

用微电子计算机纠正测试数据的系统误差

苏建堂

(广西区域地质调查队)

当发现测试数据有系统误差时,不必对每个样品都重新分析,只抽取30%以上有代表性的样品重新分析即可,然后用微电子计算机校正原分析,这样可以收到快、好、省的效果。

数学模型的建立

回归分析是研究变量间的相关关系,找出相关变量之间的内部规律,以定量形式建立变量之间关系的数学表达式,从而可以根据某个变量的观测值预测另一个变量的估计值。把从有误差的数据随机地抽取一部分进行重新分析得到确认是准确的数据设为 Y ,其原分析数据设为 X ,那么 Y 和 X 应存在不确定的相关关系。这样就可利用回归分析法研究 Y 和 X 的规律性,确定 Y 和 X 关系的数学表达式,从原分析数据预测出估计值。本文只给出PC-1500机的计算程序。

程序使用说明

1. 功能

本程序可用于方差分析和一元线性回归和非线性回归(幂、指数、倒数指数、对数等4种常用曲线)分析。绘出散点图,计算印出回归系数 A , B 和预报自变的估计值。

2. 方法简介

变量 Y 和 X 的回归关系是何种函数事先是无法知道的,但可以过通散点图的形式来决定。线性回归系数是通过最小二乘法估算出来,当回归函数是非线性时,可通过数学变换为线性模型。其变换方法为:

(1) 幂函数(图1) $Y = dX^b$ 。令 $Y' = \lg Y$, $X' = \lg X$, $a = \lg d$, 则有 $Y' = a + bX'$ 。

(2) 指数函数(图2) $Y = de^{bx}$ 。令 $Y' = \ln Y$, $a = \ln d$, 则有 $Y' = a + bX$ 。

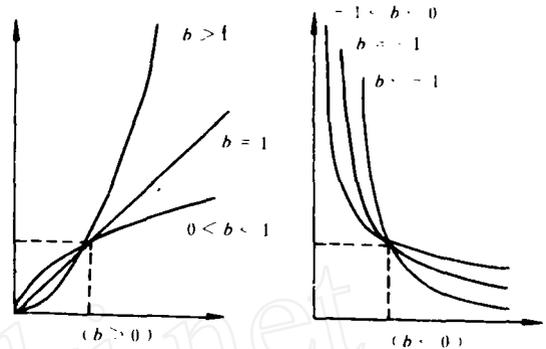


图 1

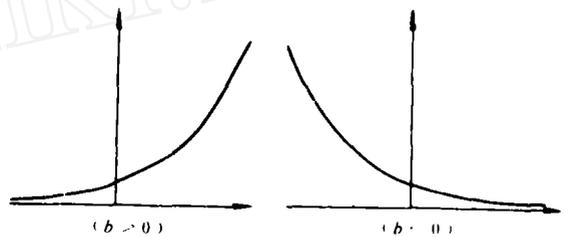


图 2

(3) 倒数指数函数(图3) $Y = de^{\frac{b}{x}}$ 。令 $Y' = \ln Y$, $X' = \frac{1}{X}$, $a = \ln d$, 则有 $Y' = a + bX'$ 。

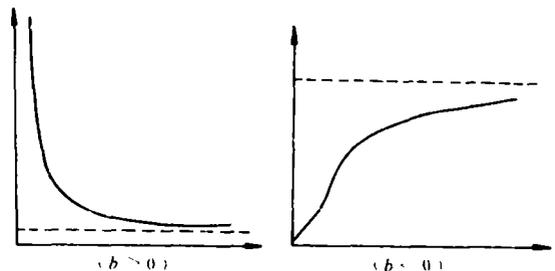


图 3

(4) 对数函数(图4) $Y = a + b \lg X$ 。

令 $X' = \lg X$, 则有 $Y = a + bX'$ 。

3. 计算公式

$$l_{xy} = \sum X^2 - \frac{1}{N} (\sum X)^2;$$

$$l_{xy} = \sum XY - \frac{1}{N} (\sum X)(\sum Y);$$

$$l_{yy} = \sum Y^2 - \frac{1}{N} (\sum Y)^2; \quad b = l_{xy}/l_{xx};$$

$$a = \bar{Y} - b\bar{X}$$

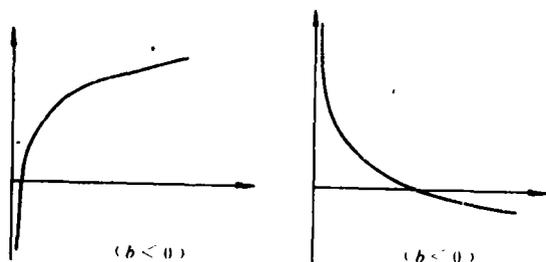


图 4

4. 标识符意义

N —观测数据组数; RR —原始数据输入形式控制参数; CC —选择线性或非线形模型控制参数; $Z(1, N)$ —原始数据矩阵; F —方差分析检验值; A, B —回归方程系数。

5. 操作步骤

数据输入顺序是自变量输完后再输因变量。

具体操作如下 (\boxed{E} 是 \boxed{ENTER}):

(1) 按 $\boxed{DEF} \boxed{A}$ 显 N ? 问观测值组数; 按 $N \boxed{E}$, (2) 显 RR ? 问数据输入方式,

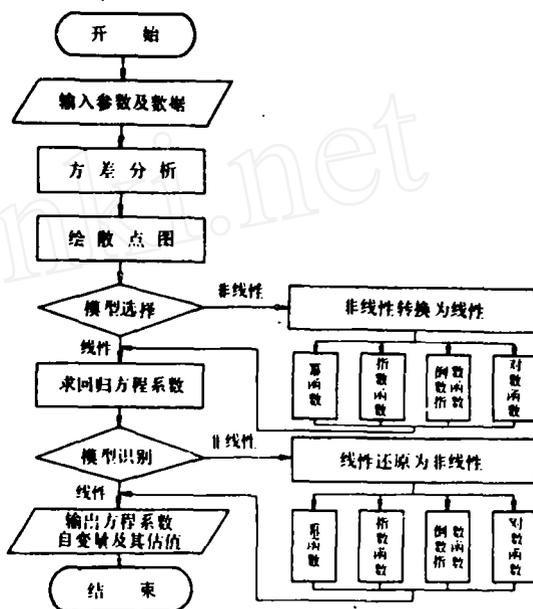
- $$RR = \begin{cases} \text{① 磁带} \\ \text{② DATA 语句;} \\ \text{③ 键盘} \end{cases}$$

按 $RR \boxed{E}$, (3) 打印出方差检验值 F 后显 $BREAK \text{ IN } 80$, 查 F 分布表, 当检验值 $F > F_{\alpha}$ 时, 则原分析数据有误差, (4) 按 $CONT \boxed{E}$ 绘散点图完后显 CC ? 选择回归模型,

- $$CC = \begin{cases} \text{① 幂函数} \\ \text{② 指数函数} \\ \text{③ 倒数指数函数 ; 按 } CC \boxed{E}, \\ \text{④ 对数函数} \\ \text{⑤ 线性函数} \end{cases}$$

打印回归系数 A, B 后, (5) 显 $X=?$ 问自变量观测值, 按 $X \boxed{E}$ 。打印出 X 和估计值 Y 后重复 (5) 步。当 X (原分析数据) 打完后, 按 \boxed{E} , 显 $>$ 结束。若另选择回归模型, 可按 $\boxed{DEF} \boxed{S}$, 各操作步序从显 CC ? 开始。

6. 程序框图



实例计算

本例是 $\times \times$ 地区水系沉积物 Zn 元素数据。光谱分析交报告后抽取 42 个样作化学分析。经过方差分析发现有系统误差, 光谱分析报出数据普遍高。因为化学分析较准确, 故命其为因变量 Y , 光谱分析命为自变量 X , 见表:

X	214	90	56	95	73	200	80	167	134	140	575	222	133	240	154	128	295	185	171	161	250
Y	86	50	36	51	50	90	51	71	72	60	122	101	80	78	76	58	88	91	72	82	116

X	113	107	180	185	194	161	233	283	216	300	300	220	125	88	82	200	220	275	300	124	200
Y	68	50	67	68	72	73	74	97	113	114	167	150	61	57	66	76	150	175	167	84	85

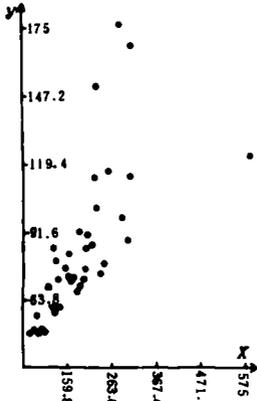
计算输出方差检验值 $F \approx 44.79 > F_{0.05}(80, 1) = 3.96$ 。散点图的形状比较符合于指数

函数曲线 (不够明显)。即给 $CC = 2$ 。算出回归系数 $A \approx 46.68, B = 0.0029$, 回归方程:

$$Y = 46.68 e^{0.0029X}$$

给出的几个预测值按“光谱半定量要求允许误差

计算输出:
 $r = 44.7900087$



A = 46.67923327
 B = 2.903923163E-03

Estimated value
 214 86.80574104
 140 70.032700373
 154 73.00124976
 161 74.00030370
 195 79.47771495
 216 82.40191459
 88 60.26911027

程序清单

```
10: "A".CLEAR :
    WAIT 0: INPUT "
    N?":N,"RR?":RR:
    N=N-1
20: DIM Z(1,N),Z1(
    1,N):ON RR
    GOSUB 570,580,
    590
30: X1=0:X2=0:X3=0
    :Y1=0:Y2=0
40: FOR I=0 TO N:X1
    =X1+Z(0,I):X2=
    X2+Z(1,I):X3=X
    3+Z(0,I)+Z(1,I
    ):Y1=Y1+Z(0,I)
    ^2
50: Y2=Y2+Z(1,I)^2
    :NEXT I
60: P=X3^2/(2*(N+1
```

本程序的其他用途

凡用一元回归模型来处理的问题都可以调用本程序。如区域化探中对新老数据进行等效处理；炼钢过程中炼毕碳对时间的预报等。本程序中的

范围是准确值的三倍”可以达到要求

计算输出和程序清单：

```
70
220: LINE (X2-5, Y2-
    5)-(X2, Y2)-(X2
    +5, Y2-5):
    LPRINT "Y"
230: ROTATE 0: FOR I
    =Y0 TO 300 STEP
    60: Y1=I: X1=X0:
    Y2=Y1: X2=5+X0
232: IF I=Y0 GOTO 23
    6
234: GOSUB 270: I1=I
    *B+D: LPRINT I1
236: NEXT I
240: ROTATE 0: FOR I
    =Y0 TO 0 STEP -6
    0: Y1=I: X1=X0: Y
    2=Y1: X2=5+X0
242: IF I=Y0 GOTO 24
    6
244: GOSUB 270: I1=I
    *B+D: LPRINT I1
246: NEXT I
248: GOTO 250
250: FOR E=0 TO N: X3
    =(Z(0,E)-C)/A:
    Y3=(Z(1,E)-D)/
    B: GOSUB 280:
    NEXT E
260: TEXT : LF 18:
    GOTO 290
270: LINE (X1, Y1)-(
    X2, Y2): RETURN
280: LINE (X3, Y3)-(
    X3+2, Y3+2), 0, 1
    , B: LPRINT E+1:
    RETURN
290: "S": COLOR 0:
    INPUT "CC?": CC
295: FOR J=0 TO 1:
    FOR I=0 TO N: Z1
    (J, I)=Z(J, I):
    NEXT I: NEXT J
300: ON CCGOSUB 510
    , 540, 530, 520
310: X=0: X1=0: Y=0: Y
    2=0
320: FOR I=0 TO N
330: X=X+Z1(0, I): X1
    =X1+Z1(1, I)
340: Y=Y+Z1(0, I)^2.
    Y2=Y2+Z1(0, I)*
    Z1(1, I)
350: NEXT I
360: L=Y-X^2/(N+1):
    L2=Y2-X*X1/(N+
    1)
370: B=L2/L: A=(X1/(
    N+1))-B*(X/(N+
    1))
380: ON CCGOSUB 550
    , 560, 560
390: LPRINT "A=": A
400: LPRINT "B=": B:
    LF 1: LPRINT "E
    stimated value
    "
410: INPUT "X=?": X:
    CLS: GOTO 430
420: GOTO 500
430: ON CCGOSUB 450
    , 460, 470, 480
440: Y=A+B*X: GOTO 4
    90
450: Y=A*X^0.8: GOTO 4
    90
460: Y=A*EXP (0*X):
    GOTO 490
470: Y=A*EXP (2*X/
    X): GOTO 490
480: Y=A+B*LOG (X)
490: LPRINT X: " ";
    Y: I=I+1: GOTO 4
    10
500: END
510: FOR J=0 TO N: Z1
    (1, J)=LOG (Z(1
    , J)): NEXT J
520: FOR J=0 TO N: Z1
    (0, J)=LOG (Z(0
    , J)): NEXT J:
    RETURN
530: FOR J=0 TO N: Z1
    (0, J)=1/Z(0, J)
    : NEXT J
540: FOR J=0 TO N: Z1
    (1, J)=LN (Z(1,
    J)): NEXT J:
    RETURN
550: A=10^A: RETURN
560: A=EXP A: RETURN
570: INPUT #Z(*).
    RETURN
580: FOR J=0 TO 1:
    FOR I=0 TO N:
    READ Z(J, I):
    NEXT I: NEXT J:
    RETURN
590: FOR J=0 TO 1:
    FOR I=0 TO N
600: A$=" "+STR$ (
    +1)+": "+STR$ (
    I+1)+": "=
610: PRINT A$:
620: INPUT Z(J, I):
    CLS: BEEP 1, 40
    : NEXT I: NEXT J
    : RETURN
```

绘散点图子程序经适当修改，可以绘制二维数据中的行（或列）两两之间的散点相关图，如因子载荷相关图。限于篇幅，本文不一一举例说明。