

旋转设计试验数据处理程序的编制

胡立成 姚 远

(黑龙江省农业科学院大豆研究所)

一、引 言

近年来广泛应用的旋转设计方法是以组合设计理论为基础的。它是把正交设计和回归分析结合起来进行的一种设计方法。其特点是试验次数少,得到的回归系数彼此独立,预测值方差仅与测点的离心度有关。因此这种方法被用于农业、畜牧业、化工、区划等各个领域,是进行多因素综合分析的有效方法,可以建立目标函数的多元二次方程,以方程为基础进行模拟预测。

在农业科学试验研究中,数学的应用愈来愈深入,计算量和精度就成为不可忽视的问题。计算机的应用和发展为解决这一课题提供了有效的工具,也是实现科技与生产现代化的一种趋势。本程序根据田间试验分析的实际要求而设计,也可用于其它领域,为其提供相应的最优决策信息。对于正交回归旋转组合设计的试验结果,可建立多元回归方程,进行多参数分析,确定回归方程的有效性,以回归方程为依据,定量地揭示出各因素(农艺措施、生产条件)的效应,各因素的交互作用,定量地描述其复杂的内在联系,为模拟各种生产条件,预测农作物产量,寻求最佳农艺措施,提供了可靠的依据。

二、设计思想与功能

本程序系统选用APPLE—II微型计算机,内存64K,一台143K磁盘驱动器,一

台屏幕监视器,一台SUPER5—CP80打印机。

全部程序由APPLESOFT(浮点BAS-IO)语言写成,总体采用模块结构,清晰易懂,便于查错,修改,扩充。模块按级别和功能分成若干程序文件存在一个磁盘上,用户可按需要调用。通过系统管理调用,实现程序复盖,解决了内存不足问题,程序运行中不断给出运行结果,使用户消除心理上的“着急感”,集中了注意力。系统采用“菜单”对话方式,机器用汉语拼音向用户发问,了解用户要求,满足其意愿。进行结果尽可能输出习惯的表达方式,自然、正齐、美观。

本系统可自动处理3~5个因素部分实施或全实施的“正交旋转设计”和“通用的旋转设计”的试验数据结果。首先求出回归方程,并作方差分析和回归系数显著性检验。当方程检验通过,得到用户认可时,系统即以回归方程为基础,继续完成以下功能:

1. 求最大、最小值,找出最优产量和最佳生产条件。
2. 打印频数表,在给定的产量值内统计出各因素水平值分布,高产概率,各农艺措施水平范围,为控制高产提供信息。
3. 进行主元素分析,确定各因素对产量贡献顺序,以及因素间的联系。
4. 进行双因素互作分析,即求出两个因素对产量的互作效应。

三、系统结构与实现

本系统由数据准备、统计分析、频数分析及优化、主元素分析、单因素分析和双因素交互分析等六个模块组成。结构见图1。

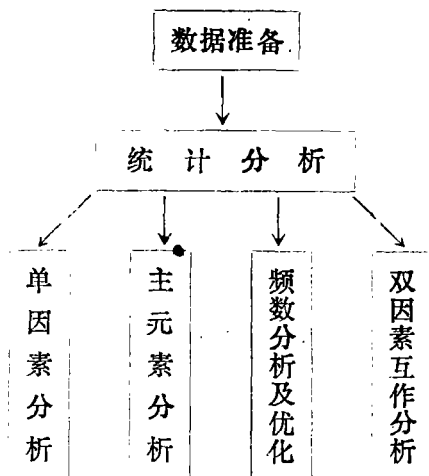


图1 系统结构图

各模块工作实现如下:

1. 数据准备模块

选定试验因素个数为 N , 小区个数 (即试验次数), G_9 , 用循环语句和 INPUT 语句完成输入 Y 矩阵 (即产量值) 的工作。

```

10 FOR I=1 TO G9
20 INPUT Y(I)
30 NEXT I
  
```

输完 G_9 个数据, 检查无错后, 模块自动按 N 值调用统计分析模块中不同因素个数的试验程序进行运算。

2. 统计分析模块

本模块根据因子数, 是否全实施, 小区个数自动按要求产生结构矩阵 $X(I, J)$:

```

10 INPUT N
20 PRINT "QUAN SHI Y/N"
30 INPUT M$: N1 = N
40 IF M$ = "N" THEN N1 = N - 1
50 R = 2^(N1/4)
60 PRINT "SHU RU XIAO QU SHU";
70 INPUT G9
  
```

```

80 DIM(50, G9), Y(G9)
90 FOR I=1 TO N1
100 G = 2*(N1 - I); P1 = 1
110 IF G = 0 THEN G = 1; C = G * 2
120 IF C = 0 THEN C = 2
130 FOR J=1 TO 2^(I-1):
    P3 = -1
140 FOR K=1 TO G: P3 = P3 + 1
150 X(I, P1 + P3) = 1; X(I, P1 + P3 + G) = -1
160 NEXT K: P1 = P1 + C
170 NEXT J
180 NEXT I
  
```

继续运算可得到交互项、平方项的中心化结构矩阵。再继续运算求出: $B_i = \sum X_{ij} Y_j$, $d_i = \sum X_i^2$, $b_i = B_i/d_i$, $Q = \sum B_i \cdot b_i$ 。根据计算出的系数 b_i 建立方程式:

$$Y = b_0 + \sum_{i=1}^m b_i X_i + \sum_{i < j} b_{ij} X_i X_j + \sum_{i=1}^m b_{ii} X_i^2$$

对上述方程进行检验求出平方和: $D(\text{总})$ 、 $D(\text{回})$ 、 $D(\text{剩})$ 、 $D(\text{误})$ 、 $D(L_f)$, 并进行 F_1 , F_2 显著性检验和系数 b 的 t 检验。

一般 $f(x)$ 为定义在 m 维欧氏空间内区域 R 中的 m 元二次函数, 向量 $x^T = (x_1, x_2, \dots, x_m)$ 就是函数的自变量。则函数 $f(x)$ 可在 R 域内任一点 x_0 附近展开为:

$$f(x) = f(x_0) + \nabla f(x_0)^T \Delta x + \frac{1}{2} \Delta x^T A \Delta x$$

$$\text{其中 } \nabla f(x_0) = \begin{pmatrix} \frac{\partial f(x_0)}{\partial x_1} \\ \frac{\partial f(x_0)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x_0)}{\partial x_m} \end{pmatrix}$$

是函数 $f(x)$ 在点 x_0 处的一阶偏导数向量, 也称梯度。而 A 为 $m \times m$ 对称矩阵

$$A = \begin{pmatrix} \frac{\partial^2 f(x_0)}{\partial x_1^2} & \frac{\partial^2 f(x_0)}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f(x_0)}{\partial x_1 \partial x_m} \\ \frac{\partial^2 f(x_0)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x_0)}{\partial x_2^2} & \dots & \frac{\partial^2 f(x_0)}{\partial x_2 \partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x_0)}{\partial x_m \partial x_1} & \frac{\partial^2 f(x_0)}{\partial x_m \partial x_2} & \dots & \frac{\partial^2 f(x_0)}{\partial x_m^2} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{pmatrix}$$

若在驻点 x^* 矩阵 A 正定, 则 x^* 为极小点, 矩阵 A 为负定, 则 x^* 为极大点; 矩阵 A 为不定, 则它为多维鞍形曲面, 无极值点。

3. 频数分析及优化模块

此模块为寻求最大产量出现频率, 以及各因素相应水平值分布。在给定量指标范围内, 利用循环语句在 $-r \sim r$ 约束区域, 按给定步长代入方程, 对方程求出的产量按不同指标归类, 再按相应水平值累加到相应的数组中, 具体通过以下程序实现:

```

10 FOR X1 = -r TO r STEP M
20 R1 = R1 + 1; R2 = 0
30 FOR X2 = -r TO r STEP M
40 R2 = R2 + 1; R3 = 0
  :
  :
140 FOR Xn = -r TO r STEP M
150 Rn = Rn + 1
160 Y = O(0) + O(1) * X1 + ...
    + O(N) * Xn
170 IF Y ≥ 500 THEN I = 1; GOTC
    200
180 IF Y ≥ 400 THEN I = 2; GOTO
    200
190 IF Y ≥ 300 THEN I = 3
200 IF X1 = -r + R2 * M THEN
    A(20 * I + R1) = A(20 * I + R1)
    + 1
210 IF X2 = -r + R1 * M THEN
    B(20 * I + R2) = B(20 * I + R2) + 1
  :
  :
240 IF Xn = -r + Rn * M THEN

```

$$F(20 * I + R_n) = F(20 * I + R_n) + 1$$

```

250 NEXT Xn
  :
290 NEXT X2
300 NEXT X1

```

运行后打印出频数表。

在频数分析中, 以找出的最高产量的相应水平值为基础, 对各因素水平值增加或减少 Δx 值, 计算其产量 y 值后与原产量 y_0 比较, 进行逐步逼近, 求出极大值为最优产量。

4. 主元素分析模块

此模块对方程进行线性变换, 消去交互项, 给出变换方程组及标准方程。

我们得到的回归方程是包含一次项的二次多项式, 不能直接套用“二次齐式标准化”的算法, 为解决此问题, 采取对一次项、二次项分别对待的方法。首先对“二次齐式”部分进行线性变换, 即正交标准化过程, 然后把变换式代入到一次项中, 形成新的一次项, 把新的一次项系数和特征值写到一个方程里, 即是标准方程, 具体是: 把方程

$$Y = b_0 + \sum_{i=1}^m b_i X_i + \sum_{i < j} b_{ij} X_i X_j + \sum_{i=1}^m b_i X_i^2, \text{ 中的 } \sum b_{ij} X_i X_j + \sum b_i X_i^2 \text{ 即二次项按齐次方程处}$$

理得:

$$|A| = \begin{vmatrix} b_{11} & b_{12}/2 & \dots & b_{1m}/2 \\ b_{21}/2 & b_{22} & \dots & b_{2m}/2 \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1}/2 & b_{m2}/2 & \dots & b_{mm} \end{vmatrix}$$

由 $|\lambda E - A| = 0$ 得:

$$\begin{vmatrix} \lambda - b_{11} & -b_{12}/2 & \dots & -b_{1m}/2 \\ -b_{21}/2 & \lambda - b_{22} & \dots & -b_{2m}/2 \\ \vdots & \vdots & \ddots & \vdots \\ -b_{m1}/2 & -b_{m2}/2 & \dots & \lambda - b_{mm} \end{vmatrix} = 0$$

求出特征值 λ , 即标准方程二次项系数, 对 n 个特征值 λ 求其特征向量, 得到特征向量方程组, 再经正交化和标准化, 即得:

$$Y_1 = B_{11}y_1 + B_{12}y_2 \cdots B_{1m}y_m$$

$$Y_2 = B_{21}y_1 + B_{22}y_2 \cdots B_{2m}y_m$$

$$\vdots$$

$$Y_n = B_{n1}y_1 + B_{n2}y_2 \cdots B_{nm}y_m$$

对系数 B 矩阵求逆, 代入原方程一次项中, 得标准方程一次项系数 C , 即得:

$$Y = b_0 + C_1y_1 + \cdots + C_my_m + \lambda_1Y_1^2 + \cdots + \lambda_mY_m^2$$

5. 单因素分析模块

此模块分析单因素对目标函数的作用。首先确定分析因素, 固定其它因素于一定水平上, 经计算得单因素函数方程式:

$$Y = P + AX_i + BX_i^2$$

先在屏幕定出横、纵坐标占屏幕比例, 划出横、纵轴, 在 X 横轴上确定 $-r \sim r$ 的分度, 再由 y 纵轴定出高、低产量范围。根据单因素方程, 用高分辨度绘图方式 (HGR), 于屏幕绘出曲线。

6. 双因素互作分析模块

此模块分析两个因素交互作用与产量的关系。先确定两个因素分析, 固定其它因素于一定水平上, 经计算得二元二次函数方程:

$$y_{ij} = p_i + AX_i + BX_j + CX_iX_j + DX_i^2 + EX_j^2$$

先确定等产量线图四个顶点 (x, y) 座标值, 用 $HPLLOT X_1, Y_1 TO X_2, Y_2 TO X_3, Y_3 TO X_1, Y_1$, 划出框图。

用循环语句从 $-r \sim r$ 之间先给出 X_1 值, 并代入双因素互作方程, 得一元二次方程。预先给定等产量线图的上下限值和间隔

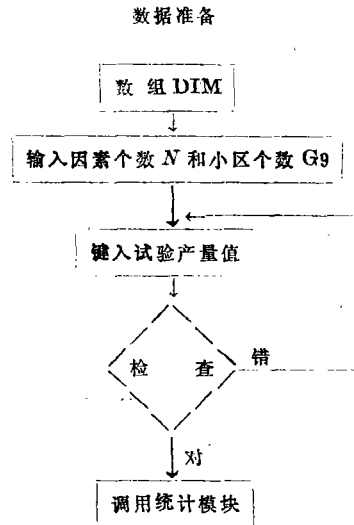
值, 求出 a, b, c , 代入 $X_2 = -b \pm \sqrt{\frac{b^2 - 4ac}{2a}}$ 公式, X_2 有三种情况: ①无根, 无点可划; ②有一个根, 只划一点; ③有二个根, 划二点。通过多次循环, 屏幕划出等产量线图。

四、结 语

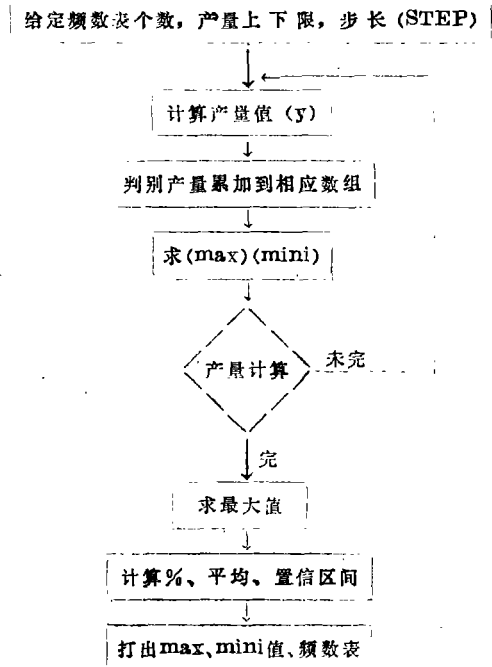
本文介绍了在 APPLE-II 微机上编制“旋转设计试验数据处理程序”的过程和系统实现内容。该程序从 1985 年开始在大豆、水稻、亚麻等作物上进行了应用, 处理了多组

试验数据, 指导了生产实际。该系统是当前开展综合试验研究进行统计分析的有实用价值的软件。

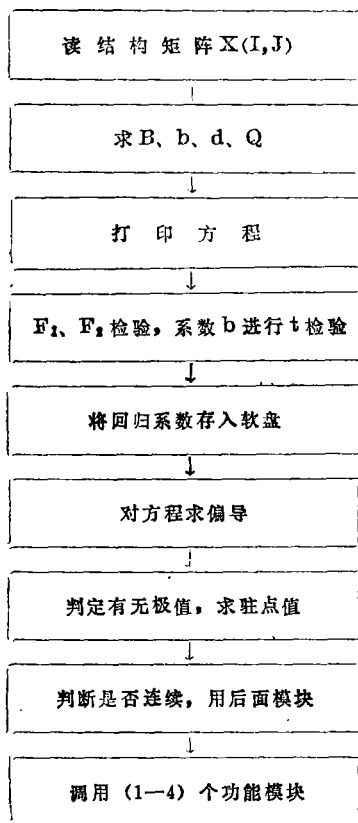
本系统各模块框图如下:



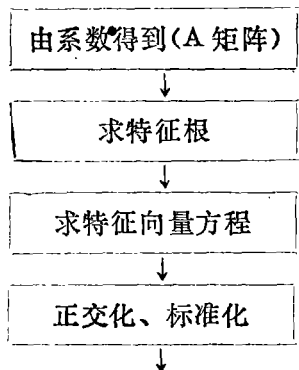
频数分析及优化



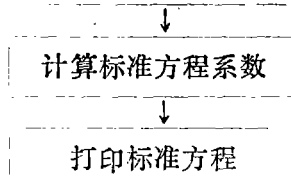
统计分析



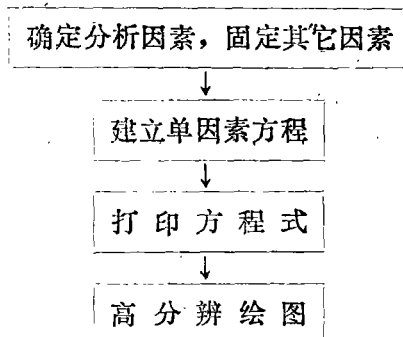
主元素分析



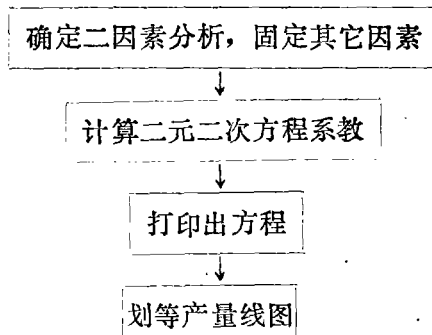
矩阵求逆



单因素分析



双因素互作分析



参考文献

1. 杨汝康、徐中儒: 生物数学讲义, 东北农学院, 1983
2. 刘福: 电子计算机在农业中的应用, 农村读物出版社 1985
3. 谭浩强等: BASIC 语言, 科学普及出版社, 1980