} 1 & Z_{11} & Z_{12} & \dots & Z_{1m} & Z_{11} Z_{12} & \cdots & Z_{1,m-1} Z_{1m} \\ 1 & Z_{21} & Z_{22} & \cdots & Z_{2m} & Z_{21} Z_{22} & \cdots & Z_{2,m-1} Z_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & Z_{N1} & Z_{N2} & \dots & Z_{Nm} & Z_{N1} Z_{N2} & \cdots & Z_{N,m-1} Z_{Nm} \end{bmatrix}$$

✓任一列的和为零

✓任两列的相应元素乘积之和为零

从数学意义上讲，也就是要使结构矩阵x具有正交性。

$L_4(2^3)$

试验号	1	2	3
1	1	1	1
2	1	2	2
3	2	1	2
4	2	2	1

$$-1 \to 1, 1 \to 2$$

表7-1 回归正交表

试验号	1	2	3
1	-1	-1	-1
2	-1	1	1
3	1	-1	1
4	1	1	-1

- 两种正交表无本质差别。
- 代换后的正交表具有正交性。

✓任一列的和为零

✓任两列的相应元素乘积之和为零

结构矩阵

$$X=\begin{bmatrix} 1 & Z_{11} & Z_{12} & \cdots & Z_{1m} & Z_{11}Z_{12} & \cdots & Z_{1,m-1}Z_{1m} \\ 1 & Z_{21} & Z_{22} & \cdots & Z_{2m} & Z_{21}Z_{22} & \cdots & Z_{2,m-1}Z_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & Z_{N1} & Z_{N2} & \cdots & Z_{Nm} & Z_{N1}Z_{N2} & \cdots & Z_{N,m-1}Z_{Nm} \end{bmatrix}$$

- 要使结构矩具有正交性，要以2水平正交表来安排试验。
- 一次回归正交设计就是应用2水平正交表来安排试验的。

$L_4(2^3)$　　$L_8(2^7)$　　$L_{12}(2^{11})$

（二）步骤

步骤

1 确定试验因素及其下水平和上水平

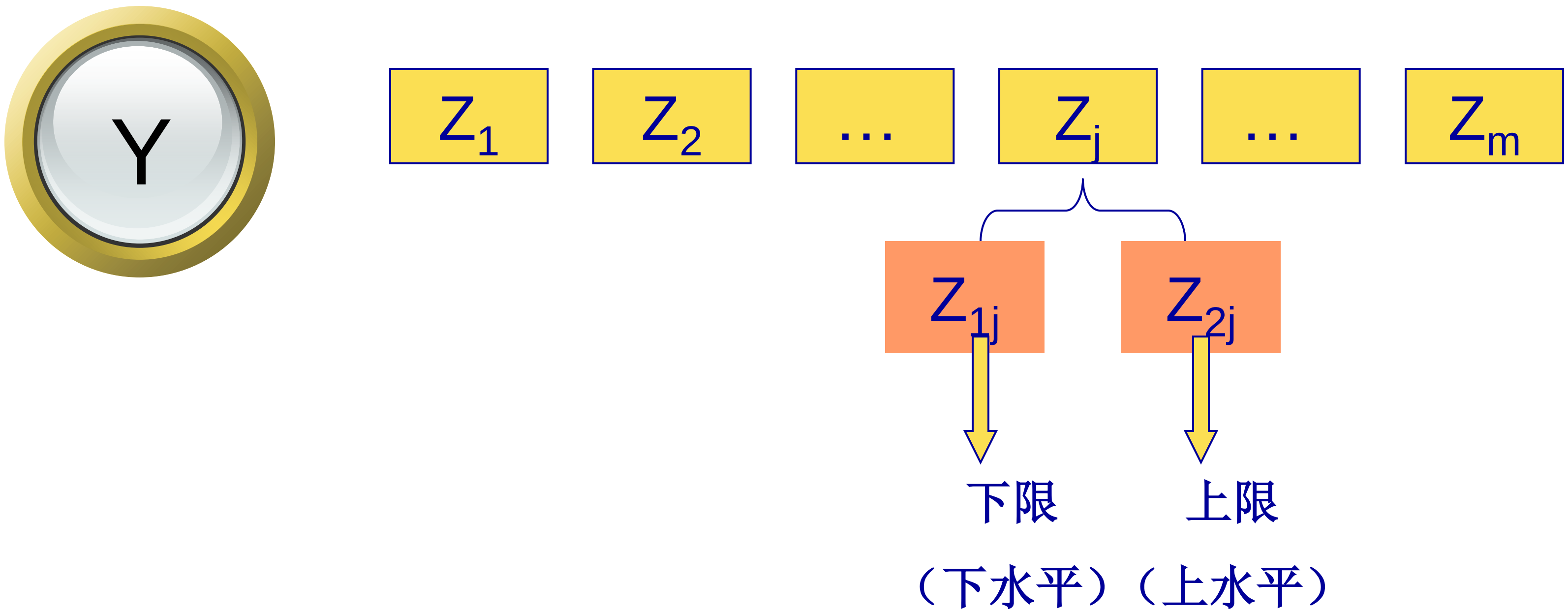
2 因素水平进行编码

3 选择适当的正交表，列出编码因素的试验方案

4 建立回归方程

（二）步骤

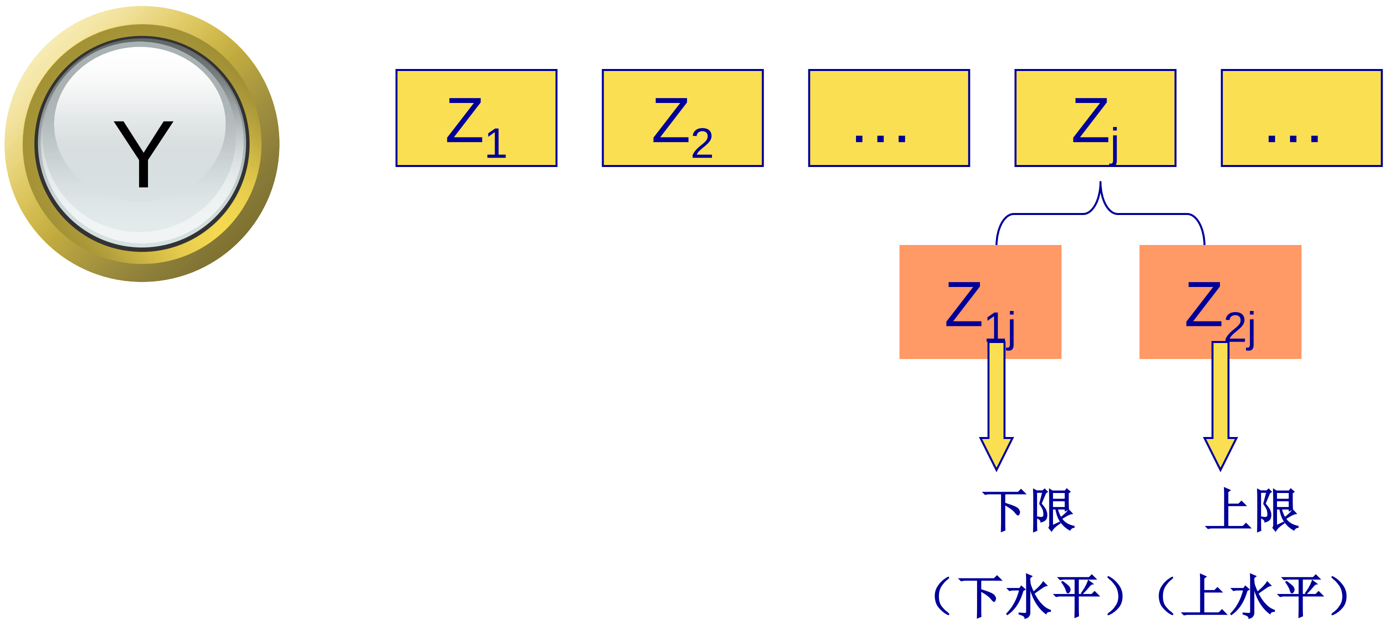
1. 确定试验因素及其下水平和上水平



- 根据试验指标y，选择要考察的因素数m;
- 确定每个因素的取值范围。

（二）步骤

1. 确定试验因素及其下水平和上水平



◆确定因素水平的上限和下限要根据专业的知识或预试验，

◆通常上下限距离越小，越接近最佳水平范围，试验求得的的回归方程的预测性就越好。

（二）步骤

1. 确定试验因素及其下水平和上水平

Y

Z_1 Z_2 … Z_j … Z_m

Z_{1j} Z_{2j}

下限（下水平） 上限（上水平）

$$Z_{oj}=\frac{Z_{1j}+Z_{2j}}{2}$$

因子Z_j的零水平

$$\Delta_j=\frac{Z_{2j}-Z_{1j}}{2}$$

因子Z_j的变化区间

实例分析1

一次回归正交设计

为了研究某作物的栽培技术，选择影响作物产量的3个主要因素：

水分状况 (全生育期土壤湿度占田间持水量的百分比)、

追施氮肥量、

密度，

试验指标为产量y（kg/小区）。进行一次回归正交设计并分析。

表7-2 作物栽培技术试验的因素水平表

水平	水分状况Z_1（%）	追氮量Z_2(kg/hm^2)	密度Z_3(万株/hm^2)
上水平	95	40	65
下水平	75	20	45

列出因素水平编码表

表7-3 作物栽培技术试验的因素水平编码表

名称	水分状况Z_1（%）	追氮量Z_2(kg/hm^2)	密度Z_3(万株/hm^2)
上水平	95	40	65
下水平	75	20	45
零水平			
变化区间			

2. 因素水平进行编码

编码变换，就是对各因素每个水平的取值作无量纲的线性变换。

Z_{aj}

x_{aj}

编码因子

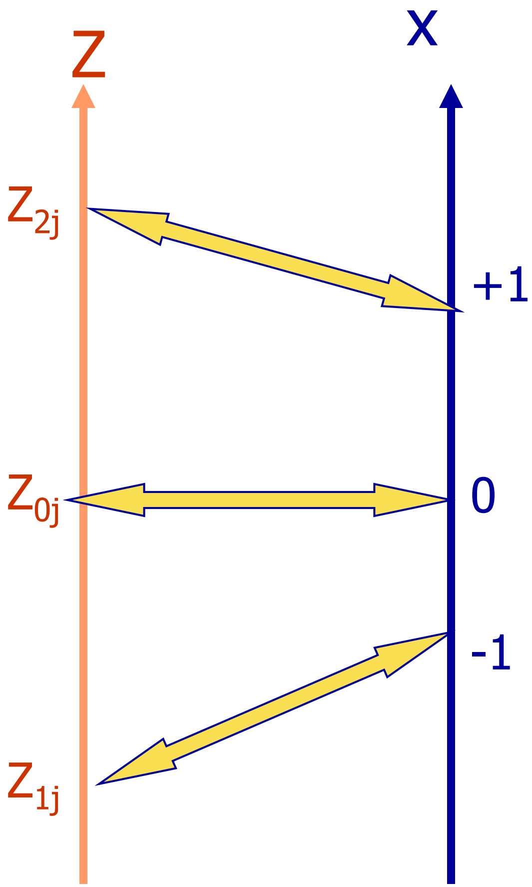
$$x_{aj}=\frac{Z_{aj}-Z_{0j}}{\Delta_j}(a=0,1,2;\quad j=1,2,\ldots,m)$$

水平	因子 Z_j	编码因子 x_j
下水平	Z_{1j}	-1
零水平	Z_{0j}	0
上水平	Z_{2j}	1

列出因素水平编码表

表7-3 作物栽培技术因素水平编码表

名称	编码xj	水分状况Z_1（%）	追氮量Z_2(kg/hm^2)	密度Z_3(万株/hm^2)
上水平		95	40	65
下水平		75	20	45
零水平		85	30	55
变化区间		10	10	10

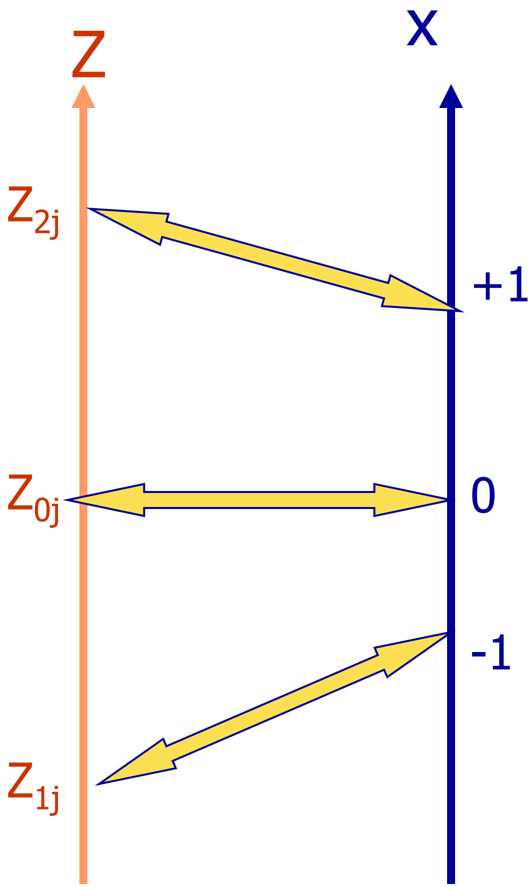


对因素各水平编码的目的，

□是为了使供试因素各水平在编码空间是“平等”的，

□即它们的取值都在**[-1,+1]**区间变化，

□而不受因素原单位和取值的影响。



- 对于各个自然因素编码后，

建立y对$Z_1,Z_2,\ldots,Z_m$的回归就转化为

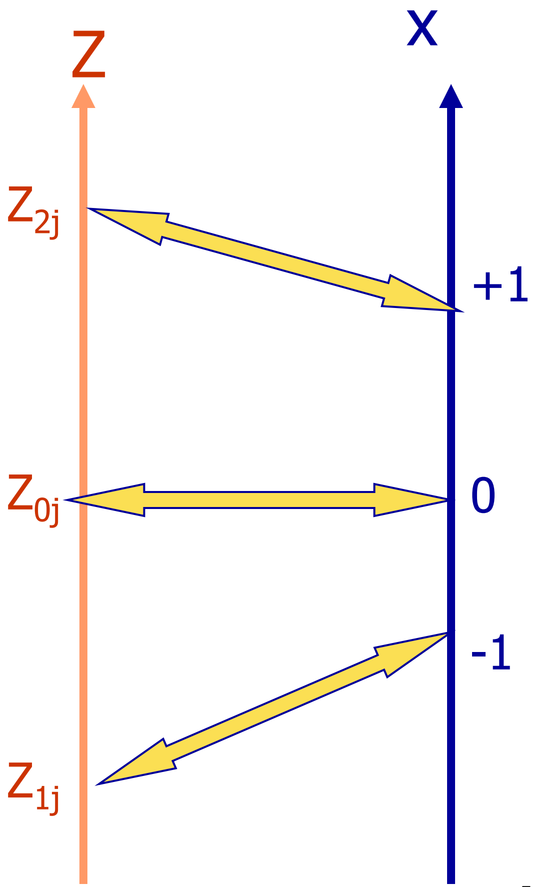
建立y对$x_1,x_2,\ldots,x_m$的回归问题。

- 在以$Z_1,Z_2,\ldots,Z_m$为坐标轴的因子空间中选择适当试验点的回归问题

就转化为以$x_1,x_2,\ldots,x_m$的坐标的编码空间中选择适当试验点的回归问题。

$$\hat{Y} = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_m x_m$$

$$\hat{Y} = b_0 + \sum_{j=1}^{m} b_j x_j + \sum_{i<j} b_{ij} x_i x_j$$



- 试验方案的优化、方程的回归及其统计检验，都相应转化为在编码空间中进行。

- 这样的设计简化了计算过程，运算更为简便。

$$\hat{Y} = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_m x_m$$

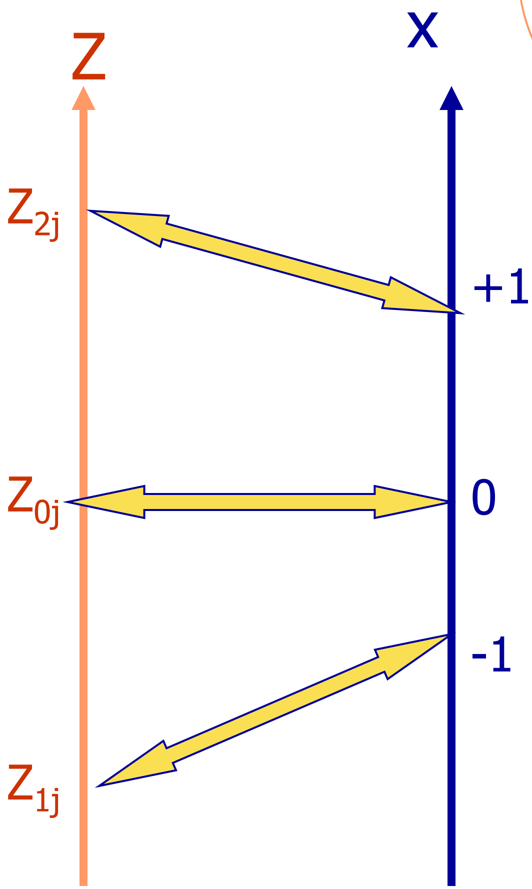
$$\hat{Y} = b_0 + \sum_{j=1}^{m} b_j x_j + \sum_{i<j} b_{ij} x_i x_j$$

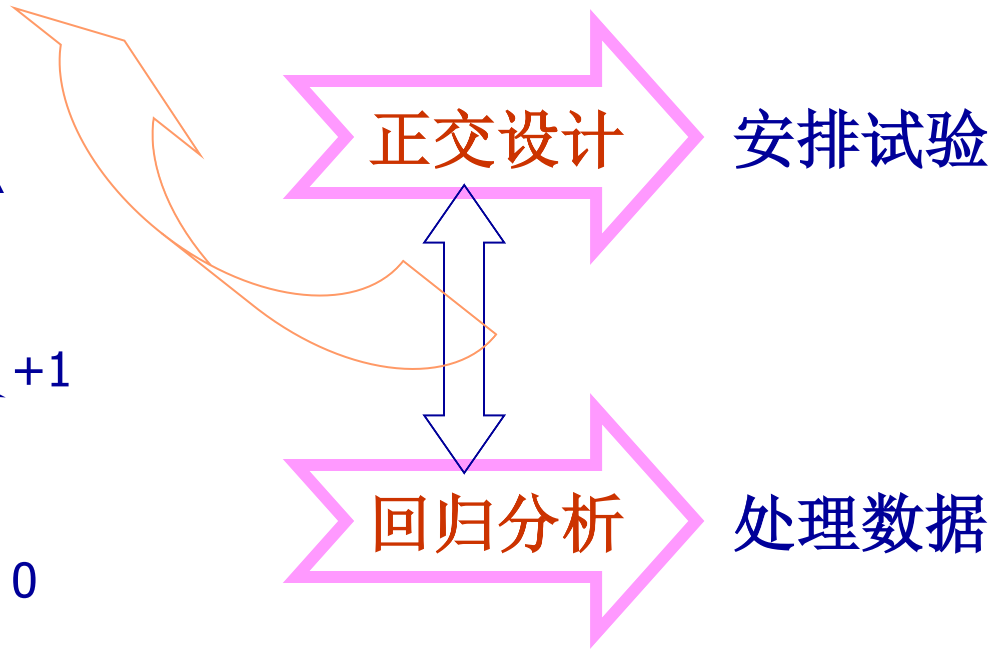
表 1　因素水平编码表

因素	Z_1	Z_2	„	Z_p
编码记号	X_1	X_2	„	X_p
下水平（-1）	Z_{11}	Z_{12}	„	Z_{1P}
零水平（0）	Z_{01}	Z_{02}	„	Z_{0P}
上水平（+1）	Z_{21}	Z_{22}	„	Z_{2P}
变化间距	$\triangle_1$	$\triangle_2$	„	$\triangle_P$

一次回归正交设计

因素水平进行编码





因此，编码是回归正交设计的关键环节，

也是回归正交试验设计与一般的正交设计的主要区别。

二、步骤

1 确定试验因素及其下水平和上水平

2 因素水平进行编码

3 选择适当的正交表，列出编码因素的试验方案

4 建立回归方程

二水平正交表

一次回归正交设计

一次回归正交设计中正交表的选择和方案设计同正交设计相类似。

- **(1)** 根据因素个数和交互作用选择适当的正交表，
- **(2)** 将各因素及交互作用分别安排到正交表的相应列，
- **(3)** 将各因素的每一水平真实值填入相应的编码中，
- **(4)** 得到一次回归正交设计的试验方案。

3个因素 $L_8(2^7)$

试验号	1	2	3	4	5	6	7
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2

以“1”代换正交表中的“2”，

以“-1”代换正交表中的“1”。

均衡分散

整齐可比

表7-4　三因素一次回归正交设计与实施方案

处理号	试验设计			实施方案		
	X_1	X_2	X_3	Z_1	Z_2	Z_3
1	-1	-1	-1	Z_{11}	Z_{12}	Z_{13}
2	-1	-1	1	Z_{11}	Z_{12}	Z_{23}
3	-1	1	-1	Z_{11}	Z_{22}	Z_{13}
4	-1	1	1	Z_{11}	Z_{22}	Z_{23}
5	1	-1	-1	Z_{21}	Z_{12}	Z_{13}
6	1	-1	1	Z_{21}	Z_{12}	Z_{23}
7	1	1	-1	Z_{21}	Z_{22}	Z_{13}
8	1	1	1	Z_{21}	Z_{22}	Z_{23}

表7-4 三因素一次回归正交设计与实施方案

处理号	试验设计			实施方案		
	X_1	X_2	X_3	Z_1	Z_2	Z_3
1	1	1	1	Z_{21}	Z_{22}	Z_{23}
2	1	1	-1	Z_{21}	Z_{12}	Z_{13}
3	1	-1	1	Z_{21}	Z_{12}	Z_{23}
4	1	-1	-1	Z_{21}	Z_{22}	Z_{13}
5	-1	1	1	Z_{11}	Z_{22}	Z_{23}
6	-1	1	-1	Z_{11}	Z_{22}	Z_{13}
7	-1	-1	1	Z_{11}	Z_{12}	Z_{23}
8	-1	-1	-1	Z_{11}	Z_{12}	Z_{13}

数字1和-1既表示因子的两个水平，又同时表示每个水平变化的数量大小。

- 1既表示因子的上水平也表示该上水平的取值是1，
- -1既表示因子的下水平也表示该下水平的取值是-1。

表7-4 三因素一次回归正交设计与实施方案

处理号	试验设计		
	X_1	X_2	X_3
1	1	1	1
2	1	1	-1
3	1	-1	1
4	1	-1	-1
5	-1	1	1
6	-1	1	-1
7	-1	-1	1
8	-1	-1	-1

$L_9(3^4)$

试验号	1	2	3	4
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3	1	3	2
8	3	2	1	3
9	3	3	2	1

$L_8(2^7)$

试验号	1	2	3	4	5	6	7
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2

正交表的正交性

（1）任一列中，各水平都出现，且出现的次数相等。

每列中数字的和等于零。

$$\sum x_j = 0$$

表7-4 三因素一次回归正交设计与实施方案

处理号	试验设计		
	X_1	X_2	X_3
1	1	1	1
2	1	1	-1
3	1	-1	1
4	1	-1	-1
5	-1	1	1
6	-1	1	-1
7	-1	-1	1
8	-1	-1	-1

$L_9(3^4)$

试验号	1	2	3	4
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3	1	3	2
8	3	2	1	3
9	3	3	2	1

$L_8(2^7)$

试验号	1	2	3	4	5	6	7
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2

正交表的正交性

（2）任两列之间各种不同水平的所有可能组合都出现，且出现的次数相等。

任意两列对应数字乘积和等于零；

$$\sum x_j x_i = 0$$

表7-4 三因素一次回归正交设计与实施方案

处理号	试验设计		
	X_1	X_2	X_3
1	1	1	1
2	1	1	-1
3	1	-1	1
4	1	-1	-1
5	-1	1	1
6	-1	1	-1
7	-1	-1	1
8	-1	-1	-1

- 由于以正交表来安排试验和对变量进行了线性代换，其系数矩阵的逆矩阵计算简单；
- 同时回归系数之间也不存在相关性；
- 一次回归方程的建立简单。

每列中数字平方和 =

表7-4 三因素一次回归正交设计与实施方案

处理号	试验设计			实施方案		
	X_1	X_2	X_3	Z_1	Z_2	Z_3
1	1	1	1	Z_{21}	Z_{22}	Z_{23}
2	1	1	-1	Z_{21}	Z_{12}	Z_{13}
3	1	-1	1	Z_{21}	Z_{12}	Z_{23}
4	1	-1	-1	Z_{21}	Z_{22}	Z_{13}
5	-1	1	1	Z_{11}	Z_{22}	Z_{23}
6	-1	1	-1	Z_{11}	Z_{22}	Z_{13}
7	-1	-1	1	Z_{11}	Z_{12}	Z_{23}
8	-1	-1	-1	Z_{11}	Z_{12}	Z_{13}

二、步骤

1. 确定试验因素及其下水平和上水平
2. 因素水平进行编码
3. 选择适当的正交表，列出编码因素的试验方案
4. 建立回归方程

列出结构矩阵

由于按正交表来安排试验和对变量进行了线性代换，

- 使得系数矩阵相关的计算过程简单化，
- 且回归系数之间不存在相关性。

表7-5 三因素一次回归正交设计结构矩阵与结果计算表

处理号		X_1	X_2	X_3		y
1		1	1	1		y_1
2		1	1	-1		y_2
3		1	-1	1		y_3
4		1	-1	-1		y_4
5		-1	1	1		y_5
6		-1	1	-1		y_6
7		-1	-1	1		y_7
8		-1	-1	-1		y_8

- 常数列

表7-5 三因素一次回归正交设计结构矩阵与结果计算表

处理号	X_0	X_1	X_2	X_3		y
1	1	1	1	1		y_1
2	1	1	1	-1		y_2
3	1	1	-1	1		y_3
4	1	1	-1	-1		y_4
5	1	-1	1	1		y_5
6	1	-1	1	-1		y_6
7	1	-1	-1	1		y_7
8	1	-1	-1	-1		y_8

- 如果一次回归方程中有交互作用项，则回归方程不是线性的；
- 但交互作用项回归系数的计算和检验与线性项是相同的。

表7-5 三因素一次回归正交设计结构矩阵与结果计算表

处理号	X_0	X_1	X_2	X_3		y
						y_1
						y_2
						y_3
						y_4
						y_5
						y_6
7	1	-1	-1	1		y_7
8	1	-1	-1	-1		y_8

$$\sum_{k=1}^{N} x_{kj} = 0, (j = 1, 2, \ldots, p)$$

$$\sum_{k=1}^{N} x_{ki} x_{kj} = 0, (i, j = 1, 2, \ldots, p, i \neq j)$$

$$\sum_{k=1}^{N} x_{kj}^{2} = N, (j = 1, 2, \ldots, p)$$

□交互作用列可直接由表中相应元素列的对应水平相乘得到，原交互作用列表失去作用。

表7-5 三因素一次回归正交设计结构矩阵与结果计算表

处理号	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	y
1	1	1	1	1	1	1	1	y_1
2	1	1	1	-1	1	-1	-1	y_2
3	1	1	-1	1	-1	1	-1	y_3
4	1	1	-1	-1	-1	-1	1	y_4
5	1	-1	1	1	-1	-1	1	y_5
6	1	-1	1	-1	-1	1	-1	y_6
7	1	-1	-1	1	1	-1	-1	y_7
8	1	-1	-1	-1	1	1	1	y_8

$$\hat{Y} = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_m x_m$$

表7-5 三因素一次回归正交设计结构矩阵与结果计算表

处理号	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	y
1	1	1	1	1	1	1	1	y_1
2	1	1	1	-1	1	-1	-1	y_2
3	1	1	-1	1	-1	1	-1	y_3
4	1	1	-1	-1	-1	-1	1	y_4
5	1	-1	1	1	-1	-1	1	y_5
6	1	-1	1	-1	-1	1	-1	y_6
7	1	-1	-1	1	1	-1	-1	y_7
8	1	-1	-1	-1	1	1	1	y_8

表7-5 三因素一次回归正交设计结构矩阵与结果计算表

处理号	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	y
1	1	1	1	1	1	1	1	y_1
2	1	1	1	-1	1	-1	-1	y_2
3	1	1	-1	1	-1	1	-1	y_3
4	1	1	-1	-1	-1	-1	1	y_4
5	1	-1	1	1	-1	-1	1	y_5
	1	-1	1	-1	-1	1	-1	y_6
	1	-1	-1	1	1	-1	-1	y_7
8	1	-1	-1	-1	1	1	1	y_8
B_j	B_0	B_1	B_2	B_3	B_{12}	B_{13}	B_{23}	

$$B_j = \sum_{i=1}^{n} x_{ij} y_i$$

表7-5 三因素一次回归正交设计结构矩阵与结果计算表

处理号	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	y
1	1	1	1	1	1	1	1	y_1
2	1	1	1	-1	1	-1	-1	y_2
3	1	1	-1	1	-1	1	-1	y_3
4	1	1	-1	-1	-1	-1	1	y_4
5	1	-1	1	1	-1	-1	1	y_5
6	1	-1	1	-1	-1	1	-1	y_6
7	1	-1	-1	1	1	-1	-1	y_7
8	1	-1	-1	-1	1	1	1	y_8
B_j	B_0	B_1	B_2	B_3	B_{12}	B_{13}	B_{23}	
d_j	8	8	8	8	8	8	8	

$$d_j = \sum_{i=1}^{n} x_{ij}^2$$

表7-5 三因素一次回归正交设计结构矩阵与结果计算表

$$b=\frac{\sum xy-(\sum x)(\sum y)/n}{\sum x^2-(\sum x)^2/n}=\frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2}$$

处理号	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	y
				1	1	1	1	y_1
				-1	1	-1	-1	y_2
				1	-1	1	-1	y_3
4	1	1					1	y_4
5	1	-1					1	y_5
6	1	-1					-1	y_6
7	1	-1	-1	1	1	-1	-1	y_7
8	1	-1	-1	-1	1	1	1	y_8
B_j	B_0	B_1	B_2	B_3	B_{12}	B_{13}	B_{23}	
d_j	8	8	8	8	8	8	8	
b_j	b_0	b_1	b_2	b_3	b_{12}	b_{13}	b_{23}	

$$b_j=B_j/d_j$$

$$\hat{y}=b_0+b_1x_1+b_2x_2+b_3x_3+b_{12}x_1x_2+b_{13}x_1x_3+b_{23}x_2x_3$$

表7-5 三因素一次回归正交设计结构矩阵与结果计算表

- 确定了这些偏回归系数后，可以直接根据其绝对值的大小来判断各因素和交互作用的重要性；
- 回归正交设计中，所有因素的水平都经过了编码变换，它们在所研究的范围内是“平等的”，所求得的回归系数不受因素的单位和取值的影响，直接反映了该因素的大小；
- 回归系数的符号反映了因素对指标影响的正负。

$$b_j = B_j / d_j$$

8	1	-1	-1	-1	1	1	1	y_8
B_j	B_0	B_1	B_2	B_3	B_{12}	B_{13}	B_{23}	
d_j	8	8	8	8	8	8	8	
b_j	b_0	b_1	b_2	b_3	b_{12}	b_{13}	b_{23}	

实例分析1

一次回归正交设计

为了研究某作物的栽培技术，选择影响作物产量的3个主要因素：

水分状况 (全生育期土壤湿度占田间持水量的百分比)、

追施氮肥量、

密度，

试验指标为产量y（kg/小区）。进行一次回归正交设计并分析。

表7-2 作物栽培技术因素水平表

水平	水分状况Z_1（%）	追氮量Z_2(kg/hm^2)	密度Z_3(万株/hm^2)
水平1	95	40	65
水平2	75	20	45

试验指标为产量y（kg/小区）。

表7-2 作物栽培技术因素水平表

水平	水分状况Z_1（%）	追氮量Z_2(kg/hm^2)	密度Z_3(万株/hm^2)
水平1	95	40	65
水平2	75	20	45

$$\hat{Y} = b_0 + \sum_{j=1}^{m} b_j Z_j + \sum_{i<j} b_{ij} Z_i Z_j$$

◆根据多元线性回归的理论，用矩阵表示，以最小二乘法，可以推导出系数矩阵为

$$b = (x^T x)^{-1} x^T y$$

试验指标为产量y（kg/小区）。

表7-2 作物栽培技术因素水平表

水平	水分状况Z_1（%）	追氮量Z_2(kg/hm^2)	密度Z_3(万株/hm^2)
上水平	95	40	65
下水平	75	20	45

$$\hat{Y} = b_0 + \sum_{j=1}^{m} b_j Z_j + \sum_{i<j} b_{ij} Z_i Z_j$$

$$b = (x^T x)^{-1} x^T y$$

◆线性代数知识可知，系数矩阵如为对角阵时，其逆矩阵便于计算。

◆因此，如果能经过某种安排适当的试验点，使系数矩阵A为对角矩阵，不仅能简化其逆矩阵计算，而且还使得回归系数间不存在相关性。

◆从数学意义上讲，也就是要使结构矩阵x具有正交性。

正交设计

列出因素水平编码表

一次回归正交设计

表7-3 作物栽培技术因素水平编码表

名称	水分状况Z_1（%）	追氮量Z_2(kg/hm^2)	密度Z_3(万株/hm^2)
上水平（1）	95	40	65
下水平（-1）	75	20	45
零水平 （0）	85	30	55
变化区间	10	10	10

Z_{aj}

x_{aj}

$$x_{aj} = \frac{Z_{aj} - Z_{0j}}{\Delta_j} (a = 0,1,2; \quad j = 1,2,\ldots,m)$$

编码因子

列出因素水平编码表

$$\hat{y} = b_0 + \sum_{j=1}^{m} b_j Z_j + \sum_{k<j} b_{kj} Z_k Z_j (j = 1,2, \ldots m)$$

表7-3 作物栽培技术因素水平表

名称	水分状况Z_1（%）	追氮量Z_2(kg/hm²)	密度Z_3(万株/hm²)
上水平	95	40	65
下水平	75	20	45
零水平	85	30	55
变化区间	10	10	10

$$\hat{y} = b_0 + \sum_{j=1}^{m} b_j x_j + \sum_{k<j} b_{kj} x_k x_j (j = 1,2, \ldots m)$$

表7-3-1 作物栽培技术因素水平编码表

名称	水分状况X_1（%）	追氮量X_2(kg/hm²)	密度X_3(万株/hm²)
上水平（1）	1	1	1
下水平（-1）	-1	-1	-1
零水平 （0）	0	0	0

$L_8(2^7)$

1

-1

试验号	1	2		4			7
	A	B	AB	C			
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1						1
5	2						2
6	2						1
7	2						1
8	2						2

处理号	X_1	X_2	X_3
1	1	1	1
2	1	1	-1
3	1	-1	1
4	1	-1	-1
5	-1	1	1
6	-1	1	-1
7	-1	-1	1
8	-1	-1	-1

因素的2个水平

编码因素的2个水平

表7-3-2 作物栽培技术实施方案表

处理号	水分状况x_1（%）	追氮量x_2(kg/hm^2)	密度x_3(万株/hm^2)
1	1 (95)	1 (40)	1 (65)
2	1	1	-1 (45)
3	1	-1 (20)	1
4	1	-1	-1
5	-1 (75)	1	1
6	-1	1	-1
7	-1	-1	1
8	-1	-1	-1

表7-3-3 作物栽培技术实施方案及结果表

处理号	水分状况 x_1（%）	追氮量 x_2(kg/hm^2)	密度 x_3(万株/hm^2)	产量y（kg/小区）
1	1 (95)	1 (40)	1 (65)	2.1
2	1	1	-1 (45)	2.3
3	1	-1 (20)	1	3.3
4	1	-1	-1	4.0
5	-1 (75)	1	1	5.0
6	-1	1	-1	5.6
7	-1	-1	1	6.9
8	-1	-1	-1	7.8

$$\hat{y} = b_0 + \sum_{j=1}^{m} b_j x_j + \sum_{k<j} b_{kj} x_k x_j (j = 1,2, \dots m)$$

表7-5 三因素一次回归正交设计结构矩阵与结果计算表

处理号		X_1	X_2	X_3		y
1		1	1	1		y_1
2		1	1	-1		y_2
3		1	-1	1		y_3
4		1	-1	-1		y_4
5		-1	1	1		y_5
6		-1	1	-1		y_6
7		-1	-1	1		y_7
8		-1	-1	-1		y_8

- 常数列
- 互作列

表7-5 三因素一次回归正交设计结构矩阵与结果计算表

处理号	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	y
1	1	1	1	1	1	1	1	y_1
2	1	1	1	-1	1	-1	-1	y_2
3	1	1	-1	1	-1	1	-1	y_3
4	1	1	-1	-1	-1	-1	1	y_4
5	1	-1	1	1	-1	-1	1	y_5
6						1	-1	y_6
7						-1	-1	y_7
8						1	1	y_8

$$\sum_{k=1}^{N} x_{kj} = 0,(j=1,2,\ldots,p)$$

$$\sum_{k=1}^{N} x_{ki}x_{kj} = 0,(i,j=1,2,\ldots,p,i\neq j)$$

$$\sum_{k=1}^{N} x_{kj}^2 = N,(j=1,2,\ldots,p)$$

(3) 计算回归系数及偏回归平方和

表7-5 三因素一次正交回归设计结构矩阵与试验结果计算表

处理号	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	y
1	1	1	1	1	1	1	1	2.1
							1	2.3
							1	3.3
								4.0
5	1	-1	1	1	-1	-1	1	5.0

$$\hat{y} = 4.625 - 1.70x_1 - 0.875x_2 - 0.3x_3 + 0.15x_1x_2 + 0.075x_1x_3 + 0.1x_2x_3$$

- 对回归方程进行显著性检验，与多元线性回归方程相同，采用F检验；
- 由于试验设计具有正交性，消除了回归系数间的相关性，回归平方和为各项偏回归平方和之和。

表7-5 三因素一次回归正交设计结构矩阵与结果计算表

处理号	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	y
1	1	1	1	1	1	1	1	y_1
2	1	1	1	-1	1	-1	-1	y_2
3	1	1	-1	1	-1	1	-1	y_3
4	1	1	-1	-1	-1	-1	1	y_4
5	1	-1	1	1	-1	-1	1	y_5
6	1	-1	1	-1	-1	1	-1	y_6
7	[illegible]	[illegible]	[illegible]	1	1	-1	-1	y_7
8	1	-1	-1	-1	1	1	1	y_8
B_j	B_0	B_1	B_2	B_3	B_{12}	B_{13}	B_{23}	
d_j	8	8	8	8	8	8	8	
b_j	b_0	b_1	b_2	b_3	b_{12}	b_{13}	b_{23}	
Q_j	-	Q_1	Q_2	Q_3	Q_{12}	Q_{13}	Q_{23}	

$$Q_j = SS_j = B_j^2 / d_j$$

(3) 计算回归系数及偏回归平方和

表7-5 三因素一次正交回归设计结构矩阵与试验结果计算表

$$\hat{y} = 4.625 - 1.70x_1 - 0.875x_2 - 0.3x_3 + 0.15x_1x_2 + 0.075x_1x_3 + 0.1x_2x_3$$

处理号	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	y
1	1	1	1	1	1	1	1	2.1
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2.3
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	3.3
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	4.0
5	1	-1	1	1	-1	-1	1	5.0
6	1	-1	1	-1	-1	1	-1	5.6
7	1	-1	-1	1	1	-1	-1	6.9
8	1	-1	-1	-1	1	1	1	7.8
$B_j = \sum x_{aj} y_a$	37.0	-13.6	-7	-2.4	1.2	0.6	0.8	
$d_j = \sum x_{aj}^2$	8	8	8	8	8	8	8	
$b_j = B_j / d_j$	4.625	-1.7	-0.875	-0.3	0.15	0.075	0.1	
$Q_j = B_j^2 / d_j$	-	23.12	6.125	0.72	0.18	0.045	0.08	

表7-3 三因素一次回归正交设计结构矩阵与结果计算表

处理号	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	y
1	1	1	1	1	1	1	1	y_1
2	1	1	1	-1	1	-1	-1	y_2
						1	-1	y_3
						-1	1	y_4
5	1	-1				-1	1	y_5
6	1	-1				1	-1	y_6
7	1	-1	-1	1	1	-1	-1	y_7
8	1	-1				1	1	y_8
B_j	B_0	B_1	B_2	B_3	B_{12}	B_{13}	B_{23}	
d_j	8	8	8	8	8	8	8	
b_j	b_0	b_1	b_2	b_3	b_{12}	b_{13}	b_{23}	
Q_j	-	Q_1	Q_2	Q_3	Q_{12}	Q_{13}	Q_{23}	

(2) 回归方程及偏回归系数的显著性检验

$$SS_T = \sum (y-\bar{y})^2 = \sum y^2 - \frac{(\sum y)^2}{n} = \sum y^2 - \frac{B_0^2}{n}$$

$$SS_T = SS_R + SS_e$$

$$df_T = n-1$$

$$df_T = df_R + df_e$$

表7-5 三因素一次回归正交设计结构矩阵与结果计算表

在一次回归正交设计下，

$$d_j=\sum x_{aj}^2=n$$

因素项回归平方和

$$Q_i=b_jB_j=\frac{B_j^2}{n}=nb_j^2$$

互作项回归平方和

$$Q_{ij}=b_{ij}B_{ij}=\frac{B_{ij}^2}{n}=nb_{ij}^2$$

因素项与互作项自由度

$$df_i=df_{ij}=1$$

处理号	X_0	X	X	X	X_1X_2	X_1X_3	X_2X_3	y
1	1	1	1	1	1	1	1	y_1
2	1	1				1	-1	y_2
3	1	1					-1	y_3
4	1	1	-1	-1	-1	-1	1	y_4
5	1	-1					1	y_5
6	1	-1					-1	y_6
7	1	-1	-1	1	1	-1	-1	y_7
		-1			1	1	1	y_8
B_j	B_0	B_1	B_2	B_3	B_{12}	B_{13}	B_{23}	
d_j	8	8	8	8	8	8	8	
b_j	b_0	b_1	b_2	b_3	b_{12}	b_{13}	b_{23}	
Q_j	-	Q_1	Q_2	Q_3	Q_{12}	Q_{13}	Q_{23}	

m个因素

在一次回归正交设计下，由于偏回归系数两两相互独立，

回归平方和等于各偏回归平方和之和 $SS_R = \sum_{j=1}^{m} Q_j + \sum Q_{ij}$

回归自由度为各偏回归自由度之和

$$df_R = \sum_{j=1}^{m} df_j + \sum df_{ij} = m + C_m^2 = \frac{m(m+1)}{2}$$

剩余（误差）平方和 $SS_e = SS_T - SS_R$

剩余（误差）自由度 $df_e = df_T - df_R = (n-1) - \frac{m(m+1)}{2}$

$$F_R = \frac{MS_R}{MS_e} = \frac{SS_R / df_R}{SS_e / df_e}$$

- 显著
- 不显著

(2) 回归方程及偏回归系数的显著性检验

$$F_j = \frac{MS_j}{MS_e} = \frac{Q_j}{MS_e} \qquad F_{ij} = \frac{MS_{ij}}{MS_e} = \frac{Q_{ij}}{MS_e}$$

$$Q_i = b_j B_j = \frac{B_j^2}{n} = nb_j^2 \qquad Q_{ij} = b_{ij} B_{ij} = \frac{B_{ij}^2}{n} = nb_{ij}^2$$

由上面的计算可知，各项偏回归平方和分别与 b_i 或 b_{ij} 的平方成正比。这说明在由回归正交设计所求得的回归方程中，偏回归系数绝对值的大小表示了对应变量（因素或互作）作用的大小，其符号反映了这种作用的性质。

(3) 计算回归系数及偏回归平方和

表7-5 三因素一次正交回归设计结构矩阵与试验结果计算表

处理号	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	y
1	1	1	1	1	1	1	1	2.1
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1	2.3
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1	3.3
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	4.0
5	1	-1	1	1	-1	-1	1	5.0

$$\hat{y} = 4.625 - 1.70x_1 - 0.875x_2 - 0.3x_3 + 0.15x_1x_2 + 0.075x_1x_3 + 0.1x_2x_3$$

- 一次正交归设计消除了回归系数间的相关性，
- 在检验过程中，若某些因素或互作项的偏回归系数不显著，则这些因素或互作项可以从回归方程中剔除，此时不影响其它回归系数的数值。
- 将被剔除项的偏回归平方和、自由度并入剩余平方和和与自由度，然后再次进行相关的分析计算。

(3) 计算回归系数及偏回归平方和

表7-5 三因素一次正交回归设计结构矩阵与试验结果计算表

处理号	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	y
1	1	1	1	1	1	1	1	2.1
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
4	1	1	-1	-1	-1	-1	1	4.0
5	[illegible]	[illegible]	[illegible]	[illegible]	-1	-1	1	5.0
6	1	-1	1	-1	-1	1	-1	5.6
7	1	-1	-1	1	1	-1	-1	6.9
8	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1	7.8
$B_j=\sum x_{aj}y_a$	37.0	-13.6	-7	-2.4	1.2	0.6	0.8	
$d_j=\sum x_{aj}^2$	[illegible]	8	8	8	8	8	8	
$b_j=B_j/d_j$	4.625	-1.7	-0.875	-0.3	0.15	0.075	0.1	
$Q_j=B_j^2/d_j$	-	23.12	6.125	0.72	0.18	0.045	0.08	

$$SS_T=\sum(y-\bar{y})^2=\sum y^2-\frac{(\sum y)^2}{n}=\sum y^2-\frac{37^2}{8}=30.275$$

$$df_T=n-1=8-1=7$$

$$SS_R=23.12+6.126+\ldots+0.08=30.27$$

$$df_R=6$$

变异来源	平方和	自由度	方差	*F*	
x_1	23.12	1	23.12	4624	**
x_2	6.125	1	6.125	1225	**
x_3	0.72	1	0.72	144	
x_1x_2	0.18	1	0.18	36	
x_1x_3	0.045	1	0.045	9	
x_2x_3	0.08	1	0.08	16	
回归	30.27	6	5.045	1009	**
离回归	0.005	1	0.005		
总	30.275	7			

- 上述回归方程的检验，是相对于平均剩余平方和而言，变量部分的影响是否显著。
- 检验结果显著的，说明一次回归方程在试验点（即上、下界）上与试验结果拟合得很好；
- 但不能保证在被研究的整个回归区域内拟合得很好，即不能保证采用一次回归模型是最好的。

拟合性进行检验

- 为分析经F检验结果为显著的一次回归方程在被研究区域内的拟合情况，
- 可通过在零水平试验点所安排的重复试验值估计真正的试验误差，
- 进而检验所建立的回归方程的失拟性（亦称为拟合度检验）。
- 零水平试验点一般重复2-6次。

表7-6 三因素一次正交回归设计（零水平试验点重复3次）与试验结果计算表

处理号	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	y
1	1	1	1	1	1	1	1	2.1
2	1	1	1	-1	1	-1	-1	2.3
3	1	1	-1	1	-1	1	-1	3.3
4	1	1	-1	-1	-1	-1	1	4.0
5	1	-1	1	1	-1	-1	1	5.0
6	1	-1	1	-1	-1	1	-1	5.6
7	1	-1	-1	1	1	-1	-1	6.9
8	1	-1	-1	-1	1	1	1	7.8
9	1	0	0	0	0	0	0	y_9
10	1	0	0	0	0	0	0	y_{10}
11	1	0	0	0	0	0	0	y_{11}

误差

设在零水平试验点安排了m_0次重复试验，

试验指标的观测值分别为：$y_{01}, y_{02}, \cdots, y_{0m_0}$

差异？

$$SS = \sum_{i=1}^{m_0}(y_{0i} - \overline{y_0})^2 = \sum y_{0i}^2 - \frac{(\sum y_{0i})^2}{m_0} \quad SS_{el}$$

$$df = m_0 - 1 \quad df_{el}$$

表7-6 三因素一次正交回归设计（零水平试验点重复3次）与试验结果计算表

处理号	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	y
1	1	1	1	1	1	1	1	y_1
2	1	1	1	-1	1	-1	-1	y_2
3	1	1	-1	1	-1	1	-1	y_3
4	1	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1	y_4
5	1	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1	y_5
6	1	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-1	y_6
7	1	-1	-1	1	1	-1	-1	y_7
8	1	-1	-1	-1	1	1	1	y_8
9	1	0	0	0	0	0	0	y_9
10	1	0	0	0	0	0	0	y_{10}
11	1	0	0	0	0	0	0	y_{11}
B_i	B_0	B_1	B_2	B_3	B_{12}	B_{13}	B_{23}	
d_i	11	8	8	8	8	8	8	
b_j	b_0	b_1	b_2	b_3	b_{12}	b_{13}	b_{23}	
Q_j	-	Q_1	Q_2	Q_3	Q_{12}	Q_{13}	Q_{23}	

$$b_j = \frac{B_j}{d_j} = \sum x_{ij} y_j \Big/ \sum (x_{ij})^2$$

增加零水平试验后

表7-6 三因素一次正交回归设计（零水平试验点重复3次）与试验结果计算表

处理号	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	y
1	1	1	1	1	1	1	1	y_1
2	1	1	1	-1	1	-1	-1	y_2
3	1	1	-1	1	-1	1	-1	y_3
4	1	1	-1	-1	-1	-1	1	y_4
5	1	-1	1	1	-1	[illegible]	[illegible]	y_5
6	1	-1	1	-1	-1	[illegible]	[illegible]	y_6
7	1	-1	-1	1	1	-1	-1	y_7
8	[illegible]	[illegible]	-1	[illegible]	1	1	1	y_8
9	[illegible]	[illegible]	0	[illegible]	0	0	0	y_9
10	[illegible]	[illegible]	0	[illegible]	[illegible]	[illegible]	0	y_{10}
11	1	0	0	0	[illegible]	[illegible]	0	y_{11}
B_i	B_0	B_1	B_2	B_3	[illegible]	[illegible]	B_{23}	
d_i	11	8	8	8	8	8	8	
b_j	b_0	b_1	b_2	b_3	b_{12}	b_{13}	[illegible]	
Q_j	-	Q_1	Q_2	Q_3	Q_{12}	[illegible]	[illegible]	

$$SS_T$$

$$SS_T - SS_R = SS_e \rightarrow SS_{el}$$

$$SS_{Lf}$$

增加零水平试验后 SS_R 不变

$$SS_{Lf} = SS_e - SS_{el}$$

失拟平方和 表示了回归方程未能拟合的部分，包括未考虑的

其它因素及各自变量的高次项等所引起的差异。

$$df_{Lf} = df_e - df_{el}$$

总平方和 $SS_T = SS_R + SS_e = SS_R + SS_{Lf} + SS_{el}$

总自由度 $df_T = df_R + df_e = df_R + df_{Lf} + df_{el}$

$$df_T = n + m_0 - 1$$

失拟性F检验

$$F_{Lf}=\frac{MS_{Lf}}{MS_{el}}=\frac{SS_{Lf}}{df_{Lf}}\Big/\frac{SS_{el}}{df_{el}}$$

F_{Lf} $\xrightarrow{\text{不显著}}$ SS_{Lf} （由随机误差造成的）

$\Downarrow$

$$F_R=\frac{SS_R/df_R}{SS_e/df_e}$$

- 显著，说明回归方程是显著的，而且拟合得好
- 不显著，说明回归方程不显著的

$\left.\begin{matrix}F_{Lf}\\ F_R\end{matrix}\right\}$ $\xrightarrow{\text{显著}}$ 尽管回归方程显著，但拟合得不好，还有其它因素（未考虑的因素、因素的交互、含有自变量的非线性影响等）的影响，需要查明原因，进一步改进回归模型。

(3) 失拟性检验

表7-6 三因素一次正交回归设计（零水平试验点重复3次）与试验结果计算表

处理号	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	y
1	1	1	1	1	1	1	1	y_1
2	1	1	1	-1	1	-1	-1	y_2
3	1	1	-1	1	-1	1	-1	y_3
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
5	1	-1	1	1	-1	-1	1	y_5
6	1	-1	1	-1	-1	1	-1	y_6
7	1	-1	-1	1	1	-1	-1	y_7
8	1	-1	-1	-1	1	1	1	y_8
9	1	0	0	0	0	0	0	y_9
10	1	0	0	0	0	0	0	y_{10}
11	1	0	0	0	0	0	0	y_{11}
B_i	B_0	B_1	B_2	B_3	B_{12}	B_{13}	B_{23}	
d_i	11	8	8	8	8	8	8	
b_j	b_0	b_1	b_2	b_3	b_{12}	b_{13}	b_{23}	
Q_j	-	Q_1	Q_2	Q_3	Q_{12}	Q_{13}	Q_{23}	

$$\widehat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3$$

注意：

- 所得到的回归方程是在编码空间求得的，它表述了编码因子x与试验指标y之间实际存在的线性关系。

- 根据试验要求与实际需要，通常还应由编码空间转换到自然因素空间，寻求到试验指标y关于自然因素Z的回归方程。

- 无论用于预测、控制，还是用于调优，回归方程只在所试验的范围内有效，超出原有试验范围，就可能失去实际意义。

例 7.1

为了研究某作物的栽培技术，选择影响作物产量的3个主要因素：

水分状况 (全生育期土壤湿度占田间持水量的百分比)、

追施氮肥量、

密度，

试验指标为产量y（kg/小区）。进行一次回归正交设计并分析。

(1) 列出因素水平编码表

表7-7 作物栽培技术研究因素水平编码表

名称	编码xj	水分状况Z_1（%）	追氮量Z_2(kg/hm^2)	密度Z_3(万株/hm^2)
上水平(+1)	1	95	40	65
下水平(-1)	-1	75	20	45
零水平(0)	0	85	30	55
变化区间	Δ_j	10	10	10

(2) 列出试验方案并实施

试验要求考察3个因素及两两因素间的交互作用，并且需要对失拟性进行检验，

$L_8(2^7)$ **零水平试验点重复2次。**

表7-8　三因素一次回归正交设计试验方案与结果表

处理号	试验设计			实施方案			产量y（kg/小区）
	X_1	X_2	X_3	水分状况 Z_1（%）	追氮量 Z_2(kg/hm^2)	密度Z_3 (万株/hm^2)	
1	1	1	1	95	40	65	2.1
2	1	1	-1	95	40	45	2.3
3	1	-1	1	95	20	65	3.3
4	1	-1	-1	95	20	45	4.0
5	-1	1	1	75	40	65	5.0
6	-1	1	-1	75	40	45	5.6
7	-1	-1	1	75	20	65	6.9
8	-1	-1	-1	75	20	45	7.8
9	0	0	0	85	30	55	4.5
10	0	0	0	85	30	55	4.3

(3) 计算回归系数及偏回归平方和

表7-9 三因素一次正交回归设计结构矩阵与试验结果计算表

处理号	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	y
1	1	1	1	1	1	1	1	2.1
2	1	1	1	-1	1	-1	-1	2.3
3	1	1	-1	1	-1	1	-1	3.3
6	1	-1	1	-1	-1	1	-1	5.6
7	1	-1	-1	1	1	-1	-1	6.9
8	1	-1	-1	-1	1	1	1	7.8
9	1	0	0	0	0	0	0	4.5
10	1	0	0	0	0	0	0	4.3
$B_j=\sum x_{aj}y_a$	45.8	-13.6	-7	-2.4	1.2	0.6	0.8	
$d_j=\sum x_{aj}^2$	10	8	8	8	8	8	8	
$b_j=B_j/d_j$	4.58	-1.7	-0.875	-0.3	0.15	0.075	0.1	
$Q_j=B_j^2/d_j$	-	23.12	6.125	0.72	0.18	0.045	0.08	

$$\hat{y}=4.58-1.70x_1-0.875x_2-0.3x_3+0.15x_1x_2+0.075x_1x_3+0.1x_2x_3$$

(4) 失拟性检验与回归关系显著性检验

SS_T

$$SS_T=\sum_{a=1}^{10}y^2-\frac{1}{10}\left(\sum_{a=1}^{10}y\right)^2=240.14-\frac{45.8^2}{10}=30.376$$

df_T

$$df_T=10-1=9$$

表7-9 三因素一次正交回归设计结构矩阵与试验结果计算表

处理号	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	y
1	1	1	1	1	1	1	1	2.1
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2.3
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	3.3
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	4.0
5	1	-1	1	1	-1	-1	1	5.0
6	1	-1	1	-1	-1	1	-1	5.6
7	1	-1	-1	1	1	-1	-1	6.9
[illegible]	1	-1	-1	-1	1	1	1	7.8
9	1	0	0	0	0	0	0	4.5
10	1	0	0	0	0	0	0	4.3
$B_j=\sum x_{aj}y_a$	45.8	-13.6	-7	-2.4	1.2	0.6	0.8	
$d_j=\sum x_{aj}^2$	10	8	8	8	8	8	8	
$b_j=B_j/d_j$	4.58	-1.7	-0.875	-0.3	0.15	0.075	0.1	
$Q_j=B_j^2/d_j$	-	23.12	6.125	0.72	0.18	0.045	0.08	

(4) 失拟性检验与回归关系显著性检验

表7-9 三因素一次正交回归设计结构矩阵与试验结果计算表

SS_R

$$SS_R = 23.12 + 6.13 + 0.72 + 0.18 + 0.045 + 0.08 = 30.275$$

df_R

$$df_R = 6$$

处理号	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	y
1	1	1	1	1	1	1	1	2.1
2	1	1	1	-1	1	-1	-1	2.3
5	1	-1	1	1	-1	-1	1	5.0
6	1	-1	1	-1	-1	1	-1	5.6
7	1	-1	-1	1	1	-1	-1	6.9
8	1	-1	-1	-1	1	1	1	7.8
	1	0	0	0	0	0	0	4.5
	1	0	0	0	0	0	0	4.3
$B_j = \sum x_{aj} y_a$	45.8	-13.6	-7	-2.4	1.2	0.6	0.8	
$d_j = \sum x_{aj}^2$	10	8	8	8	8	8	8	
$b_j = B_j / d_j$	4.58	-1.7	-0.875	-0.3	0.15	0.075	0.1	
$Q_j = B_j^2 / d_j$	-	23.12	6.125	0.72	0.18	0.045	0.08	

(4) 失拟性检验与回归关系显著性检验

SS_e

$$SS_e = SS_T - SS_R = 30.376 - 30.275 = 0.101$$

df_e

$$df_e = 10 - 1 - 6 = 3$$

表7-9 三因素一次正交回归设计结构矩阵与试验结果计算表

处理号	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	y
1	1	1	1	1	1	1	1	2.1
2	1	1	1	-1	1	-1	-1	2.3
							-1	3.3
							1	4.0
5	1	-1	1	1	-1	-1	1	5.0
6	1	-1	1	-1	-1	1	-1	5.6
7	1	-1	-1	1	1	-1	-1	6.9
8	1	-1	-1	-1	1	1	1	7.8
			0	0	0	0	0	4.5
			0	0	0	0	0	4.3
$B_j = \sum x_{aj} y_a$	45.8	-13.6	-7	-2.4	1.2	0.6	0.8	
$d_j = \sum x_{aj}^2$	10	8	8	8	8	8	8	
$b_j = B_j / d_j$	4.58	-1.7	-0.875	-0.3	0.15	0.075	0.1	
$Q_j = B_j^2 / d_j$	-	23.12	6.125	0.72	0.18	0.045	0.08	

(4) 失拟性检验与回归关系显著性检验

SS_{el}

$$SS_{el}=(4.5-\frac{4.5+4.3}{2})^2+(4.3-\frac{4.5+4.3}{2})^2=0.02$$

df_{el}

$$df_{el}=m_0-1=2-1=1$$

表7-9 三因素一次正交回归设计结构矩阵与试验结果计算表

处理号	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	y
1	1	1	1	1	1	1	1	2.1
							-1	2.3
							-1	3.3
4	1	1	-1	-1	-1	-1	1	4.0
5	1	-1	1	1	-1	-1	1	5.0
6	1	-1	1	-1	-1	1	-1	5.6
7	1	-1	-1	1	1	-1	-1	6.9
			-1	-1	1	1	1	7.8
			0	0	0	0	0	4.5
10	1	0	0	0	0	0	0	4.3
$B_j=\sum x_{aj}y_a$	45.8	-13.6	-7	-2.4	1.2	0.6	0.8	
$d_j=\sum x_{aj}^2$	10	8	8	8	8	8	8	
$b_j=B_j/d_j$	4.58	-1.7	-0.875	-0.3	0.15	0.075	0.1	
$Q_j=B_j^2/d_j$	-	23.12	6.125	0.72	0.18	0.045	0.08	

(4) 失拟性检验与回归关系显著性检验

表7-9 三因素一次正交回归设计结构矩阵与试验结果计算表

SS_{Lf}

$$SS_{Lf} = SS_e - SS_{el} = 0.101 - 0.02 = 0.081$$

df_{Lf}

$$df_{Lf} = df_e - df_{el} = 3 - 1 = 2$$

处理号	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	y
1	1	1	1	1	1	1	1	2.1
								2.3
								3.3
4	1	1	-1	-1	-1	-1	1	4.0
5	1	-1	1	1	-1	-1	1	5.0
6	1	-1	1	-1	-1	1	-1	5.6
7	1	-1	-1	1	1	-1	-1	6.9
8	1	-1	-1	-1	1	1	1	7.8
					0	0	0	4.5
					0	0	0	4.3
$B_j = \sum x_{aj} y_a$	45.8	-13.6	-7	-2.4	1.2	0.6	0.8	
$d_j = \sum x_{aj}^2$	10	8	8	8	8	8	8	
$b_j = B_j / d_j$	4.58	-1.7	-0.875	-0.3	0.15	0.075	0.1	
$Q_j = B_j^2 / d_j$	-	23.12	6.125	0.72	0.18	0.045	0.08	

表7-10 三因素一次正交回归设计试验结果方差分析表

变异来源	*SS*	*df*	*MS*	F
x_1	23.12	1	23.12	680^{**}
x_2	6.13	1	6.13	180.294^{**}
x_3	0.72	1	0.72	21.176^{*}
x_1x_2	0.18	1	0.18	5.294
x_1x_3	0.045	1	0.045	1.324
x_2x_3	0.08	1	0.08	2.353
回归	30.275	6	5.046	148.41^{**}
剩余	0.101	3	0.034	
失拟	0.081	2	0.041	**2.025**
纯误差	0.02	1	0.02	
总变异	30.376	9		

(5) 将回归方程中的编码变量还原为实际变量。

表7-9 三因素一次正交回归设计结构矩阵与试验结果计算表

处理号	X_0	X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	y
1	1	1	1	1	1	1	1	2.1
2	1	1	1	-1	1	-1	-1	2.3
3	1	1	-1	1	-1	1	-1	3.3
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
5	1	-1	1	1	-1	-1	1	5.0
6	1	-1	1	-1	-1	1	-1	5.6
7	1	-1	-1	1	1	-1	-1	6.9
8	1	-1	-1	-1	1	1	1	7.8
9	1	0	0	0	0	0	0	4.5
10	1	0	0	0	0	0	0	4.3
$B_j=\sum x_{aj}y_a$	45.8	-13.6	-7	-2.4	1.2	0.6	0.8	
$d_j=\sum x_{aj}^2$	10	8	8	8	8	8	8	
$b_j=B_j/d_j$	4.58	-1.7	-0.875	-0.3	0.15	0.075	0.1	
$Q_j=B_j^2/d_j$	-	23.12	6.125	0.72	0.18	0.045	0.08	

$$\widehat{y}=4.58-1.70x_1-0.875x_2-0.3x_3+0.15x_1x_2+0.075x_1x_3+0.1x_2x_3$$

(5) 将回归方程中的编码变量还原为实际变量。

$$\hat{y} = 4.58 - 1.70x_1 - 0.875x_2 - 0.3x_3 + 0.15x_1x_2 + 0.075x_1x_3 + 0.1x_2x_3$$

$$x_1 = \frac{Z_1 - Z_{01}}{\Delta_1} = \frac{Z_1 - 85}{10}$$

，

$$x_2 = \frac{Z_2 - Z_{02}}{\Delta_2} = \frac{Z_2 - 30}{10}$$

，

$$x_3 = \frac{Z_3 - Z_{03}}{\Delta_3} = \frac{Z_3 - 55}{10}$$

$$\hat{y} = 4.58 - 1.70(\frac{Z_1 - 85}{10}) - 0.88(\frac{Z_2 - 30}{10}) - 0.30(\frac{Z_3 - 55}{10}) + 0.15 \times (\frac{Z_1 - 85}{10})(\frac{Z_2 - 30}{10}) +$$
$$0.075 \times (\frac{Z_1 - 85}{10})(\frac{Z_3 - 55}{10}) + 0.10 \times (\frac{Z_2 - 30}{10})(\frac{Z_3 - 55}{10})$$

$$\hat{y} = 25.28875 - 0.12920\, Z_1 - 0.27050\, Z_2 + 0.00375\, Z_3$$
$$+ 0.00150\, Z_1Z_2 - 0.00075\, Z_1Z_3 + 0.00100\, Z_2Z_3$$

名称	编码xj	水分状况Z_1（%）	追氮量Z_2(kg/hm^2)	密度Z_3(万株/hm^2)
上水平(+1)	1	95	40	65
下水平(-1)	-1	75	20	45
零水平(0)	0	85	30	55
变化区间	Δ_j	10	10	10

$$\hat{y} = 25.28875 - 0.12920\, Z_1 - 0.27050\, Z_2 + 0.00375\, Z_3 + 0.00150\, Z_1 Z_2 - 0.00075\, Z_1 Z_3 + 0.00100\, Z_2 Z_3$$

一次回归正交设计

例 7.2

某化工产品的产量(y) 与时间(Z_1)、温度(Z_2)、压力(Z_3)、溶液浓度(Z_4)有关。

实际生产中，时间 30-40min，

温度 50-60℃，

压力 2×10^5-6×10^5Pa，

溶液浓度为 20%-40%.

试用一次回归正交设计求其回归方程，要求考察时间与温度间的交互作用。

1. 因素编码

表7-11 因素水平编码表

名称	编码 xj	时间(Z_1)/min	温度(Z_2)/℃	压力(Z_3)/×10^5Pa	溶液浓度(Z_4)/%
上水平(+1)	1	40	60	6	40
下水平(-1)	-1	30	50	2	20
零水平(0)	0	35	55	4	30
变化区间		5	5	2	10

$$x_1 = \frac{Z_1 - 35}{5}$$

$$x_2 = \frac{Z_2 - 55}{5}$$

$$x_3 = \frac{Z_3 - 4}{2}$$

$$x_4 = \frac{Z_4 - 30}{10}$$

2. 试验方案设计

$$L_8(2^7)$$

因素x_1、 x_2、 x_3、 x_4依次安排在正交表的1、2、4、7列，

$x_1 x_2$互作安排在第3列。

试验号	x_1	x_2	x_3	x_4
1	1	1	1	1
2	1	1	-1	-1
3	1	-1	1	-1
4	1	-1	-1	1
5	-1	1	1	-1
6	-1	1	-1	1
7	-1	-1	1	1
8	-1	-1	-1	-1
9	0	0	0	0
10	0	0	0	0
11	0	0	0	0

3. 回归方程建立

表7-12 试验方案及结果计算表

试验号	x_0	x_1	x_1	x_3	x_4	x_1x_2	y
1	1	1	1	1	1	1	9.7
2	1	1	1	-1	-1	1	4.6
3	1	1	-1	1	-1	-1	10
4	1	1	-1	-1	1	-1	11
5	1	-1	1	1	-1	-1	9
6	1	-1	1	-1	1	-1	10
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]3
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]4
9	1	0	0	0	0	0	7.9
10	1	0	0	0	0	0	8.1
11	1	0	0	0	0	0	7.4
$B_j=\sum_{i=1}^{n}x_{ij}y_i$	87.4	6.6	2.6	8.0	12.0	-16.0	
$d_j=\sum_{i=1}^{n}x_{ij}^2$	11	8	8	8	8	8	
$b_j=\frac{B_j}{d_j}$	7.945	0.825	0.325	1.000	1.500	-2.000	
$Q_j=SS_j=\frac{B_j^2}{d_j}=b_jB_j=b_j^2d_j$	694.433	5.445	0.845	8.000	18.000	32.000	

$$\hat{Y}=7.495+0.825x_1+0.325x_2+x_3+1.5x_4-2x_1x_2$$

4. 回归方程检验

表7-12 四因素一次正交回归设计试验结果方差分析表

变异来源	平方和	自由度	方差	F	显著水平
x_1	5.445	1	5.445	76.26**	0.01
x_2	0.845	1	0.845	11.83*	0.05
x_3	8.000	1	8.000	112.04**	0.01
x_4	18.000	1	18.000	252.10**	0.01
x_1x_2	32.000	1	32.000	448.18**	0.01
回归	64.29	5	12.858	180.08**	0.01
剩余	0.357	5	0.0714		
失拟	0.097	3	0.032333	0.248718	<1
误差	0.26	2	0.13		
总和	64.744	10			

5. 回归方程变换

$$\hat{Y} = 7.495 + 0.825x_1 + 0.325x_2 + x_3 + 1.5x_4 - 2x_1x_2$$

名称	编码 xj	时间(Z_1)/min	温度(Z_2)/℃	压力(Z_3)/×10^5Pa	溶液浓度(Z_4)/%
上水平(+1)	1	40	60	6	40
下水平(-1)	-1	30	50	2	20
零水平(0)	0	35	55	4	30
变化区间		5	5	2	10

$$x_1 = \frac{Z_1 - 35}{5} \quad x_2 = \frac{Z_2 - 55}{5} \quad x_3 = \frac{Z_3 - 4}{2} \quad x_4 = \frac{Z_4 - 30}{10}$$

$$\hat{Y} = -162.05 + 4.57Z_1 + 2.87Z_2 + 0.50Z_3 + 0.15Z_4 - 0.08Z_1Z_2$$

- 在编制试验方案时，考虑到计算试验误差与显著性检验的需要，若将零水平重复试验一并编入试验方案中，这种设计称为整体设计，而将试验方案与零水平试验分别考虑的设计称为非整体设计。
- 整体设计的方案是一次全面制订，因此进行试验时，可以全面考虑，统一安排，既为所有试验点同时进行试验创造了条件，也为进一步缩小试验时空范围、减少试验干扰提供了可能，与非整体设计相比，整体设计可以更充分地利用零水平试验信息，提高常数项回归系数b_0的精度，而对一次项回归系数b_j的计算无任何影响。

回归方程显著性检验的主要内容有：

(1) 回归系数的检验，主要考察试验因素对试验指标是否有显著影响；

(2) 回归方程的检验，主要考察整个回归方程是否显著；

(3) 失拟检验，主要考察事先假定的回归模型是否符合实际，这是一项容易被忽视但非常重要的检验。

y 的变异

SS_R 试验因素取不同水平引起的变异

SS_e 试验误差引起的变异

SS_{Lf} 试验因素的非线性效应及其他条件因素及其交互作用等影响引起的变异

高次项

$$\hat{Y} = b_0 + \sum_{j=1}^{m} b_j Z_j + \sum_{i<j} b_{ij} Z_i Z_j$$

Y

Z_1 Z_2 … Z_j … Z_m

m个自变量时，二次回归方程的数学模型为

$$\hat{Y} = \beta_0 + \sum_{j=1}^{m} \beta_j Z_j + \sum \beta_{ij} Z_i Z_j + \sum_{j=1}^{m} \beta_{jj} Z_j^2 + \varepsilon$$

其回归方程

$$\hat{Y} = b_0 + \sum_{j=1}^{m} b_j Z_j + \sum b_{ij} Z_i Z_j + \sum_{j=1}^{m} b_{jj} Z_j^2$$

$$\hat{Y} = b_0 + \sum_{j=1}^{m} b_j Z_j + \sum b_{ij} Z_i Z_j + \sum_{j=1}^{m} b_{jj} Z_j^2$$

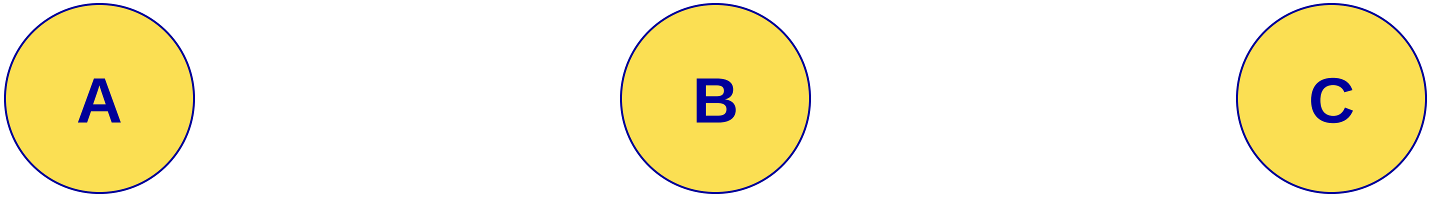
$$q = 1 + m + C_m^2 + m$$

$$q = 1 + m + \frac{m(m-1)}{2} + m = \frac{m^2 + 3m + 2}{2} = \frac{(m+2)(m+1)}{2} = C_{m+2}^2$$

$$N \geq q$$

每个因素至少要取**3**个水平

$$3^m$$



A_1, A_2, A_3 $\quad$ B_1, B_2, B_3 $\quad$ C_1, C_2, C_3

$$q = 1 + m + C_m^2 + m = \frac{(m+2)(m+1)}{2} = C_{m+2}^2$$

$$q = \frac{(3+2)(3+1)}{2} = 10$$

$N \geq 10$

$L_9(3^4) \quad L_{18}(3^7) \quad L_{27}(3^{13})$

$3^3 = 27$

表1 全面试验与组合设计的试验处理数和误差自由度

因素数m	回归系数个 数	三水平全面实施		组合设计		½实施	
		N	df_e	N	df_e	N	df_e
2	6	9	3	9	3		
3	10	27	17	15	5		
4	15	81	66	25	10		
5	21	243	222	43	22	27	6
6	28	729	701	77	49	45	17

三水平全面试验中，试验处理数和误差自由度大，

在一次回归正交设计的基础上再增加一些特定的试验点，通过适当的组合形成试验方案，即组合设计。

二次回归(组合)正交设计

研究重点：

寻找最优工艺参数、最佳配比组合、最适宜条件等；

试验结果：

二次反应方程

组合设计(Combinatorial design)

在参试因子（自变量）空间中选择几类不同特点的试验点（即分别处于不同球面上的点），适当组合而形成试验方案。

由于组合设计可选择多种类型的点，且有些类型点的数目（试验处理数）又可适当调节，故组合设计要比全面试验灵活，并且更为科学。

试验点

二水平因素全面点或其部分实施点

m_c

全面实施

$$m_c = 2^m$$

1/2实施

$$m_c = 2^{m-1}$$

1/4实施

$$m_c = 2^{m-2}$$

这些点的每一个坐标，都分别取1或-1；

轴点

原点

二水平因素全面点或其部分实施点

试验点

轴点

m_γ

$$m_\gamma = 2m$$

点在坐标轴上，且与坐标原点的距离都为γ，即这些点只有一个坐标γ或$-\gamma$，而其余坐标值都取零。

点在坐标图上通常用星号标出，故又称为星号点。星号臂的取值，根据一定的要求（如正交性、旋转性）进行调节。

原点

试验点

二水平因素全面点或其部分实施点

$$N = m_c + m_\gamma + m_0$$

轴点

原点

$$m_0$$

原点又称为中心点（基准点），即各自变量都取零的点，

中心试验点可作一次，也可作多次，

根据试验设计需求而定。

$$m = 2, m_0 = 1$$

$$N = m_c + m_\gamma + m_0 = 2^2 + 2\times 2 + 1 = 9$$

x_1 x_2

(1, 1)

(1, -1)

(-1, 1)

(-1, -1)

2个2水平因素的全面试验点

(γ, 0)

(0, -γ)

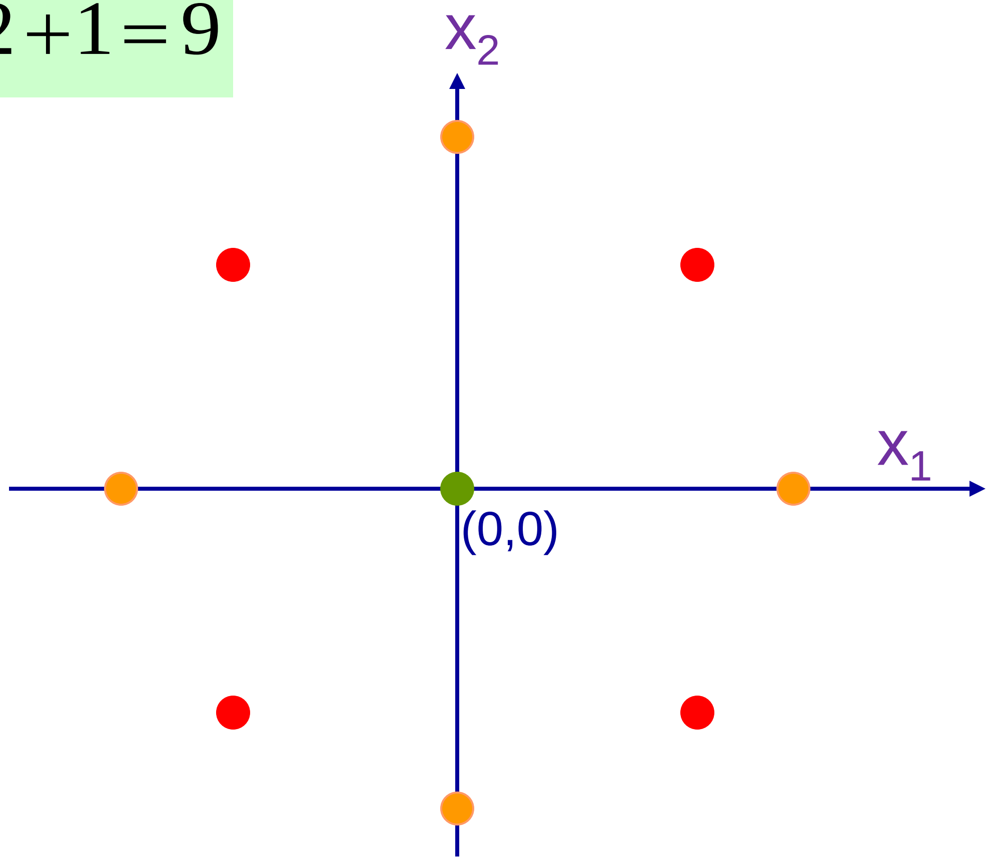
(0, γ)

(-γ, 0)

坐标轴上的点，即星号点

(0, 0)

由零水平组成的中心试验点



$x_1 \ x_2 \ x_3$

$$m = 2, m_0 = 1$$

$$N = m_c + m_\gamma + m_0 = 2^3 + 2\times3+1=15$$

(1, 1, 1)
(1, 1,-1)
(1, -1,1)
(1, -1,-1)
(-1, 1,1)
(-1, -1,1)
(-1, 1,-1)
(-1, -1,-1)

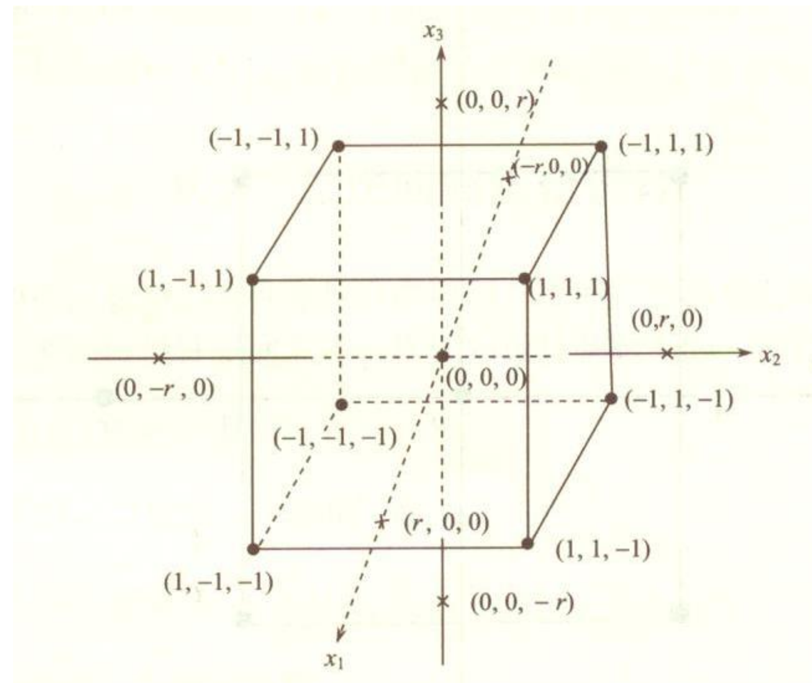
3个2水平因素的全面试验点

(γ, 0,0)
(-γ, 0,0)
(0, -γ,0)
(0, γ,0)
(0, 0,γ)
(0, 0,-γ)

坐标轴上的点，即星号点

(0, 0,0)

由零水平组成的中心试验点



组合设计的优点

- ① 减少试验点数，因素越多，试验点数减少得越多；
- ② 组合设计的试验点在因子空间中的分布是较均匀的；
- ③ 组合设计还便于在一次回归的基础上实施。

若一次回归不显著，可以在原m_c个试验点基础上，补充一些中心点与星号点试验，即可求得二次回归方程。

正交性的实现

1. 选择适当的二水平正交表，做好表头设计。

表2 二次回归组合设计试验点数

因素数m	正交表	表头设计	m_c	m_γ =2m	m_0	N	q
2	$L_4(2^3)$	1,2	$2^2=4$	4	1	9	6
3	$L_8(2^7)$	1,2,4	$2^3=8$	6	1	15	10
4	$L_{16}(2^{15})$	1,2,4,8	16	8	1	25	15
5	$L_{32}(2^{31})$	1,2,4,8,16	32	10	1	43	21
5(1/2实施)	$L_{16}(2^{15})$	1,2,4,8,15	16	10	1	27	21

表3　三因素一次回归组合设计结构矩阵（$m_0=1$）

试验号	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3
1	1	1	1	1	1	1	1
2	1	1	1	-1	1	-1	-1
3	1	1	-1	1	-1	1	-1
4	1	1	-1	-1	-1	-1	1
5	1	-1	1	1	-1	-1	1
6	1	-1	1	-1	-1	1	-1
7	1	-1	-1	1	1	-1	-1
8	1	-1	-1	-1	1	1	1
9	1	0	0	0	0	0	0

表3 三因素二次回归组合设计结构矩阵（$m_0=1$）

试验号		x_1	x_2	x_3						
1		1	1	1						
2		1	1	-1						
3		1	-1	1						
4		1	-1							
5		-1	1							
6		-1	1							
7		-1	-1	1						
8		-1	-1	-1						
9		γ	0	0						
10		$-\gamma$	0	0						
11		0	γ	0						
12		0	$-\gamma$	0						
13		0	0	γ						
14		0	0	$-\gamma$						
15		0	0	0						

二次组合设计是在一次回归正交设计的基础上加入星号点试验设计。

表3　三因素二次回归组合设计结构矩阵（$m_0=1$）

试验号	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	x_1^2	x_2^2	x_3^2
1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	-1	1	-1	-1	1	1	1
3	1	1	-1	1	-1	1	-1	1	1	1
4	1	1	-1	-1	-1	-1	1	1	1	1
5	1	-1	1	1	-1	-1	1	1	1	1
6	1	-1	1	-1	-1	1	1	1	1	1
7	1	-1	-1	1	1	[illegible]	[illegible]	1	1	1
8	1	-1	-1	-1	1	[illegible]	[illegible]	1	1	1
9	1	γ	0	0	0	[illegible]	[illegible]	γ^2	0	0
10	1	$-\gamma$	0	0	0	[illegible]	[illegible]	γ^2	0	0
11	1	0	γ	0	0	[illegible]	[illegible]	0	γ^2	0
12	1	0	$-\gamma$	0	0	0	0	0	γ^2	0
13	1	0	0	γ	0	0	0	0	0	γ^2
14	1	0	0	$-\gamma$	0	0	0	0	0	γ^2
15	1	0	0	0	0	0	0	0	0	0

常数列

一次项

互作项

二次项

表3　三因素二次回归组合设计结构矩阵（$m_0=1$）

试验号	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	x_1^2	x_2^2	x_3^2
1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	-1	1	-1	-1	1	1	1
3	1	1	-1	1	-1	1	-1	1	1	1
4	1	1	-1	-1	-1	-1	1	1	1	1
5	1	-1	1	1	-1	-1	1	1	1	1
6	1	-1	1	-1	-1	1	-1	1	1	1
7	1	-1	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
8	1	-1	-1	-1	1	1	1	1	1	1
9	1	γ	0	0	0	0	0	γ^2	0	0
10	1	$-\gamma$	0	0	0	0	0	γ^2	0	0
11	1	0	γ	0	0	0	0	0	γ^2	0
12	1	0	$-\gamma$	0	0	0	0	0	γ^2	0
13	1	0	0	γ	0	0	0	0	0	γ^2
14	1	0	0	$-\gamma$	0	0	0	0	0	γ^2
15	1	0	0	0	0	0	0	0	0	0

一次变量与交互作用列仍具有正交性

表3　三因素二次回归组合设计结构矩阵（m_0=1）

常数项列入正交表只是为了计算方便，并没有真正参加正交表的设计，无需要对其进行转换。

常数项与平方项不具有正交性

试验号	x_0	x_1								
1	1	1	1	1	1	1	1	1	1	1
2								1	1	1
3	1	1	-1	1	-1	1	-1	1	1	1
4			-1					1	1	1
5			1					1	1	1
6			1	-1	-1	1	-1	1	1	1
7								1	1	1
8								1	1	1
9			0					γ^2	0	0
10							0	γ^2	0	0
11							0	0	γ^2	0
12							0	0	γ^2	0
13	1	0	0							
14	1	0	0							
15	1	0	0	0						

$$\sum_{a=1}^{N} x_{aj}^2 = ?$$

$$\sum_{a=1}^{N} x_{aj}^2 = m_c + 2\gamma^2 \neq 0$$

$$\sum_{a=1}^{N} x_0 x_{aj}^2 = ?$$

$$\sum_{a=1}^{N} x_0 x_{aj}^2 = m_c + 2\gamma^2 \neq 0$$

$$\sum_{a=1}^{N} x_{am}^2 x_{aj}^2 = ?$$

$$\sum_{a=1}^{N} x_{am}^2 x_{aj}^2 = m_c \neq 0$$

为了使组合设计的结构矩阵具有正交性，必须使上述三式全部为零。

表3　三因素二次回归组合设计结构矩阵（$m_0=1$）

试验号	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	x_1^2	x_2^2	x_3^2
1	1	1	1	1						1
2	1	1	1	-1						1
					-1	1	-1	1	1	1
4	1	1	-1	-1				1	1	1
5	1	-1	1	1				1	1	1
	1	-1	1	-1	-1	1	-1	1	1	1
	1	-1	-1	1	1	-1	-1	1	1	1
9										
11										
12	1	0	-γ	0				0	γ^2	0
	1	0	0	γ				0	0	γ^2
	1	0	0	-γ				0	0	γ^2
	1	0	0	0	0	0	0	0	0	0

2. 转化以实现正交性。

（1）对平方项列的元素中心化

（2）$m, m_0 \rightarrow \gamma$

x_{aj}^2

$$x'_{aj} = x_{aj}^2 - \frac{\sum_a x_{aj}^2}{N} = x_{aj}^2 - \frac{(m_c + 2\gamma^2)}{N} \quad (j = 1, \cdots, m; a = 1, \cdots N)$$

$$\sum_{a=1}^{N} x'_{aj} ?$$

$$\sum_{a=1}^{N} x_0 x'_{aj} = ?$$

$$= 0 (j = 1, 2, \cdots, m)$$

$$\sum_{a=1}^{N} x_0 x'_{aj} = 0 (j = 1, 2, \cdots, m)$$

正交性的实现

$$\sum_{a=1}^{N} x'_{ai} = 0 \qquad \sum_{a=1}^{N} x_0 x'_{ai} = 0$$

2. 转化以实现正交性。

（1）对平方项列的元素中心化

（2）$m, m_0 \to \gamma$

$$\sum_{a=1}^{N} x'_{ai} x'_{aj} = 0\,?$$

$$\gamma = \sqrt{\frac{\sqrt{(m_c + 2m + m_0)m_c} - m_c}{2}}$$

$$\sum_{a=1}^{N} x'_{ai} x'_{aj} = 0 \quad (i \neq j,\ j = 1, 2, \cdots, m)$$

表4 γ 值表

m_0	m							
	2	3	4	5(1/2实施)	5	6(1/2实施)	6	7(1/2实施)
1	1.00000	1.21541	1.41421	1.54671	1.59601	1.72443	1.76064	1.88488
2	1.07809	1.28719	1.48258	1.60717	1.66183	1.78419	1.82402	1.94347
3	1.14744	1.35313	1.54671	1.66443	1.72443	1.84139	1.88488	2.00000
4	1.21000	1.41421	1.60717	1.71885	1.78419	1.89629	1.94347	2.05464
5	1.26710	1.47119	1.66443	1.77074	1.84139	1.94910	2.00000	2.10754
6	1.31972	1.52465	1.71885	1.82036	1.89629	2.00000	2.05464	2.15884
7	1.36857	1.57504	1.77074	1.86792	1.94910	2.04915	2.10754	2.20866
8	1.41421	1.62273	1.82036	1.91361	2.00000	2.09668	2.15884	2.25709
9	1.45709	1.66803	1.86792	1.95759	2.04915	2.14272	2.20866	2.30424
10	1.49755	1.71120	1.91361	2.00000	2.09668	2.18738	2.25709	2.35018
11	1.53587	1.75245	1.95759	2.04096	2.14272	2.23073	2.30424	2.39498

$$\gamma = \sqrt{\frac{\sqrt{(m_c + 2m + m_0)m_c} - m_c}{2}}$$

$m = 3, m_0 = 3$

⇩

$\gamma = 1.353$

$$N = m_c + 2m + m_0 = 2^3 + 2\times 3 + 3 = 17$$

表5 三因素二次回归正交组合设计的结构矩阵（$m_0=3$）

$$\gamma = 1.353$$

$$x_{aj}^{'} = x_{aj}^2 - \frac{\sum_a x_{aj}^2}{N} = x_{aj}^2 - \frac{(m_c + 2\gamma^2)}{N} \quad (j = 1, \cdots, m; a = 1, \cdots N)$$

$$\sum x_{aj}^2 = m_c + 2\gamma^2 = 8 + 2\times 1.831$$

$$x_{aj}^{'} = x_{aj}^2 - \frac{8 + 2\times 1.831}{17} = x_{aj}^2 - 0.686$$

$$\sum x_{aj}^{'} = 0$$

试验号	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	x'_1	x'_2	x'_3
1	1	[illegible]	1	1	1	1	1	0.314	0.314	0.314
2	1	1	1	1	1	1	1	0.314	0.314	0.314
3	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
4	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
6	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	1	-1	0.314	0.314	0.314
7	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-1	-1	0.314	0.314	0.314
8	1	-1	-1	-1	1	1	1	0.314	0.314	0.314
9	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	1.145	-0.686	-0.686
10	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	1.145	-0.686	-0.686
11	1	0	1.353	0	0	0	0	-0.686	1.145	-0.686
12	1	0	-1.353	0	0	0	0	-0.686	1.145	-0.686
13	1	0	0	1.353	0	0	0	-0.686	-0.686	1.145
14	1	[illegible]	[illegible]	-1.353	0	0	0	-0.686	-0.686	1.145
15	1	[illegible]	[illegible]	0	0	0	0	-0.686	-0.686	-0.686
16	1	0	0	0	0	0	0	-0.686	-0.686	-0.686
17	1	0	0	0	0	0	0	-0.686	-0.686	-0.686

二次回归正交组合设计与统计分析

1 确定试验因素及其上、下水平

2 将因素水平编码

3 确定实施方案

4 回归系数的计算与检验

与一次回归正交设计类似，

二次回归正交组合设计的方法，同样是在确定试验因素的基础上拟定每个因素的上、下水平，两者之算术平均数为零水平。

$$Z_{0j} = \frac{Z_{1j} + Z_{2j}}{2}$$

二次回归正交组合设计与统计分析

1 确定试验因素及其上、下水平

2 将因素水平编码

3 确定实施方案

4 回归系数的计算与检验

因素水平编码

Z_j

Z_{1j} → 下限

Z_{2j} → 上限

$$Z_{0j} = \frac{Z_{1j} + Z_{2j}}{2}$$

$$x_{aj} = \frac{Z_{aj} - Z_{0j}}{\Delta j}$$

$$m, m_0 \rightarrow \gamma$$

$$x_{aj} = \frac{Z_{aj} - Z_{0j}}{\Delta j}$$

因素水平编码 I

Z_j

Z_{1j} $\rightarrow$ 下限 $-\gamma$

Z_{2j} $\rightarrow$ 上限 γ

$$\Delta j = \frac{Z_{2j} - Z_{0j}}{\gamma}$$

$$\Delta j = \frac{Z_{2j} - Z_{0j}}{\gamma}$$

$$x_{aj} = \frac{Z_{aj} - Z_{0j}}{\Delta j}$$

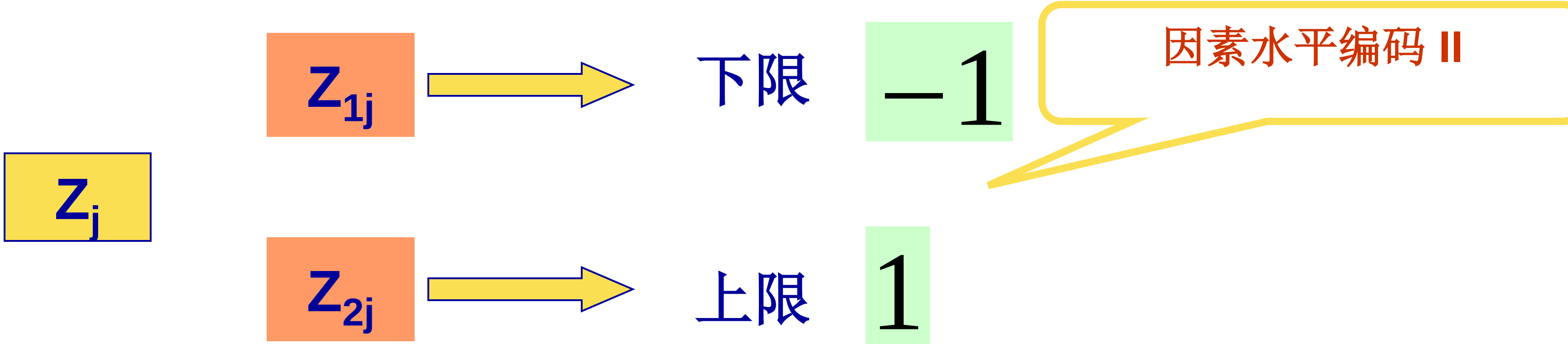
$$Z_{aj} = Z_{0j} + x_{aj}\Delta j$$

$$Z_{(x_j=1)j} = Z_{0j} + \Delta j$$

$$Z_{(x_j=-1)j} = Z_{0j} - \Delta j$$

表6 因素水平编码表（方法 I）

x_j	Z_1	Z_2	……	Z_m
γ	Z_{21}	Z_{22}	……	Z_{2m}
1	$Z_{01} + \Delta 1$	$Z_{02} + \Delta 2$	……	$Z_{0m} + \Delta m$
0	Z_{01}	Z_{02}	……	Z_{0m}
-1	$Z_{01} - \Delta 1$	$Z_{02} - \Delta 2$	……	$Z_{0m} - \Delta m$
$-\gamma$	Z_{11}	Z_{12}	……	Z_{1m}



$$\Delta j = Z_{2j} - Z_{0j} = \frac{Z_{2j} - Z_{1j}}{2}$$

二次回归连贯设计

在一次回归正交设计中，补充星号点，制订组合设计方案，寻找二次回归方程，可充分利用一次回归正交设计中试验的信息。

$$\Delta j = \frac{Z_{2j} - Z_{1j}}{2}$$

$$x_j = \frac{Z_j - Z_{0j}}{\Delta j}$$

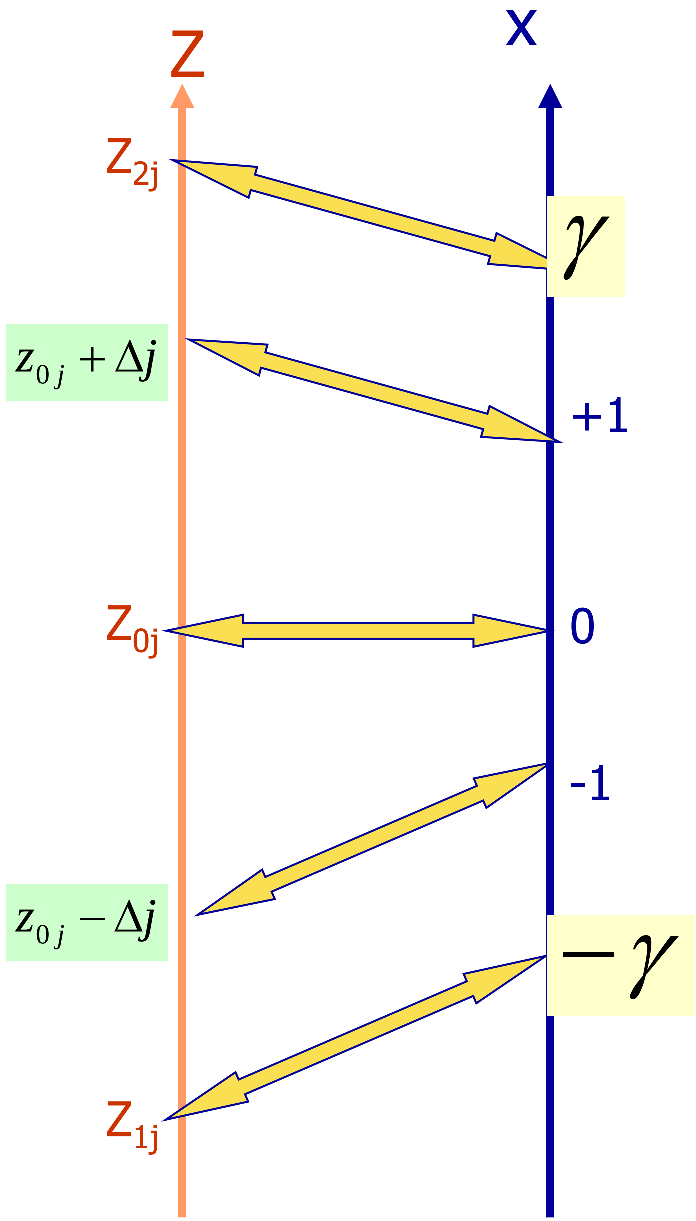
$$Z_j = Z_{0j} + x_j \Delta j$$

$$Z_{(x_j=\gamma)j} = Z_{0j} + \gamma\Delta j$$

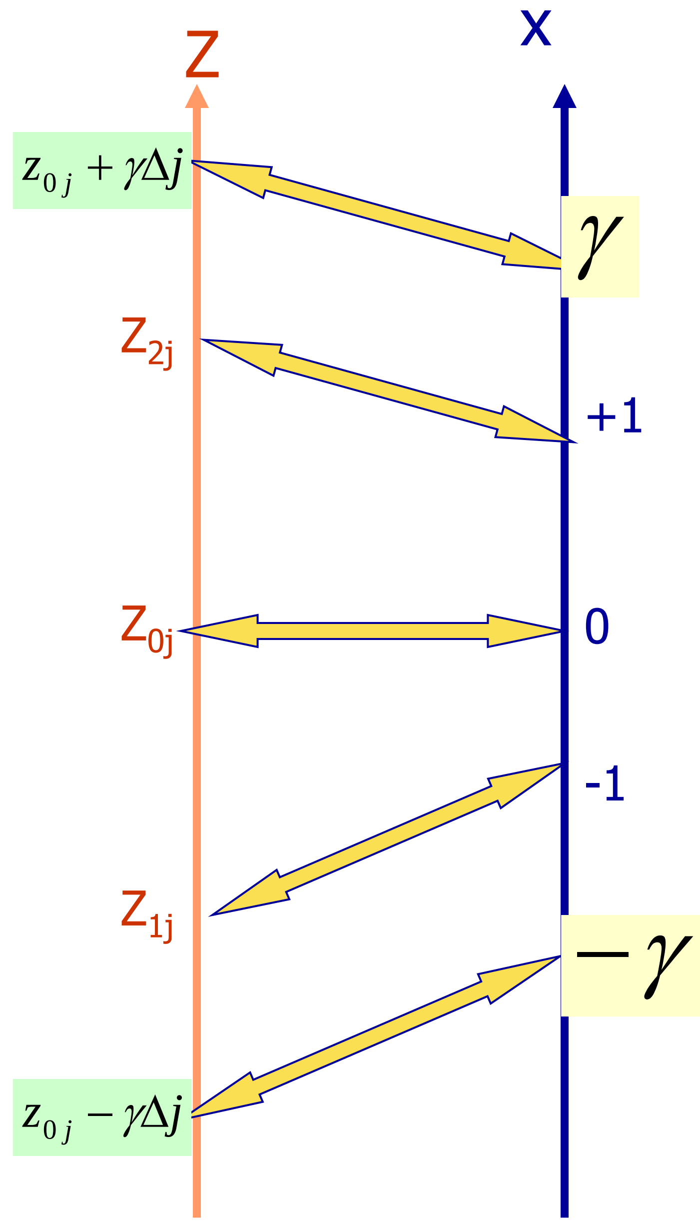
$$Z_{(x_j=-\gamma)j} = Z_{0j} - \gamma\Delta j$$

表7 因素水平编码表（方法 II）

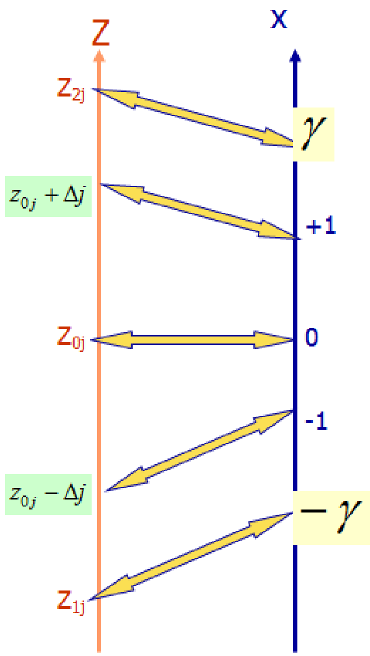
x_j	Z_1	Z_2	……	Z_m
γ	$Z_{01} + \gamma\Delta 1$	$Z_{02} + \gamma\Delta 2$	……	$Z_{0m} + \gamma\Delta m$
1	Z_{21}	Z_{22}	……	Z_{2m}
0	Z_{01}	Z_{02}	……	Z_{0m}
-1	Z_{11}	Z_{12}	……	Z_{1m}
$-\gamma$	$Z_{01} - \gamma\Delta 1$	$Z_{02} - \gamma\Delta 2$	……	$Z_{0m} - \gamma\Delta m$



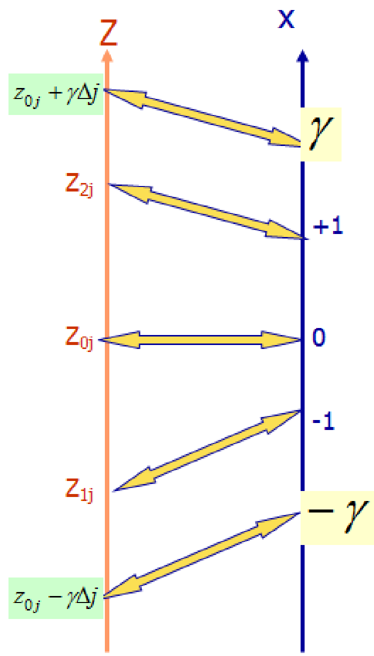
二次回归组合设计



二次回归连贯设计



1 二次回归组合设计



2 二次回归连贯设计

- 二次回归连贯设计，自然因素的实际上、下限不是在试验方案编制前预先选定的，而是根据编码因素的星号臂，利用一次回归设计编码公式计算确定的。
- 在设计时，应注意实际试验范围是否容许。

二次回归正交组合设计与统计分析

1 确定试验因素及其上、下水平

2 将因素水平编码

3 确定实施方案

4 回归系数的计算与检验

$m=3, m_c=2^3, m_0=3$

表8 三因素二次回归正交组合设计的结构矩阵（m_0=3）

试验号	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	x'_1	x'_2	x'_3
1	1	1	1	1	1	1	1	0.314	0.314	0.314
2	1	1	1	-1	1	-1	-1	0.314	0.314	0.314
3	1	1	-1	1	-1	1	-1	0.314	0.314	0.314
4	1	1	-1	-1	-1	-1	1	0.314	0.314	0.314
5	1	-1	1	1	-1	-1	1	0.314	0.314	0.314
6	1	-1	1	-1	-1	1	-1	0.314	0.314	0.314
7	1	-1	-1	1	1	-1	-1	0.314	0.314	0.314
8	1	-1	-1	-1	1	1	1	0.314	0.314	0.314
9	1	1.353	0	0	0	0	0	1.145	-0.686	-0.686
10	1	-1.353	0	0	0	0	0	1.145	-0.686	-0.686
11	1	0	1.353	0	0	0	0	-0.686	1.145	-0.686
12	1	0	-1.353	0	0	0	0	-0.686	1.145	-0.686
13	1	0	0	1.353	0	0	0	-0.686	-0.686	1.145
14	1	0	0	-1.353	0	0	0	-0.686	-0.686	1.145
15	1	0	0	0	0	0	0	-0.686	-0.686	-0.686
16	1	0	0	0	0	0	0	-0.686	-0.686	-0.686
17	1	0	0	0	0	0	0	-0.686	-0.686	-0.686

二次回归正交组合设计与统计分析

1. 确定试验因素及其上、下水平
2. 将因素水平编码
3. 确定实施方案
4. 回归系数的计算与检验

$$B = \sum xy \qquad d = \sum x^2 \qquad b = B/d \qquad Q = bB = B^2/d$$

$$\widehat{y} = b_0^{'} + \sum_{j=1}^{m} b_j x_j + \sum b_{ij} x_i x_j + \sum b_{jj} x_j^{'}$$

$$x_j^{'} = x_j^2 - \frac{\sum x_j^2}{N}$$

$$\widehat{y} = b_0 + \sum_{j=1}^{m} b_j x_j + \sum b_{ij} x_i x_j + \sum b_{jj} x_j^2$$

$$x_j = \frac{Z_j - Z_{0j}}{\Delta j}$$

影响茶叶出汁率的主要因素有：

榨法压力P，加压速度R，物料量R，榨汁时间t；

各因素对出汁率的影响不是简单的线性关系，

而且各因素间存在不同程度的交互作用，

以二次回归正交组合设计安排试验，

建立出汁率与各因素的回归方程。

(1) 根据初步试验，确定各因素的下、上水平

压力P(at): 5, 8

加压速度R (at/s): 1, 8

物料量W (g): 100, 400

榨汁时间t (min): 2, 4

(2) 因素水平编码

根据星号臂长的值（计算得出或查表得出），

对因素水平进行编码，得到编码变量。

$m = 4, m_0 = 3$ $\gamma = 1.547$

γ 值表

m_0	m							
	2	3	4	5(1/2实施)	5	6(1/2实施)	6	7(1/2实施)
1	1.00000	1.21541	1.41421	1.54671	1.59601	1.72443	1.76064	1.88488
2	1.07809	1.28719	1.48258	1.60717	1.66183	1.78419	1.82402	1.94347
3	1.14744	1.35313	1.54671	1.66443	1.72443	1.84139	1.88488	2.00000
4	1.21000	1.41421	1.60717	1.71885	1.78419	1.89629	1.94347	2.05464
5	1.26710	1.47119	1.66443	1.77074	1.84139	1.94910	2.00000	2.10754
6	1.31972	1.52465	1.71885	1.82036	1.89629	2.00000	2.05464	2.15884
7	1.36857	1.57504	1.77074	1.86792	1.94910	2.04915	2.10754	2.20866
8	1.41421	1.62273	1.82036	1.91361	2.00000	2.09668	2.15884	2.25709
9	1.45709	1.66803	1.86792	1.95759	2.04915	2.14272	2.20866	2.30424
10	1.49755	1.71120	1.91361	2.00000	2.09668	2.18738	2.25709	2.35018
11	1.53587	1.75245	1.95759	2.04096	2.14272	2.23073	2.30424	2.39498

(2) 因素水平编码

根据星号臂长的值（计算得出或查表得出），对因素水平进行编码，得到编码变量。

$$m = 4, m_0 = 3 \qquad \gamma = 1.547$$

表9 茶叶出汁率的因素水平编码表（方法I）

x_j	(P)	$Z_{(x_j=1)j} = Z_{0j} + \Delta j$	$Z_{(x_j=-1)j} = Z_{0j} - \Delta j$	(t)
$1.547(\gamma)$	8	8	400	4
1	7.47	6.76	347	3.646
0	6.5	4.5	250	3
-1	5.53	2.24	153	2.354
$-1.547(-\gamma)$	5	1	100	2
Δj	0.97	2.26	97	0.646

表2 二次回归组合设计试验点数

因素数m	正交表	表头设计	m_c	$m_\gamma=2m$	m_0	N	q
2	$L_4(2^3)$	1,2	$2^2=4$	4	1	9	6
3	$L_8(2^7)$	1,2,4	$2^3=8$	6	1	15	10
4	$L_{16}(2^{15})$	1,2,4,8	16	8	1	25	15
5	$L_{32}(2^{31})$	1,2,4,8,16	32	10	1	43	21
5(1/2实施)	$L_{16}(2^{15})$	1,2,4,8,15	16	10	1	27	21

(3) 列出试验实施方案。

试验号	试验设计				实施方案			
	x_1	x_2	x_3	x_4	$Z_1(P)$	$Z_2(R)$	$Z_3(W)$	$Z_4(t)$
1	1	1	1	1	7.47	6.76	347	3.646
2	1	1	1	−1	7.47	6.76	347	2.354
3	1	1	−1	1	7.47	6.76	153	3.646
4	1	1	−1	−1	7.47	6.76	153	2.354
5	1	[illegible]	[illegible]	[illegible]	7.47	2.24	347	3.646
6	1	−1	1	−1	7.47	2.24	347	2.354
7	1	−1	−1	1	7.47	2.24	153	3.646
8	1	−1	−1	−1	7.47	2.24	153	2.354
9	−1	1	1	1	5.53	6.76	347	3.646
10	−1	1	1	−1	5.53	6.76	347	2.354
11	−1	1	−1	1	5.53	6.76	153	3.646
12	−1	1	−1	−1	5.53	6.76	153	2.354
13	−1	−1	1	1	5.53	2.24	347	3.646
14	−1	−1	1	−1	5.53	2.24	347	2.354
15	−1	−1	−1	1	5.53	2.24	153	3.646
16	−1	−1	−1	−1	5.53	2.24	153	2.354
17	1.547	0	0	0	8	4.5	250	3
18	−1.547	0	0	0	5	4.5	250	3
19	0	1.547	0	0	6.5	8	250	3
20	0	−1.547	0	0	6.5	1	250	3
21	0	0	1.547	0	6.5	4.5	400	3
22	0	0	−1.547	0	6.5	4.5	100	3
23	0	0	0	1.547	6.5	4.5	250	4
24	0	0	0	−1.547	6.5	4.5	250	2
25	0	0	0	0	6.5	4.5	250	3
26	0	0	0	0	6.5	4.5	250	3
27	0	0	0	0	6.5	4.5	250	3

$$\gamma = 1.547$$

$$L_{16}(2^{15})$$

1,2,4,8　列

$$m_\gamma = 8$$

$$m_0 = 3$$

(4) 试验结果与统计分析

试验	x_0	x_1	x_2	x_3	x_4	x_1x_2	x_1x_3	x_1x_4	x_2x_3	x_2x_4	x_3x_4	x'_1	x'_2	x'_3	x'_4	y_a
1	1	1	1	1	1	1	1	1	1	1	1	0.23	0.23	0.23	0.23	43.26
2	1	1	1	1	−1	1	1	−1	1	−1	−1	0.23	0.23	0.23	0.23	39.60
3	1	1	1	−1	1	1	−1	1	−1	1	−1	0.23	0.23	0.23	0.23	48.73
4	1	1	1	−1	−1	1	−1	−1	−1	−1	1	0.23	0.23	0.23	0.23	48.73
5	1	1	−1	1	1	−1	1	1	−1	−1	1	0.23	0.23	0.23	0.23	47.26
6	1	1	−1	1	−1	−1	1	−1	−1	1	−1	0.23	0.23	0.23	0.23	42.97
7	1	1	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	−1	0.23	0.23	0.23	0.23	50.73
8	1	1	−1	−1	−1	−1	−1	−1	1	1	1	0.23	0.23	0.23	0.23	45.33
9	1	−1	1	1	1	−1	−1	−1	1	1	1	0.23	0.23	0.23	0.23	41.86
10	1	−1	1	1	−1	−1	−1	1	1	−1	−1	0.23	0.23	0.23	0.23	40.11
11	1	−1	1	−1	1	−1	1	−1	−1	1	−1	0.23	0.23	0.23	0.23	49.40
12	1	−1	1	−1	−1	−1	1	1	−1	−1	1	0.23	0.23	0.23	0.23	45.73
13	1	−1	−1	1	1	1	−1	−1	−1	−1	1	0.23	0.23	0.23	0.23	45.83
14	1	−1	−1	1	−1	1	−1	1	−1	1	−1	0.23	0.23	0.23	0.23	40.06
15	1	−1	−1	−1	1	1	1	−1	1	−1	−1	0.23	0.23	0.23	0.23	46.40
16	1	−1	−1	−1	−1	1	1	1	1	1	1	0.23	0.23	0.23	0.23	45.13
17	1	1.547	0	0	0	0	0	0	0	0	0	1.623	−0.77	−0.77	−0.77	48.72
18	1	−1.547	0	0	0	0	0	0	0	0	0	1.623	−0.77	−0.77	−0.77	45.48
19	1	0	1.547	0	0	0	0	0	0	0	0	−0.77	1.623	−0.77	−0.77	46.24
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	0	0	0	0	0	−0.77	1.623	−0.77	−0.77	47.52
27	1	0	0	0	0	0	0	0	0	0	0	−0.77	−0.77	−0.77	−0.77	48.06
B_j	1235.1	17.10	−8.27	−40.27	31.01	−5.65	−1.63	0.89	−16.29	−7.65	5.13	−4.476	−5.529	−24.745	−2.083	
d_j	27	20.79	20.79	20.79	20.79	16	16	16	16	16	16	11.45	11.45	11.45	11.45	
b_j	45.74	0.823	−0.398	−1.937	1.492	−0.353	−0.102	0.056	−1.018	−0.478	0.321	−0.391	−0.483	−2.161	−0.182	
Q_j	—	14.065	3.290	78.000	46.254	1.995	0.166	0.050	16.585	3.658	1.645	1.750	2.670	53.477	0.379	

$$x'_{aj} = x_{aj}^2 - \frac{\sum_a x_{aj}^2}{N} = x_{aj}^2 - \frac{(m_c + 2\gamma^2)}{N}$$

$$= x_{aj}^2 - \frac{(16 + 2\times 1.547^2)}{27} = x_{aj}^2 - 0.770$$

$$1.547^2 - 0.770 = 1.623$$

$$\hat{y} = 45.74 + 0.823x_1 - 0.398x_2 - 1.937x_3 + 1.492x_4$$
$$- 0.353x_1x_2 - 0.102x_1x_3 + 0.056x_1x_4 - 1.018x_2x_3 - 0.478x_2x_4 + 0.321x_3x_4$$
$$- 0.391x'_1 - 0.483x'_2 - 2.161x'_3 - 0.182x'_4$$

(5) 回归方差分析

拟合效果好，回归关系极显著。

表10　回归关系方差分析表

变异来源	*SS*	*df*	*MS*	*F*	临界F值
x_1	14.065	1	14.065	5.858*	4.75($F_{0.05}$)
x_2	3.290	1	3.290	1.370	
x_3	78.000	1	78.000	32.486**	9.33($F_{0.01}$)
x_4	46.254	1	46.254	19.264**	
x_1x_2	1.995	1	1.995	<1	
x_1x_3	0.166	1	0.166	<1	
x_1x_4	0.050	1	0.050	<1	
x_2x_3	16.585	1	16.585	6.908*	
x_2x_4	3.658	1	3.658	1.524	1.46($F_{0.25}$)
x_3x_4	1.645	1	1.645	<1	
x'_1	1.750	1	1.750	<1	
x'_2	2.670	1	2.670	1.112	
x'_3	53.477	1	53.477	22.273**	
x'_4	0.379	1	0.379	<1	
回归	223.984	14	15.999	6.663**	4.05($F_{0.01}$)
剩余	28.806	12	2.401		
失拟	28.301	10	2.830	11.186^{ns}	19.39($F_{0.05}$)
纯误	0.505	2	0.253		
总变异	252.790	26			

(5) 回归方差分析

表11 回归关系的第二次方差分析表

变异来源	SS	df	MS	F	临界F值
x_1	14.065	1	14.065	6.901*	4.35($F_{0.05}$)
x_3	78.000	1	78.000	38.273**	8.10($F_{0.01}$)
x_4	46.254	1	46.254	22.696**	
x_2x_3	16.585	1	16.585	8.138*	
x_2x_4	3.658	1	3.658	1.795	1.40($F_{0.25}$)
x'_3	53.477	1	53.477	26.240**	
回归	212.039	6	35.340	17.341**	3.87($F_{0.01}$)
剩余	40.751	20	2.038		
失拟	40.246	18	2.236	8.838^{ns}	9.43($F_{0.10}$)
纯误	0.505	2	0.253		
总变异	252.790	26			

$$\widehat{y} = 45.74 + 0.823x_1 - 1.937x_3 + 1.492x_4 - 1.018x_2x_3 - 0.478x_2x_4 - 2.161x'_3$$

$$\hat{y} = 45.74 + 0.823x_1 - 1.937x_3 + 1.492x_4 - 1.018x_2x_3 - 0.478x_2x_4 - 2.161x_3'$$

$$x_3' = x_3^2 - 0.77 \qquad x_1 = \frac{Z_1 - 6.5}{0.97} \qquad x_2 = \frac{Z_2 - 4.5}{2.26} \qquad x_3 = \frac{Z_3 - 250}{97}$$

$$\hat{y} = 15.954 + 0.848Z_1 + 2.413Z_2 + 0.116Z_3 + 3.783Z_4 - 4.644\times10^{-3}Z_2Z_3 - 0.327Z_2Z_4 - 2.297\times10^{-4}Z_3^2$$

Z_1:压力P(at): 5, 8

Z_2:加压速度R (at/s): 1, 8

Z_3:物料量W (g): 100, 400

Z_4:榨汁时间t (min): 2, 4

试建立溶液电导率与A、B两种物质浓度的线性关系。

A考查的范围是：Z_1：30-70 g/L;

B考查的范围是：Z_2：90-150 g/L。

为建立线性回归模型，采用回归设计。

$$L_4(2^3)$$

$$m_0 = 4$$

一次回归正交设计

表12 溶液电导率与物质浓度研究中的因素编码表

$z_j(x_j)$	A/（g/L） z_1	B/（g/L） z_2
$z_{2j}(1)$	**70**	**150**
$z_{0j}(0)$	**50**	**120**
$z_{1j}(-1)$	**30**	**90**
$\Delta j = \dfrac{z_{2j} - z_{1j}}{2}$	**20**	**30**
$x_j = \dfrac{z_j - z_{0j}}{\Delta j}$	$x_1 = \dfrac{z_1 - 50}{20}$	$x_2 = \dfrac{z_1 - 120}{30}$

表13 溶液电导率与物质浓度结构矩阵与结果

试验号	x_0	$x_1(z_1)$	$x_2(z_2)$	x_1x_2	y
1	1	1	1	1	5.0
2	1	1	-1	-1	6.7
3	[illegible]	[illegible]	[illegible]	[illegible]	8.5
4	[illegible]	[illegible]	[illegible]	[illegible]	2.0
5	[illegible]	[illegible]	[illegible]	[illegible]	2.8
6	[illegible]	[illegible]	[illegible]	[illegible]	3.2
7	1	0	0	0	3.4
8	1	0	0	0	3.0
a_j	8	4	4	4	
B_j	34.60	[illegible]	[illegible]	[illegible]	
b_j	4.325	[illegible]	[illegible]	[illegible]	
Q_j	149.645	0.36	5.76	16.81	
F_j		0.12	1.89	5.51	

$$F_{Lf}=\frac{SS_{Lf}/df_{Lf}}{SS_e/df_e}=180.08^{**}$$

二次回归连贯设计

- 该方程严重失拟，需要考虑试验因素的非线性效应。
- 可以在此基础上，补做星号试验来进行二次回归设计。

m_0	m							
	2	3	4	5(1/2实施)	5	6(1/2实施)	6	7(1/2实施)
1	1.00000	1.21541	1.41421	1.54671	1.59601	1.72443	1.76064	1.88488
2	1.07809	1.28719	1.48258	1.60717	1.66183	1.78419	1.82402	1.94347
3	1.14744	1.35313	1.54671	1.66443	1.72443	1.84139	1.88488	2.00000
4	1.21000	1.41421	1.60717	1.71885	1.78419	1.89629	1.94347	2.05464
5	1.26710	1.47119	1.66443	1.77074	1.84139	1.94910	2.00000	2.10754
6	1.31972	1.52465	1.71885	1.82036	1.89629	2.00000	2.05464	2.15884
7	1.36857	1.57504	1.77074	1.86792	1.94910	2.04915	2.10754	2.20866
8	1.41421	1.62273	1.82036	1.91361	2.00000	2.09668	2.15884	2.25709
9	1.45709	1.66803	1.86792	1.95759	2.04915	2.14272	2.20866	2.30424
10	1.49755	1.71120	1.91361	2.00000	2.09668	2.18738	2.25709	2.35018
11	1.53587	1.75245	1.95759	2.04096	2.14272	2.23073	2.30424	2.39498

$$N = m_c + m_r + m_0 = 4 + 4 + 4 = 12$$

表14 溶液电导率与物质浓度组合设计因素编码表（二次连贯设计）

$z_j(x_j)$	z_1/（g/L）	z_2/（g/L）
$z'_{2j}(\gamma)$	74.2	156.3
$z_{2j}(1)$	70	150
$z_{0j}(0)$	50	120
$z_{2j}(-1)$	30	90
$z'_{1j}(-\gamma)$	25.8	83.7
$\Delta j = \dfrac{z_{2j} - z_{1j}}{2}$	20	30
$x_j = \dfrac{z_j - z_{0j}}{\Delta j}$	$x_1 = \dfrac{z_1 - 50}{20}$	$x_2 = \dfrac{z_2 - 120}{30}$

表15 溶液电导率与物质浓度组合设计结构矩阵及结果

试验号		x_0	$x_1(z_1)$	$x_2(z_2)$	x_1x_2	$x_1^2(x_1')$	$x_2^2(x_2')$	y
m_c	1	1	1	1	1	1（0.423）	1（0.423）	5.0
	2	1	1	-1	-1	1（0.423）	1（0.423）	6.7
	3	1	-1	1	-1	1（0.423）	1（0.423）	8.5
	4	1	-1	-1	1	1（0.423）	1（0.423）	2.0
m_r	9	1	1.21	0	0	1.464（0.887）	0（-0.577）	5.9
	10	1	-1.21	0	0	1.464（0.887）	0（-0.577）	4.9
	11	1	0	1.21	0	0（-0.577）	1.464（0.887	5.8
	12	1	0	-1.21	0	0（-0.577）	1.464（0.887	2.9
m_o	5	1	0	0	0	0（-0.577）	0（-0.577）	2.8
	6	1	0	0	0	0（-0.577）	0（-0.577）	3.2
	7	1	0	0	0	0（-0.577）	0（-0.577）	3.4
	8	1	0	0	0	0（-0.577）	0（-0.577）	3
	dj	12	6.928	6.928	4	4.287	4.287	
	Bj	54.10	2.41	8.31	-8.20	6.778	3.703	
	bj	4.51	0.35	1.20	-2.05	1.58	0.86	
	Qj	243.90	0.838	9.965	16.810	10.716	3.199	
	Fj		22.76	270.54	456.38	290.92	86.85	

表16 溶液电导率与物质浓度组合设计结果的方差分析

变异来源	平方和	自由度	均方	F值	显著水平
x_1	0.838	1	0.838	22.76**	0.0031
x_2	9.965	1	9.965	270.54**	＜0.0001
x_1x_2	16.810	1	16.810	456.38**	＜0.0001
$x_1^2(x_1')$	10716	1	10.716	290.92**	＜0.0001
$x_2^2(x_2')$	3.199	1	3.199	86.85**	＜0.0001
回归	41.528	5	8.305633	225.49**	＜0.0001
剩余	0.22	6	0.036834		
误差e	0.2	3	0.066667		
失拟	0.02	3	0.00700	0.19	0.9470
总和	41.749	11			

回归方程不失拟，拟合效果好，可用于预测。

表15 溶液电导率与物质浓度组合设计结构矩阵及结果

试验号		x_0	$x_1(z_1)$	$x_2(z_2)$	x_1x_2	$x_1^2(x_1^{'})$	$x_2^2(x_2^{'})$	y
m_c	1	1	1	1	1	1（0.423）	1（0.423）	5.0
	2	1	1	-1	-1	1（0.423）	1（0.423）	6.7
	3	1	-1	1	-1	1（0.423）	1（0.423）	8.5
	4	1	-1	-1	1	1（0.423）	1（0.423）	2.0
m_r	9	1	1.21	0	0	1.464（0.887）	0（-0.577）	5.9
	10	1	-1.21	0	0	1.464（0.887）	0（-0.577）	4.9
	11	1	0	1.21	0	0（-0.577）	1.464（0.887	5.8
m_o	6	1	0	0	0	0（-0.577）	0（-0.577）	3.2
	7	1	0	0	0	0（-0.577）	0（-0.577）	3.4
	8	1	0	0	0	0（-0.577）	0（-0.577）	3
	dj	12	6.928	6.928	4	4.287	4.287	
	Bj	54.10	2.41	8.31	-8.20	6.778	3.703	
	bj	4.51	0.35	1.20	-2.05	1.58	0.86	
	Qj	243.90	0.838	9.965	16.810	10.716	3.199	
	Fj		22.76	270.54	456.38	290.92	86.85	

$$y = 4.51 + 0.35x_1 + 1.20x_2 - 2.05x_1x_2 + 1.58x_1^{'} + 0.86x_2^{1}$$

表14 溶液电导率与物质浓度组合设计因素编码表（二次连贯设计）

$z_j(x_j)$	z_1/（g/L）	z_2/（g/L）
$z_{2j}(1)$	70	150
$z_{0j}(0)$	50	120
$\Delta j = \dfrac{z_{2j} - z_{1j}}{2}$	20	
$x_j = \dfrac{z_j - z_{0j}}{\Delta j}$	$x_1 = \dfrac{z_1 - 50}{20}$	$x_2 = \dfrac{z_2 - 120}{30}$

$$y = 4.51 + 0.35x_1 + 1.20x_2 - 2.05x_1x_2 + 1.58x_1' + 0.86x_2'$$

$$x_1 = \frac{z_1 - 50}{20}$$

$$x_2 = \frac{z_2 - 120}{30}$$

$$x_1' = x_1^2 - 0.577$$

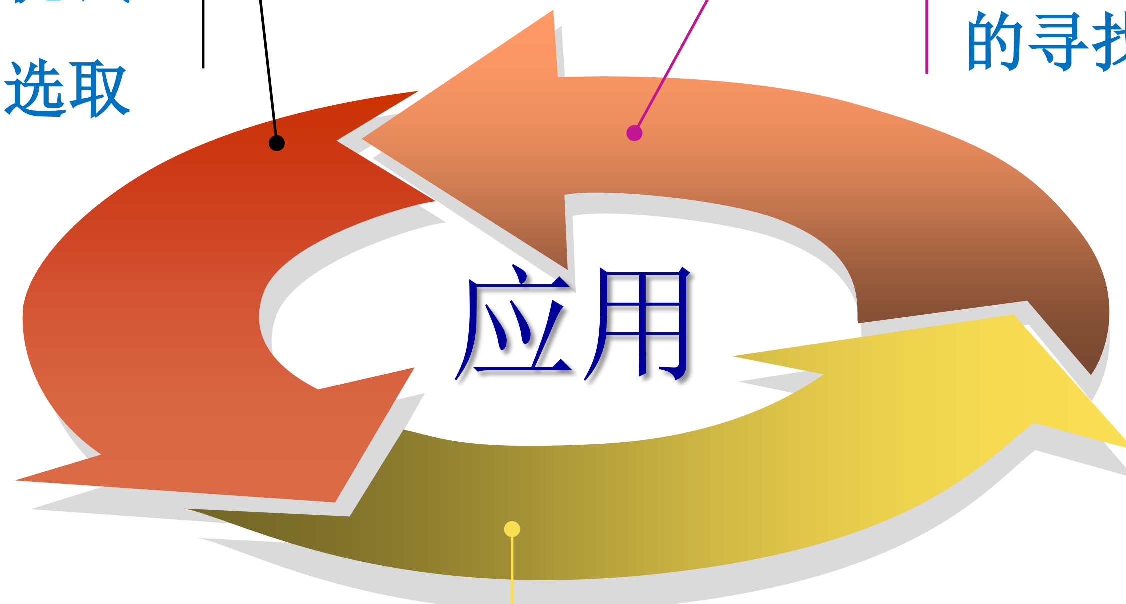
$$x_2' = x_2^2 - 0.577$$

$$y = 0.721 + 0.031z_1 - 0.023z_2 - 0.003z_1z_2 + 0.004z_1^2 + 0.001z_2^2$$

回归方程的应用

1. 在方程设计范围内最优试验因子的选取

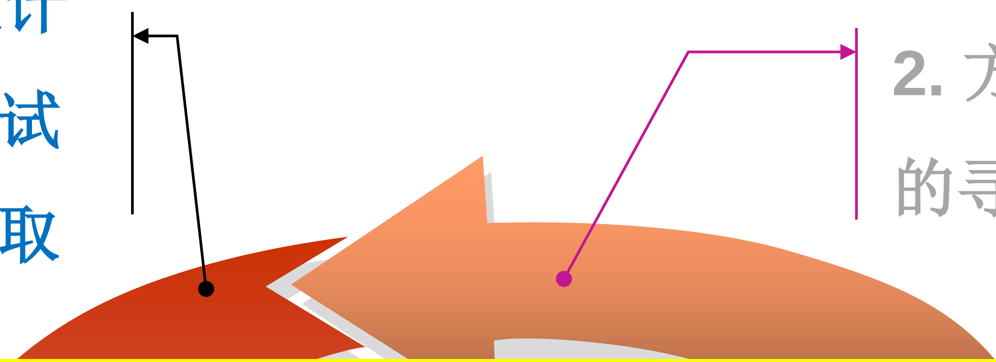
2. 方程局部最优点的寻找



3. 寻求最佳试验方案

回归方程的应用

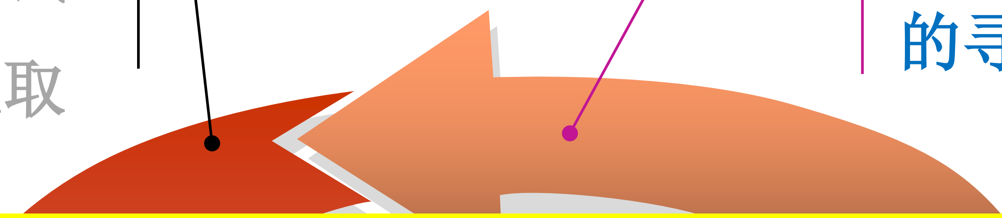
1. 在方程设计范围内最优试验因子的选取



2. 方程局部最优点的寻找

- 可通过软件程序对试验所设计的因子编码进行自动寻找，找出每一个因素已有的局部最优点，作为本试验的优化试验点。
- 找出的局部最优水平编码并不是理论上的最优点，不能代表真正的最优。
- 后续正交旋转和通用旋转设计中介绍。

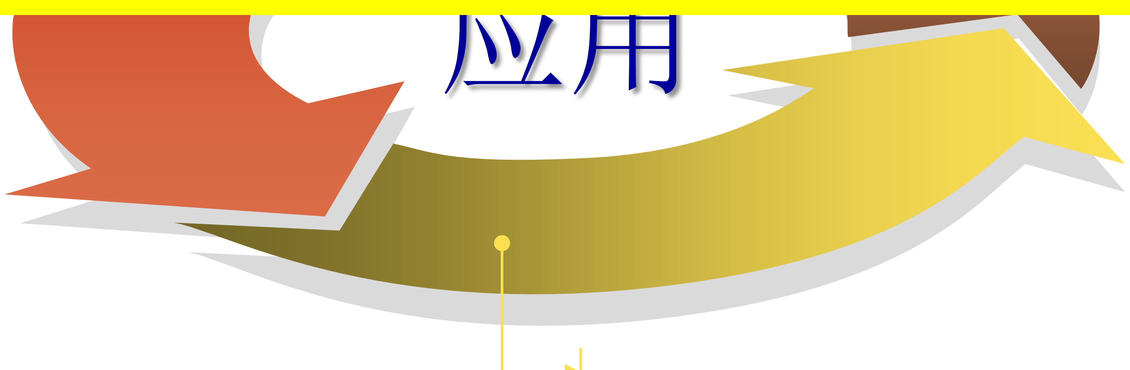
回归方程的应用



- 利用数学求极值的方法，寻找所得到的二次优化回归方程的局部最优解；
- 对回归方程求一阶偏导数，达到局部最优点时得到相应的编码水平，还原为因素的最优水平 。

回归方程的应用

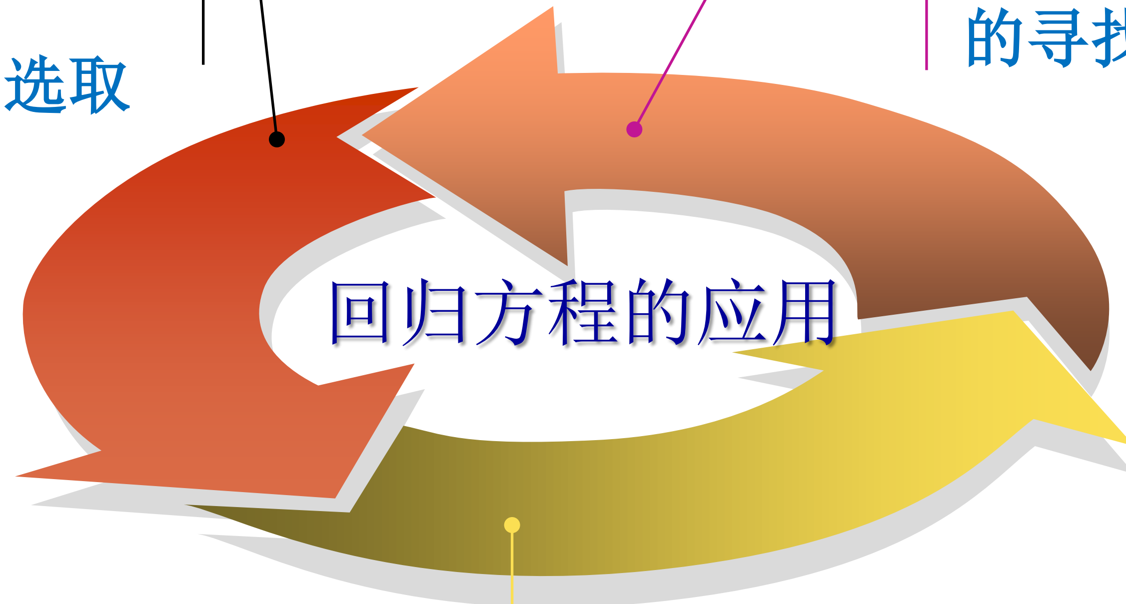
- 最佳试验方案并不一定是前面所寻求的局部最优点，
- 而是根据试验所耗费的人力、物力、财力，综合平衡，在保证试验质量和结果真实可靠的前提下，
- 寻求经济、高效的试验方案。



3. 寻求最佳试验方案

1. 在方程设计范围内最优试验因子的选取

2. 方程局部最优点的寻找



3. 寻求最佳试验方案

评价一个回归方程的精度，可以用预测值的方差来表示；

而预测值的方差与试验点在空间的位置及结构矩阵有关。

二次回归正交设计

- 试验规模小；
- 计算简便；
- 避免了回归系数间的相关性。

- 与一般回归分析一样，试验点在因子空间的位置不同（各因素所取水平不同），
- 对应的各个预测值 $\hat{y}$ 的方差也就不相同，
- 预测值的方差强烈依赖于试验点在因子空间的位置。

- 由于误差的干扰，致使在各个方向上不能提供等精度的估计，
- 不能对不同试验点预测值之间进行直接比较，
- 不易根据预测值寻找最优区域。

回归旋转设计

回归旋转设计

(regression-rotata design)

回归旋转设计

一次回归旋转设计

二次回归旋转设计

三次回归旋转设计

所谓旋转性是指试验因素空间中与试验中心距离相等的球面上各处理组合的预测值 $\hat{y}$ 的方差具有几乎相等的特性，具有这种性质的回归设计称回归旋转设计。

利用具有旋转性的回归方程进行预测时，对于同一球面上的点可直接比较其预测值的好坏，从而找出预测值较优区域。

以二元二次回归方程为例

在二个变量情况下，二次回归模型为：

$$Y_\alpha = \beta_0 + \beta_1 X_{\alpha 1} + \beta_2 X_{\alpha 2} + \beta_{12} X_{\alpha 1} X_{\alpha 2} + \beta_{11} X_{\alpha 1}^2 + \beta_{22} X_{\alpha 2}^2 + \varepsilon_\alpha$$

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2 + b_{11} x_1^2 + b_{22} x_2^2$$

（6项）结构矩阵为：

$$X = \begin{pmatrix} 1 & x_{11} & x_{12} & x_{11}x_{12} & x_{11}^2 & x_{12}^2 \\ 1 & x_{21} & x_{22} & x_{21}x_{22} & x_{21}^2 & x_{22}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{N1} & x_{N2} & x_{N1}x_{N2} & x_{N1}^2 & x_{N2}^2 \end{pmatrix}$$

信息矩阵$A = X'X$

$$X = \begin{pmatrix} 1 & x_{11} & x_{12} & x_{11}x_{12} & x_{11}^2 & x_{12}^2 \\ 1 & x_{21} & x_{22} & x_{21}x_{22} & x_{21}^2 & x_{22}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{N1} & x_{N2} & x_{N1}x_{N2} & x_{N1}^2 & x_{N2}^2 \end{pmatrix}$$

$$X' = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \cdots & x_{N1} \\ x_{12} & x_{22} & \cdots & x_{N2} \\ x_{11}x_{12} & x_{21}x_{22} & \cdots & x_{N1}x_{N2} \\ x_{11}^2 & x_{21}^2 & \cdots & x_{N1}^2 \\ x_{12}^2 & x_{22}^2 & \cdots & x_{N2}^2 \end{pmatrix}$$

对应的信息矩阵 A（6阶对称方阵） 为：

$$
A = X'X = \begin{pmatrix}
N & \sum x_{\alpha 1} & \sum x_{\alpha 2} & \sum x_{\alpha 1}x_{\alpha 2} & \sum x_{\alpha 1}^2 & \sum x_{\alpha 2}^2 \\
 & \sum x_{\alpha 1}^2 & \sum x_{\alpha 1}x_{\alpha 2} & \sum x_{\alpha 1}^2 x_{\alpha 2} & \sum x_{\alpha 1}^3 & \sum x_{\alpha 1}x_{\alpha 2}^2 \\
 & & \sum x_{\alpha 2}^2 & \sum x_{\alpha 1}x_{\alpha 2}^2 & \sum x_{\alpha 1}^2 x_{\alpha 2} & \sum x_{\alpha 2}^3 \\
 & & & \sum x_{\alpha 1}^2 x_{\alpha 2}^2 & \sum x_{\alpha 1}^3 x_{\alpha 2} & \sum x_{\alpha 1}x_{\alpha 2}^3 \\
 & \text{对称部分} & & & \sum x_{\alpha 1}^4 & \sum x_{\alpha 1}^2 x_{\alpha 2}^2 \\
 & & & & & \sum x_{\alpha 2}^4
\end{pmatrix}
$$

α的含义：试验点从第1个到第N个

$$A = X'X = \begin{pmatrix} N & \sum x_{\alpha 1} & \sum x_{\alpha 2} & \sum x_{\alpha 1}x_{\alpha 2} & \sum x_{\alpha 1}^2 & \sum x_{\alpha 2}^2 \\ & \sum x_{\alpha 1}^2 & \sum x_{\alpha 1}x_{\alpha 2} & \sum x_{\alpha 1}^2 x_{\alpha 2} & \sum x_{\alpha 1}^3 & \sum x_{\alpha 1}x_{\alpha 2}^2 \\ & & \sum x_{\alpha 2}^2 & \sum x_{\alpha 1}x_{\alpha 2}^2 & \sum x_{\alpha 1}^2 x_{\alpha 2} & \sum x_{\alpha 2}^3 \\ & & & \sum x_{\alpha 1}^2 x_{\alpha 2}^2 & \sum x_{\alpha 1}^3 x_{\alpha 2} & \sum x_{\alpha 1}x_{\alpha 2}^3 \\ & \text{对称部分} & & & \sum x_{\alpha 1}^4 & \sum x_{\alpha 1}^2 x_{\alpha 2}^2 \\ & & & & & \sum x_{\alpha 2}^4 \end{pmatrix}$$

信息矩阵A中的所有元素，可以下式表示：

$$\sum_{\alpha} X_{\alpha 1}^{a_1} X_{\alpha 2}^{a_2}$$

式中 x 指数a_1、a_2 可分别取0、1、2、3、4等数，但它们的和不超过 4，即

$$0 \le a_1 + a_2 \le 4$$

$$A = X'X = \begin{pmatrix} N & \sum X_{\alpha 1} & \sum X_{\alpha 2} & \sum X_{\alpha 1}X_{\alpha 2} & \sum X_{\alpha 1}^2 & \sum X_{\alpha 2}^2 \\ & \sum X_{\alpha 1}^2 & \sum X_{\alpha 1}X_{\alpha 2} & \sum X_{\alpha 1}^2X_{\alpha 2} & \sum X_{\alpha 1}^3 & \sum X_{\alpha 1}X_{\alpha 2}^2 \\ & & \sum X_{\alpha 2}^2 & \sum X_{\alpha 1}X_{\alpha 2}^2 & \sum X_{\alpha 1}^2X_{\alpha 2} & \sum X_{\alpha 2}^3 \\ & & & \sum X_{\alpha 1}^2X_{\alpha 2}^2 & \sum X_{\alpha 1}^3X_{\alpha 2} & \sum X_{\alpha 1}X_{\alpha 2}^3 \\ & \text{对称部分} & & & \sum X_{\alpha 1}^4 & \sum X_{\alpha 1}^2X_{\alpha 2}^2 \\ & & & & & \sum X_{\alpha 2}^4 \end{pmatrix}$$

$$\sum_{\alpha} X_{\alpha 1}^{a_1} X_{\alpha 2}^{a_2}$$

$$a_1 = a_2 = 0$$

$$a_1 = 0,\ a_2 = 4$$

m个因素d次回归方程中，共有C_{m+d}^{d}项

对应的信息矩阵是 C_{m+d}^{d}阶对称方程

矩阵中的元素的一般形式为 $\sum_{\alpha} x_{\alpha 1}^{a_1} x_{\alpha 2}^{a_2} \cdots x_{\alpha m}^{a_m}$

$$a_1, a_2, \cdots, a_m \to (0,1,2,\cdots,2d)$$

$$0 \le a_1 + a_2 + \cdots a_m \le 2d$$

m个因素d次回归方程中，

旋转设计

回归设计旋转条件

回归设计非退化条件

$$A = X'X = \begin{pmatrix} N & \sum X_{\alpha 1} & \sum X_{\alpha 2} & \sum X_{\alpha 1}X_{\alpha 2} & \sum X_{\alpha 1}^2 & \sum X_{\alpha 2}^2 \\ & \sum X_{\alpha 1}^2 & \sum X_{\alpha 1}X_{\alpha 2} & \sum X_{\alpha 1}^2 X_{\alpha 2} & \sum X_{\alpha 1}^3 & \sum X_{\alpha 1}X_{\alpha 2}^2 \\ & & \sum X_{\alpha 2}^2 & \sum X_{\alpha 1}X_{\alpha 2}^2 & \sum X_{\alpha 1}^2 X_{\alpha 2} & \sum X_{\alpha 2}^3 \\ & & & \sum X_{\alpha 1}^2 X_{\alpha 2}^2 & \sum X_{\alpha 1}^3 X_{\alpha 2} & \sum X_{\alpha 1}X_{\alpha 2}^3 \\ & \text{对称部分} & & & \sum X_{\alpha 1}^4 & \sum X_{\alpha 1}^2 X_{\alpha 2}^2 \\ & & & & & \sum X_{\alpha 2}^4 \end{pmatrix}$$

$$\sum_{\alpha} X_{\alpha 1}^{a_1} X_{\alpha 2}^{a_2}$$

旋转性

指数 a_1、a_2中至少有一个为奇数	指数a_1、a_2都是偶数或者为零

$$\sum_{a} X_{\alpha 1}^{a_1} X_{\alpha 2}^{a_2} \cdots\cdots X_{\alpha p}^{a_p} = 0$$

$$\sum_{a} x_{\alpha j}^{2} = \lambda_2 N$$

$$\sum_{a} x_{\alpha j}^{4} = 3\sum x_{\alpha i}^{2} \alpha_{\alpha j}^{2} = 3\lambda_4 N$$

***i*，*j* =1，2···, *m*;** λ_α 待定参数

为了使旋转设计成为可能，

要求信息矩阵 **A** 不退化，为此，必须有不等式：

$$\frac{\lambda_4}{\lambda_2^2} \neq \frac{p}{P+2}$$

- 已证明，只要使 **N** 个试验点不在同一个球面上，就能满足非退化条件。
- 或者说只要使 **N** 个试验点至少分布于两个半径不等的球面上，就有可能获得旋转设计方案。

P元二次旋转设计方案

旋转性条件

非退化条件

- 旋转性条件是旋转设计的必要条件；
- 非退化条件是使旋转设计成为可能的充分条件；
- 二者结合起来才能构成旋转设计方案。
- 该方案在实际操作时主要借助于组合设计来实现。

P元二次旋转设计方案

旋转性条件

非退化条件

组合设计中 N 个试验点，

$$N = m_c + m_\gamma + m_0$$

分布在三个半径不相等的球面上，

它们有固定的组合搭配。

P元二次
旋转设计方案

采用组合设计选取的试验点，完全能够满足非退化条件。

旋转性条件

非退化条件

$$N = m_c + m_\gamma + m_0$$

m_c 个点分布在半径为 $\rho_c = \sqrt{P}$ 的球面上；

m_γ 个点分布在半径为 $\rho_\gamma = \gamma$ 的球面上；

m_0 个点分布在半径为 $\rho_0 = 0$ 的球面上；

采用组合设计，其信息矩阵 **A** 的元素中：

$$\sum_{a} x_{\alpha j} = \sum_{\alpha} x_{\alpha i} x_{\alpha j}^{2} = \sum_{\alpha} x_{\alpha i}^{2} x_{\alpha j} = 0$$

偶次方元素：

$$\sum_{a} x_{\alpha i}^{2} = m_{c} + 2\gamma^{2} \neq 0$$

$$\sum_{a} x_{\alpha i}^{4} = m_{c} + 2\gamma^{4} \neq 0$$

$$\sum_{a} x_{\alpha i}^{2} x_{\alpha j}^{2} = m_{c} \neq 0$$

完全符合旋转性条件的要求。

为了获得旋转设计方案，可根据旋转性条件确定 γ 值：

$$\sum x_{\alpha j}^{4} = 3\sum x_{\alpha i}^{2} x_{\alpha j}^{2}$$

求出 γ 就可以了。在组合设计下，当 $m_c = 2^P$（全实施），则上式变为：

$$2^P + 2\gamma^4 = 3 \times 2^P$$

解此方程，即可建立全实施时 γ 值的计算式，即：

$$\gamma = 2^{m/4}$$

$m_c = 2^{m-1}$（1 / 2实施）　$\gamma = 2^{(m-1)/4}$

$m_c = 2^{m-2}$（1 / 4实施）　$\gamma = 2^{(m-2)/4}$

$m_c = 2^{m-3}$（1 / 8实施）　$\gamma = 2^{(m-3)/4}$

$m = 2,\ m_c = 2^2,$

$\gamma = 2^{2/4} = 1.414$

表1 m个因素不同实施下二次正交旋转组合设计参数表

m	m_c	m_γ	m_0	N	γ
2 (全实施)	4	4	8	16	1.414
3 (全实施)	8	6	9	23	1.682
4 (全实施)	16	8	12	36	2.000
5 (全实施)	32	10	17	59	2.378
5(1/2实施)	16	10	10	36	2.000
6(1/2实施)	32	12	15	59	2.378
6(1/4实施)	16	12	8	36	2.000
7(1/2实施)	64	14	22	100	2.828
7(1/4实施)	32	14	13	59	2.378
8(1/2实施)	128	16	33	177	3.364
8(1/4实施)	64	16	20	100	2.828
8(1/8实施)	32	16	11	59	2.378

P元二次
旋转设计方案

旋转性条件

非退化条件

$$N = m_c + m_\gamma + m_0$$

分布在三个半径不相等的球面上：

m_c 个点分布在半径为 $\rho_c = \sqrt{P}$ 的球面上；

m_γ 个点分布在半径为 $\rho_\gamma = \gamma$ 的球面上；

m_0 个点分布在半径为 $\rho_0 = 0$ 的球面上；

回归旋转设计

$$N = m_c + m_\gamma + m_0$$

- 通常，中心点不安排试验，不会影响旋转性；
- 中心点附近区域往往是试验者所关心的区域，
- 同时选择适当的中心点还可使二次旋转组合设计具有正交性。

正交性的获得

在二次旋转组合计划中，

一次项和交互项的回归系数 b_j，b_{ij} 仍保持正交，

但 b_0 与 b_{jj} 之间，以及 b_{ii} 与 b_{jj} 之间都存在相关，

即不具正交性，

它们之间的相关矩分别为：

$$\operatorname{cov}(b_0, b_{jj}) = -2\lambda_2\lambda_4 t\sigma^2 / N$$

$$\operatorname{cov}(b_{ii}, b_{jj}) = (\lambda_2^2 - \lambda_4)t\sigma^2 / N$$

消除 b_0 与 b_{jj} 之间的相关性，只需对平方项进行中心化变换，即

$$X'_{\alpha j} = X_{\alpha j}^2 - \frac{1}{N}\sum_{i=1}^{N} X_{\alpha j}^2$$

平方项之间相关性的消除，需使

$$\operatorname{cov}(b_{ii}, b_{jj}) = (\lambda_2^2 - \lambda_4) t\sigma^2 / N$$

$\lambda_4 = \lambda_2^2$ 或者 $\lambda_4 / \lambda_2^2 = 1$ 就会使

$\operatorname{cov}(b_{ii}, b_{jj}) = 0$ 即消除了相关性

$$\lambda_4 / \lambda_2^2 = 1$$

$$\frac{\lambda_4}{\lambda_2^2} = \frac{m_c m + 2\gamma^4}{(m_c + 2\gamma^2)^2} \cdot \frac{N}{m + 2}$$

对于m个因素的二次旋转组合设计中的 m、m_c、γ **都是固定的，只有适当调整** N**，才能** $\lambda_4 / \lambda_2^2 = 1$**，**

$$N = m_c + m_\gamma + m_0$$

$$\frac{N}{\lambda_4 / \lambda_2^2} = \frac{(m_c + 2\gamma^2)^2 (m + 2)}{m_c \times m + 2\gamma^2}$$

表1 m个因素不同实施下二次正交旋转组合设计参数表

m	m_c	m_γ	m_0	N	γ
2 (全实施)	4	4	8	16	1.414
3 (全实施)	8	6	9	23	1.682
4 (全实施)	16	8	12	36	2.000
5 (全实施)	32	10	17	59	2.378
5(1/2实施)	16	10	10	36	2.000
6(1/2实施)	32	12	15	59	2.378
6(1/4实施)	16	12	8	36	2.000
7(1/2实施)	64	14	22	100	2.828
7(1/4实施)	32	14	13	59	2.378
8(1/2实施)	128	16	33	177	3.364
8(1/4实施)	64	16	20	100	2.828
8(1/8实施)	32	16	11	59	2.374

适当地选择m_0，就能使二次旋转组合设计具有一定的正交性。

表2 三个因素正交旋转组合设计结构矩阵

试验号	x_0	x_1	x_2	x_3	x_1x_1	x_1x_3	x_2x_3	x_1'	x_2'	x_3'
1	1	1	1	1	1	1	1	0.406	0.406	0.406
2	1	1	1	-1	1	-1	-1	0.406	0.406	0.406
3	1	1	-1	1	-1	1	-1	0.406	0.406	0.406

平方项中心化变化：

消除常数项β_0的估计值b_0与平方项β_{jj}的估计值b_{jj}的相关性

试验号	x_0	x_1	x_2	x_3	x_1x_1	x_1x_3	x_2x_3	x_1'	x_2'	x_3'
9	1	1.682	0	0	0	0	0	2.235	-0.594	-0.594
10	1	-1.682	0	0	0	0	0	2.235	-0.594	-0.594
11	1	0	1.682	0	0	0	0	-0.594	2.235	-0.594

$$x'_{aj} = x_{aj}^2 - \frac{\sum x_{aj}^2}{N} = x_{aj}^2 - \frac{13.65825}{23} = x_{aj}^2 - 0.594$$

试验号	x_0	x_1	x_2	x_3	x_1x_1	x_1x_3	x_2x_3	x_1'	x_2'	x_3'
16	1	0	0	0	0	0	0	-0.594	-0.594	-0.594
17	1	0	0	0	0	0	0	-0.594	-0.594	-0.594
18	1	0	0	0	0	0	0	-0.594	-0.594	-0.594
19	1	0	0	0	0	0	0	-0.594	-0.594	-0.594
20	1	0	0	0	0	0	0	-0.594	-0.594	-0.594
21	1	0	0	0	0	0	0	-0.594	-0.594	-0.594
22	1	0	0	0	0	0	0	-0.594	-0.594	-0.594
23	1	0	0	0	0	0	0	-0.594	-0.594	-0.594

$$N = m_c + m_\gamma + m_0$$

γ, m_0

回归组合设计

回归旋转设计

二次旋转组合设计的**通用性**

二次回归旋转组合设计，

具有同一球面上各试验点的预测值 $\hat{y}$ 的方差相等的优点，

但不同半径球面上各试验点的预测值 $\hat{y}$ 的方差不等。

为了解决这一问题，于是提出了旋转设计的通用性问题。

二次旋转组合设计的通用性

所谓“通用性”，就是试验除了仍保持其旋转性外，还具有各试验点与中心的距离 ρ 在因子空间编码值区间 $0<\rho<1$ 的范围内，其预测值 $\hat{y}$ 的方差基本相等的性质，即同时具有旋转性与通用性。

这种设计称为通用旋转组合设计。

只有恰当确定 λ_4，才能满足通用性的要求。

预测值 $\hat{y}$ 的方差

$$D(\hat{y}) = \frac{(m+2)\sigma^2}{[(m+2)\lambda_4 - m](N/\lambda_4)} \times \left[1 + \frac{\lambda_4 - 1}{\lambda_4}\rho^2 + \frac{(m+1)\lambda_4 - (m-1)}{2\lambda_4^2(m+2)}\rho^4\right]$$

对 λ_4 的要求

使上式中 ρ_i 处的值与 $\rho = 0$ 处值相差的平方和为最小。

即：

$$Q(\lambda_4) = f_0^2(\lambda_4)\sum_{i=1}^{n}\left[f_1(\lambda_4)\rho_i^2 + f_2(\lambda_4)\rho_i^4\right]^2 = \text{最小}$$

式中：

$$f_0(\lambda_4) = \frac{m+2}{\left[(m+2)\lambda_4 - m\right](N/\lambda_4)}$$

$$f_1(\lambda_4) = \frac{\lambda_4 - 1}{\lambda_4}$$

$$f_2(\lambda_4) = \frac{(m+1)\lambda_4 - (m-1)}{2\lambda_4^2(m+2)}$$

可对不同的 m，计算λ_4， λ_4确定后，可由关系式

$$N = \frac{\left(m_C + 2\gamma^2\right)^2 (m + 2)\lambda_4}{m_c m + 2\gamma^4}$$

计算出不同 m 的 处理数 ***N***， 当计算结果不是整数时， ***N*** 可取其最靠近的整数。然后由

$$m_0 = N - m_c - m_\gamma$$

计算出不同 m 值的 m_0。

表3　二次通用旋转组合设计参数表

m	m_c	m_γ	γ	λ_4	N	m_0
2（全实施）	4	4	1.414	0.81	13	5
3（全实施）	8	6	1.682	0.86	20	6
4（全实施）	16	8	2.000	0.86	31	7
4（1/2全实施）	8	8	1.682	0.86	20	4
5（1/2全实施）	16	10	2.000	0.89	32	6
6（1/2全实施）	32	12	2.378	0.90	53	9
7（1/2全实施）	64	14	2.828	0.92	92	14
8（1/2全实施）	128	16	3.364	0.93	165	21
8（1/4全实施）	64	16	2.828	0.93	93	13

为了满足通用性要求，主要在于确定出适当的 m_0 ,因此，只要在中心点安排如上表所列的 m_0 次试验，即可使二次旋转组合设计获得通用性。

通常，通用旋转设计的m_0有所减少。

但二次通用旋转组合设计中，常数项b_0与平方项回归系数b_{jj} 、平方项回归系数间b_{jj}还存在相关，因此，通用旋转设计是牺牲了部分正交性而达到一致精度的要求。

- 正交旋转的优点在于正交性，
- 它是通过增加中心点的试验次数得到的，但有时并不合算。
- 在某些实际问题中，反倒不如选用通用旋转设计。因为通用旋转设计，既能在 $0<\rho<1$ 的较实用区域使预测值基本不变，又在一定程度上减少了试验次数。

二次正交旋转组合设计及其统计分析

设研究因素为 m 个，分别以 $Z_1, Z_2, \cdots, Z_m$, 表示。

(1) 首先确定每个因素的上、下水平，进而计算零水平，以及变化间距。

某因素零水平及变化间距的计算式为

$$Z_{0j}=(Z_{1j}+Z_{2j})/2 \qquad \Delta_j=(Z_{2j}-Z_{0j})/\gamma$$

式中 γ 为待定参数，其值可以查出。

对每个因素 Z_j 各水平的取值进行线性变换，以实现其编码

$$x_{\alpha j}=(Z_{\alpha j}-Z_{0j})/\Delta_j$$

这样，就将有单位的自然变量 Z_j 变成了无单位的规范变量 x_j（$j=1, 2, \cdots, m$），并可编制出因素水平的编码值表。

表5　二次正交旋转设计因素水平编码值表

编码	Z_1	Z_2	…	Z_m
$+\gamma$	Z_{21}	Z_{22}	…	Z_{2m}
$+1$	$Z_{01}+\Delta_1$	$Z_{02}+\Delta_2$	…	$Z_{0m}+\Delta_m$
0	Z_{01}	Z_{02}	…	Z_{0m}
-1	$Z_{01}-\Delta_1$	$Z_{02}-\Delta_2$	…	$Z_{0m}-\Delta_m$
$-\gamma$	Z_{11}	Z_{12}	…	Z_{1m}

试验因素 $Z_1, Z_2, \cdots, Z_m$ 经因素水平编码后，

以变量 $x_1, x_2, \cdots, x_m$ 表示，

选用适当的二水平正交表，

即可设计出二次回归正交旋转组合方案。

表6　二元二次正交旋转组合设计的结构矩阵

处理号		x_0	x_1	x_2	x_1x_2	x_1'	x_2'
m_c	1	1	1	1	1	0.5	0.5
	2	1	1	-1	-1	0.5	0.5
	3	1	-1	1	-1	0.5	0.5
	4	1	-1	-1	1	0.5	0.5
m_γ	5	1	1.414	0	0	1.5	-0.5
	6	1	-1.414	0	0	1.5	-0.5
	7	1	0	1.414	0	-0.5	1.5
	8	1	0	-1.414	0	-0.5	1.5
m_0	9	1	0	0	0	-0.5	-0.5
	10	1	0	0	0	-0.5	-0.5
	11	1	0	0	0	-0.5	-0.5
	12	1	0	0	0	-0.5	-0.5
	13	1	0	0	0	-0.5	-0.5
	14	1	0	0	0	-0.5	-0.5
	15	1	0	0	0	-0.5	-0.5
	16	1	0	0	0	-0.5	-0.5
$\alpha_j = \sum x_{\alpha j}^2$		16	8	8	4	8	8

表7 三元二次正交旋转组合设计的结构矩阵

处理号		x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	x_1'	x_2'	x_3'
m_c	1	1	1	1	1	1	1	1	0.406	0.406	0.406
	2	1	1	1	-1	1	-1	-1	0.406	0.406	0.406
	3	1	1	-1	1	-1	1	-1	0.406	0.406	0.406
	4	1	1	-1	-1	-1	-1	1	0.406	0.406	0.406
	5	1	-1	1	1	-1	-1	1	0.406	0.406	0.406
	6	1	-1	1	-1	-1	1	-1	0.406	0.406	0.406
	7	1	-1	-1	1	1	-1	-1	0.406	0.406	0.406
	8	1	-1	-1	-1	1	1	1	0.406	0.406	0.406
m_γ	9	1	1.682	0	0	0	0	0	2.234	-0.594	-0.594
	10	1	-1.682	0	0	0	0	0	2.234	-0.594	-0.594
	11	1	0	1.682	0	0	0	0	-0.594	2.234	-0.594
	12	1	0	-1.682	0	0	0	0	-0.594	2.234	-0.594
	13	1	0	0	1.682	0	0	0	-0.594	-0.594	2.234
	14	1	0	0	-1.682	0	0	0	-0.594	-0.594	2.234

（未完）

（续前表）

处理号		x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	x_1'	x_2'	x_3'
m_0	15	1	0	0	0	0	0	0	-0.594	-0.594	-0.594
	16	1	0	0	0	0	0	0	-0.594	-0.594	-0.594
	17	1	0	0	0	0	0	0	-0.594	-0.594	-0.594
	18	1	0	0	0	0	0	0	-0.594	-0.594	-0.594
	19	1	0	0	0	0	0	0	-0.594	-0.594	-0.594
	20	1	0	0	0	0	0	0	-0.594	-0.594	-0.594
	21	1	0	0	0	0	0	0	-0.594	-0.594	-0.594
	22	1	0	0	0	0	0	0	-0.594	-0.594	-0.594
	23	1	0	0	0	0	0	0	-0.594	-0.594	-0.594
$\alpha_j=\sum x_{\alpha j}^2$		23	13.658	13.658	13.658	8	8	8	15.887	15.887	15.887

二次回归正交旋转组合设计试验结果的统计分析，

与二次回归正交组合设计试验结果的统计分析方法相似。

三因素实施正交旋转组合设计示例

表1 m个因素不同实施下二次正交旋转组合设计参数表

m	m_c	m_γ	m_0	N	γ
2 (全实施)	4	4	8	16	1.414
3 (全实施)	8	6	9	23	1.682
4 (全实施)	16	8	12	36	2.000
5 (全实施)	32	10	17	59	2.378
5(1/2实施)	16	10	10	36	2.000
6(1/2实施)	32	12	15	59	2.378
6(1/4实施)	16	12	8	36	2.000
7(1/2实施)	64	14	22	100	2.828
7(1/4实施)	32	14	13	59	2.378
8(1/2实施)	128	16	33	177	3.364
8(1/4实施)	64	16	20	100	2.828
8(1/8实施)	32	16	11	59	2.374

三因素实施正交旋转组合设计示例

采用三因素二次正交旋转设计组合设计，其试验因素水平编码如下表。

表8 试验因素水平编码表

编码	Z_1	Z_2	Z_3
+1.682	51.0	16.0	10000
+1	48.6	14.4	8580
0	45.0	12.0	6500
-1	41.4	9.6	4420
-1.682	39.0	8.0	3000
Δ_j			

（1）建立回归方程。三因素二次回归正交旋转组合设计结构矩阵与结果计算如下表。初步得回归方程为：

表9　三因素二次回归正交旋转组合设计结构矩阵与结果计算表

处理号	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	x_1'	x_2'	x_3	y
1	1	1	1	1	1	1	1	0.406	0.406	0.406	78
2	1	1	1	-1	1	-1	-1	0.406	0.406	0.406	84
3	1	1	-1	1	-1	1	-1	0.406	0.406	0.406	73
4	1	1	-1	-1	-1	-1	1	0.406	0.406	0.406	77
5	1	-1	1	1	-1	-1	1	0.406	0.406	0.406	81
6	1	-1	1	-1	-1	1	-1	0.406	0.406	0.406	88
7	1	-1	-1	1	1	-1	-1	0.406	0.406	0.406	80
8	1	-1	-1	-1	1	1	1	0.406	0.406	0.406	73
9	1	1.682	0	0	0	0	0	2.234	-0.594	-0.594	74
10	1	-1.682	0	0	0	0	0	2.234	-0.594	-0.594	71
11	1	0	1.682	0	0	0	0	-0.594	2.234	-0.594	86
12	1	0	-1.682	0	0	0	0	-0.594	2.234	-0.594	69
13	1	0	0	1.682	0	0	0	-0.594	-0.594	2.234	84
14	1	0	0	-1.682	0	0	0	-0.594	-0.594	2.234	80

（未完）

（续前表）

处理号	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	x_1'	x_2'	x_3	y	
15	1	0	0	0	0	0	0	-0.594	-0.594	-0.594	83	
16	1	0	0	0	0	0	0	-0.594	-0.594	-0.594	85	
17	1	0	0	0	0	0	0	-0.594	-0.594	-0.594	83	
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	78	
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	83	
20	1	0	0	0	0	0	0	-0.594	-0.594	-0.594	79	
2[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	
2[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	
23	1	0	0	0	0	0	0	-0.594	-0.594	-0.594	83	
$\alpha_j=\sum x_j^2$	23	13.658	13.658	13.658	8	8	8	15.887	15.887	15.887	$\sum y^2=147118$	
$\beta_j=\sum x_j y$	1836	-4.954	156.594	-3.272	-4	-10	-16	-46.524	-18.244	7.208	$SS_y=557.3044$	
$b_j=B_j/a$	79.8261	-0.3627	4.1437	-0.2396	-0.50000	-1.2500	-2.0000	-2.9284	-1.1484	0.4537	$SS_R=444.0500$	
$Q_j=B_j^2/a_j$	—	1.7969	234.5058	0.78839	2.0000	12.5000	32.0000	136.2424	20.9507	3.2703	$SS_r=113.2544$	

初步得回归方程为：

$$\hat{y}=79.8261-0.3627x_1+4.1437x_2-0.2396x_3-0.5000x_1x_2-1.2500x_1x_3-2.0000x_2x_3-2.9284x_1'-1.1484x_2'+0.4537x_3'$$

（2）回归方程的显著性测验：对所得三元二次回归方程 $\hat{y}$ 进行方差分析。

表10 三因素二次回归正交旋转组合设计试验结果方差分析表

变异来源	平方和SS	自由度df	均方MS	F值	F_α
x_1	1.7969	1	1.7969	＜1	$F_{0.01(1,13)}$=3.14
x_2	234.5058	1	234.5058	26.918**	$F_{0.05(1,13)}$=4.67
x_3	0.7839	1	0.7839	＜1	
x_1x_2	2.0000	1	2.0000	＜1	
x_1x_3	12.5000	1	12.5000	1.435	
x_2x_3	32.0000	1	32.0000	3.673**	
x_1'	136.2424	1	136.2424	15.639**	
x_2'	20.9507	1	20.9507	2.405	
x_3'	3.2703	1	3.2703	＜1	
回归	444.0500	?	?	5.663?	$F_{0.01(?,?)}$=4.17
剩余	113.2544	?	?		
误差	40.0000	?	?		
失拟	73.2544	?	?	2.930?	$F_{0.05(?,?)}$=3.69
总变异	557.3044	?			

剔除 x_1，x_3，x_1x_2，x_1x_3，x_2'和 x_3'，回归方程变为：

$$\hat{y}=79.8261+4.1437x_2-2.0000x_2x_3-2.9284x_1'$$

将中心化变换还原为 x_j^2，得：

$$\hat{y}=81.5656+4.1437x_2-2.0000x_2x_3-2.9284x_1^2$$

通用旋转组合设计及其统计分析

通用旋转组合设计与正交旋转组合设计基本相同，其组合计划中试验处理组合数 N，也是由 3 部分组成，即：

$$N = m_c + m_\gamma + m_0$$

上式中 m_c 和 m_γ 的数值与正交旋转组合设计完全相同，只是 N 和 m_0 有所不同，其值可由表查出。

表11 三元二次通用旋转组合设计的结构矩阵

处理号		x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	x_1'	x_2'	x_3'
m_c	1	1	1	1	1	1	1	1	1	1	1
	2	1	1	1	-1	1	-1	-1	1	1	1
	3	1	1	-1	1	-1	1	-1	1	1	1
	4	1	1	-1	-1	-1	-1	1	1	1	1
	5	1	-1	1	1	-1	-1	1	1	1	1
	6	1	-1	1	-1	-1	1	-1	1	1	1
	7	1	-1	-1	1	1	-1	-1	1	1	1
	8	1	-1	-1	-1	1	1	1	1	1	1
m_γ	9	1	1.682	0	0	0	0	0	2.282	0	0
	10	1	-1.682	0	0	0	0	0	2.282	0	0
	11	1	0	1.682	0	0	0	0	0	2.282	0
	12	1	0	1.682	0	0	0	0	0	2.282	0
	13	1	0	0	1.682	0	0	0	0	0	2.282
	14	1	0	0	-1.682	0	0	0	0	0	2.282
m_0	15	1	0	0	0	0	0	0	0	0	0
	16	1	0	0	0	0	0	0	0	0	0
	17	1	0	0	0	0	0	0	0	0	0
	18	1	0	0	0	0	0	0	0	0	0
	19	1	0	0	0	0	0	0	0	0	0
	20	1	0	0	0	0	0	0	0	0	0

四元二次通用旋转组合

鸡肉乳酸发酵试验，对鸡肉乳酸发酵的产酸条件（盐浓度、糖浓度、发酵温度和发酵时间）进行优化试验，

采用四元二次通用旋转组合试验寻求最优发酵条件。

表3 二次通用旋转组合设计参数表

m	m_c	m_γ	γ	λ_4	N	m_0
2（全实施）	4	4	1.414	0.81	13	5
3（全实施）	8	6	1.682	0.86	20	6
4（全实施）	16	8	2.000	0.86	31	7
4（1/2全实施）	8	8	1.682	0.86	20	4
5（1/2全实施）	16	10	2.000	0.89	32	6
6（1/2全实施）	32	12	2.378	0.90	53	9
7（1/2全实施）	64	14	2.828	0.92	92	14
8（1/2全实施）	128	16	3.364	0.93	165	21
8（1/4全实施）	64	16	2.828	0.93	93	13

四元二次通用旋转组合

试验因素及水平编码见下表。

表13　鸡肉乳酸发酵产酸条件的四元二次通用旋转组合设计因素水平表

编码	盐浓度 x_1 /%	糖浓度 x_2 /%	发酵温度 x_3 /℃	发酵时间 x_4 /h
+2	8.0	6.0	37.0	48
+1	7.0	5.0	34.0	44
0	6.0	4.0	31.0	40
-1	5.0	3.0	28.0	36
-2	4.0	2.0	25.0	32

表14 鸡肉乳酸发酵产酸条件的四元二次通用旋转组合设计方案及结果

处理号	x_1	x_2	x_3	x_4	含酸量 y_α / %
1	1	1	1	1	0.654
2	1	1	1	-1	0.433
3	1	1	-1	1	0.538
4	1	1	-1	-1	0.321
5	1	-1	1	1	0.314
6	1	-1	1	-1	0.279
7	1	-1	-1	1	0.295
8	1	-1	-1	-1	0.242
9	-1	1	1	1	0.779
10	-1	1	1	-1	0.594
11	-1	1	-1	1	0.710
12	-1	1	-1	-1	0.529
13	-1	-1	1	1	0.481
14	-1	-1	1	-1	0.307
15	-1	-1	-1	1	0.328

（未完）

（续前表）

处理号	x_1	x_2	x_3	x_4	含酸量 y_α / %
16	-1	-1	-1	-1	0.291
17	2	0	0	0	0.125
18	-2	0	0	0	0.648
19	0	2	0	0	0.785
20	0	-2	0	0	0.213
21	0	0	2	0	0.429
22	0	0	-2	0	0.198
23	0	0	0	2	0.842
24	0	0	0	-2	0.486
25	0	0	0	0	0.797
26	0	0	0	0	0.709
27	0	0	0	0	0.759
28	0	0	0	0	0.694
29	0	0	0	0	0.728
30	0	0	0	0	0.738
31	0	0	0	0	0.746

（1）建立四元二次回归方程。

$$\hat{y}=0.7448-0.0829x_1+0.1319x_2+0.0437x_3+0.0786x_4-$$

$$0.0243x_1x_2-0.0012x_1x_2-0.0032x_1x_4+0.0086x_2x_3+$$

$$0.0316x_2x_4+0.0079x_3x_4-0.0934x_1^2-0.0652x_2^2-$$

$$0.1116x_3^2-0.0239x_4^2$$

（2）回归方程的显著性检验。

变异原因	平方和SS	自由度df	均方MS	F值	显著程度
x_1	0.16484	1	0.16484	49.28?	$F_{0.01(1,?)}$=8.53
x_2	0.41738	1	0.41738	127.79?	
x_3	0.04585	1	0.04585	13.71?	
x_4	0.13726	1	0.13726	41.04?	
$x_1 x_2$	0.00946	1	0.00946	2.83	
$x_1 x_3$	0.00002	1	0.00002	＜1	
$x_1 x_4$	0.00016	1	0.00016	＜1	
$x_2 x_3$	0.00117	1	0.00117	＜1	
$x_2 x_4$	0.01594	1	0.01594	4.77?	$F_{0.05(1,?)}$=4.49
$x_3 x_4$	0.00101	1	0.00101	＜1	
x_1'	0.16884	1	0.16884	50.48?	
x_2'	0.07959	1	0.07959	23.79?	
x_3'	0.34411	1	0.34411	102.88?	
x_4'	0.01648	1	0.01648	4.93?	
回归	1.40211		0.10015	29.94?	$F_{0.01(?,?)}$=3.56
剩余	0.05352		0.00334		
误差	0.00853		0.00142		
失拟	0.04499		0.00450	3.17	$F_{0.05(?,?)}$=4.74
总变异	1.45563				

多因素试验

正交试验设计

一次回归正交设计

二次回归正交设计

回归旋转设计

明确试验目的，确定试验指标。

选因素、定水平，列因素水平表

选择合适的正交表

表头设计 交互

编制方案，实施

统计分析
- 直观分析
- 方差分析

多因素试验

正交试验设计

一次回归正交设计

二次回归正交设计

回归旋转设计

明确试验目的，确定试验指标。

选因素、定水平，列因素水平表

二水平（上、下），因素水平编码

选择合适的正交表

表头设计　交互

编制方案，实施

统计分析
- 直观分析
- 方差分析

建立回归方程

方程检验

回归系数检验

失拟性检验

多因素试验

正交试验设计

一次回归正交设计

二次回归正交设计

回归旋转设计

明确试验目的，确定试验指标。

选因素、定水平，列因素水平表

二水平（上、下），因素水平编码

选择合适的正交表

γ点试验

表头设计

$$N = m_c + m_\gamma + m_0$$

正交性的实现

编制方案，实施

统计分析

直观分析

方差分析

建立回归方程

方程检验

回归系数检验

失拟性检 验

多因素试验

正交试验设计

一次回归正交设计

二次回归正交设计

回归旋转设计

明确试验目的，确定试验指标。

选因素、定水平，列因素水平表

二水平（上、下），因素水平编码

选择合适的正交表

γ点试验

$$N = m_c + m_\gamma + m_0$$

表头设计

正交性的实现

编制方案，实施

统计分析
- 直观分析
- 方差分析

建立回归方程

方程检验

回归系数检验

失拟性检验

注意问题

◆(1) 试验因子只属于一些可控或部分可控因子，对非可控因子则难以进行试验。同于试验水平有严格的要求，因子必须是数量化的；

◆(2) 设计是已组装好的，虽然满足了正交性和旋转性的要求，但缺乏灵活性，因而应用起来较为死板；

◆(3) 试验本身不设重复，除了试验组合中m_0有一定的重复外，m_c、m_γ均为单个试验点，所以，试验要求精度较高，要尽量控制其它因子的影响为一致。同时，由于没有重复，要求每一个试验都不能失败，否则无法进行缺区估计；

◆(4) 由于设计的特性及分析方法的要求，这种试验一般很少考虑中间过程，也就是只需要因子的输入，因而这种模型是一种静态模型，当其它因子变化时，按照模型不能立即适应。因此，应用试验结果时，往往会有一定的偏差，这就是为什么有些成果试验时效果很好，待推广应用时不一定能取得良好效果的原因之一。

应用问题

1 关于试验设计

多元二次回归正交旋转组合设计的结构矩阵虽然不是唯一的，但也基本上是固定不变的，一旦定好因子数目，它的试验设计的结构矩阵也就基本固定了。

按照设计要求，四因子以下的试验均为全实施方案，而五因子以上则可以用1/2实施方案，八因子以上则可用1/4实施方案。这样就存在一个方案选择问题。

因素数m	正交表	表头设计	m_c	m_γ =2m	m_0	N	q
2	$L_4(2^3)$	1,2	4	4	1	9	6
3	$L_8(2^7)$	1,2,4	6	6	1	15	10
4	$L_{16}(2^{15})$	1,2,4,8	16	8	1	25	15
5	$L_{32}(2^{31})$	1,2,4,8,16	32	10	1	43	21
5(1/2实施)	$L_{16}(2^{15})$	1,2,4,8,15	16	10	1	27	21

应用问题

表3 γ 值表

m_0	m							
	2	3	4	5(1/2实施)	5	6(1/2实施)	6	7(1/2实施)
1	1.00000	1.21541	1.41421	1.54671	1.59601	1.72443	1.76064	1.88488
2	1.07809	1.28719	1.48258	1.60717	1.66183	1.78419	1.82402	1.94347
3	1.14744	1.35313	1.54671	1.66443	1.72443	1.84139	1.88488	2.00000
4	1.21000	1.41421	1.60717	1.71885	1.78419	1.89629	1.94347	2.05464
5	1.26710	1.47119	1.66443	1.77074	1.84139	1.94910	2.00000	2.10754
6	1.31972	1.52465	1.71885	1.82036	1.89629	2.00000	2.05464	2.15884
7	1.36857	1.57504	1.77074	1.86792	1.94910	2.04915	2.10754	2.20866
8	1.41421	1.62273	1.82036	1.91361	2.00000	2.09668	2.15884	2.25709
9	1.45709	1.66803	1.86792	1.95759	2.04915	2.14272	2.20866	2.30424
10	1.49755	1.71120	1.91361	2.00000	2.09668	2.18738	2.25709	2.35018
11	1.53587	1.75245	1.95759	2.04096	2.14272	2.23073	2.30424	2.39498

应用问题

2 因子的选择

多元二次回归正交旋转组合设计因子的选择，需要考虑因子本身、组合设计要求和因子编码是否可行等。

从**试验目的**出发，应选择对目标性状影响大的可控因子或特殊试验目的的因子作为试验因子。例如，小麦栽培体系试验可以考虑将施氮量、旋磷量、播种密度、灌水量、有机肥量、施肥时期、播期、中耕次数、灌水次数、根外喷施微肥浓度、喷施时期、收获时期等因子作为试验因子。

从**设计要求**出发，象品种、施肥种类等非数值的因子则不能作为试验因子。

对于γ不是整数的组合设计，象灌水次数、中耕次数等则不能作为试验因子，而播期、施肥时间等可将日期转化为天数，根据γ的计算结果可近似取某日期作为试验的，一般来说，实施这类因子不会遇到太大的困难。

3 模型建立

注意问题

进行多元二次回归正交旋转组合设计试验，希望建立良好的回归模型，但有相当一些试验建立的模型达不到显著标准。

- （1）设计本身不设重复，小区数目少，误差自由度小，显著水平高，不易达到显著标准；
- （2）因子取值不当：因子取值范围太小，其试验效应往往较小，获得的回归方程就不会显著；因子取值范围太大或最佳点偏离，虽然有时也会使方程显著，但不可能捕捉到真正的最佳点；
- （3）试验质量不高，导致试验误差增大，使方程达不到显著水平。

响应面分析

在多因素数量处理试验的分析中，

可以分析试验指标(依变量)与多个试验因素(自变量)间

的回归关系，

这种回归可能是曲线或曲面的关系。

响应面分析

- 在响应面分析中，首先要得到回归方程：

$$\hat{y} = f(x_1,\ x_2,\cdots,\ x_l)$$

然后通过对自变量 $x_1,\ x_2,\cdots,\ x_l$ 的合理取值，求得使 $\hat{y} = f(x_1,\ x_2,\cdots,\ x_l)$ 最优的值，这就是响应面分析的目的。

有一大麦氮磷肥配比试验，

施氮肥量为每亩尿素0, 3, 6, 9, 12, 15, 18kg 7个水平，

施磷肥量为每亩过磷酸钙0, 7, 14, 21, 28, 35, 42kg 7个水平，

共49个处理组合，试验结果如下。

试进行产量对于氮、磷施肥量的响应面分析。

磷 肥	氮 肥						
	0	3	6	9	12	15	18
0	86.9	162.5	216.4	274.7	274.3	301.4	270.3
7	110.4	204.4	276.7	342.8	343.4	368.4	335.1
14	134.3	238.9	295.9	363.3	361.7	345.4	351.5
21	162.5	275.1	325.3	336.3	381.0	362.4	382.2
28	158.2	237.9	320.5	353.7	369.5	388.2	355.3
35	144.3	204.5	286.9	322.5	345.9	344.6	353.5
42	88.7	192.5	219.9	278.0	319.1	290.5	281.2

磷 肥	氮 肥						
	0	3	6	9	12	15	18
0	86.9	162.5	216.4	274.7	274.3	301.4	270.3
7	110.4	204.4	276.7	342.8	343.4	368.4	335.1
14	134.3	238.9	295.9	363.3	361.7	345.4	351.5
21	162.5	275.1	325.3	336.3	381.0	362.4	382.2
28	158.2	237.9	320.5	353.7	369.5	388.2	355.3
35	144.3	204.5	286.9	322.5	345.9	344.6	353.5
42	88.7	192.5	219.9	278.0	319.1	290.5	281.2

$$y_{ij} = b_0 + b_1 N_i + b_2 P_j + b_3 N_i P_j + b_4 N_i^2 + b_5 P_j^2 + \varepsilon_{ij}$$

表 二元二次多项式回归分析的方差分析(全模型)

变异来源	*DF*	*SS*	*MS*	*F*	
回　归	5	332061.25	66412.25	352.08**	$F_{0.05(5,43)}$=2.44; $F_{0.01(5,43)}$=3.49
b_1	1	219217.93	219217.93	1162.16**	$F_{0.05(1,43)}$=4.07; $F_{0.01(1,43)}$=7.27
b_2	1	754.29	754.29	4.00	
b_3	1	69.31	69.31	0.37	
b_4	1	61688.63	61688.63	327.04**	

结果表明b_2和b_3这两个偏回归系数不显著，

应该将模型缩减，

逐步去掉不显著的回归系数。

表　二元二次多项式回归的方差分析(缩减模型)

变异来源	*DF*	*SS*	*MS*	*F*	
回归平方和	4	331991.95	82997.99	446.42^{**}	$F_{0.05(5,44)}=2.58$; $F_{0.01(5,44)}=3.78$
b_1	1	219217.93	219217.93	1179.11^{**}	$F_{0.05(1,44)}=4.06$; $F_{0.01(1,44)}=7.24$
b_2	1	754.29	754.29	4.06^{*}	
b_4	1	61688.63	61688.63	331.81^{**}	
b_5	1	50331.10	50331.10	270.72^{**}	
误　差	44	8180.37	185.92		
总 变 异	48	340172.32			

表　二元二次多项式回归的回归系数及其显著性测验
（缩减模型）

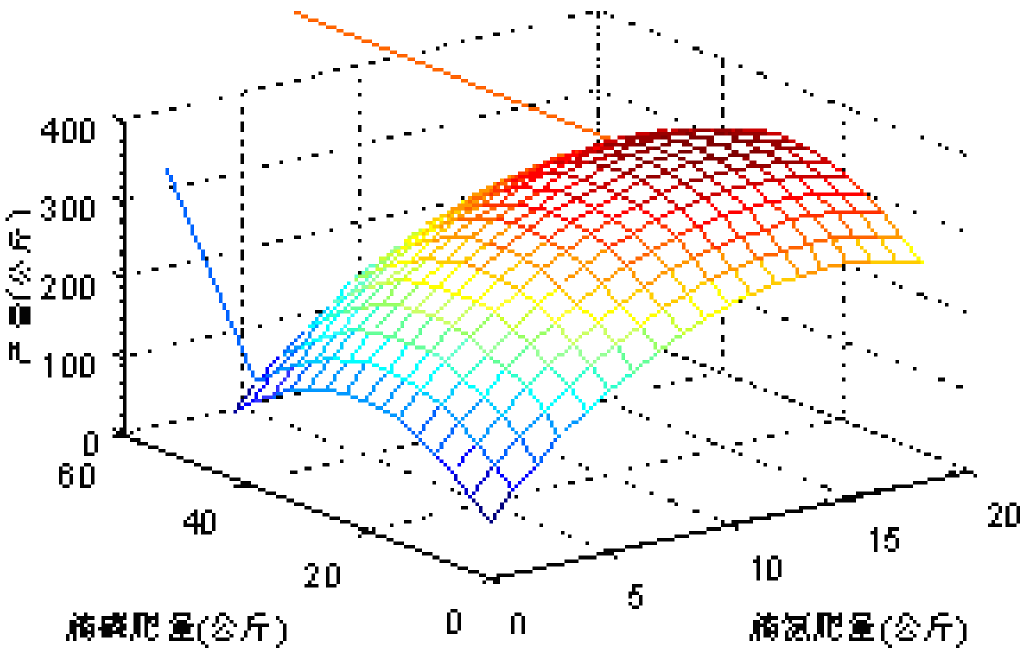
参数	回归系数估计值	标准误	t
b_0	76.70	6.06	12.66^{**}
b_1	31.63	1.17	27.02^{**}
b_2	8.21	0.50	16.37^{**}
b_4	-1.14	0.06	-18.22^{**}
b_5	-0.19	0.01	-16.45^{**}

该模型的回归变异占总变异的98%，因此可以较好地说明施用N、P对产量的影响。

- 产量对N、P施用量的回归方程为：

$$\hat{y}=76.70+31.63N+8.21P-1.14N^{2}-0.19P^{2}$$

- 由回归方程，可以作出产量对N、P施用量的响应曲面图。



$$\hat{y} = 76.70 + 31.63N + 8.21P - 1.14N^2 - 0.19P^2$$

- 分别对回归方程求对N和P的偏导数，并令偏导数等于0，可以求得极值：

$$\frac{\partial \hat{y}}{\partial N} = 31.63 - 2.28N = 0 \qquad N = 13.87(\mathrm{kg})$$

$$\frac{\partial \hat{y}}{\partial P} = 8.21 - 0.38P = 0 \qquad P = 21.61(\mathrm{kg})$$

由回归方程估计得尿素施用量为13.87kg，过磷酸钙施用量为21.61kg时产量最高。

响应面分析中通过回归方程进行预测时一般不能超过自变量的取值范围，例如氮肥的取值范围为0至18kg/亩，而磷肥的取值范围为0至42kg/亩。推论合理的处理组合时，也应该这样。