

Sharp PC-1500计算机应用实例

中国地质科学院岩矿测技术研究所

梁 汉 文

在分析实验室中为数不少的测定方法需要用检量线来求得欲测成份的含量,其过程大致是:在测定试样的同时对一组标准进行测定,然后以标准数据在直角坐标平面上绘制检量线,最后查出欲测定成份的含量,这种实验数据的作图处理法至今普遍被人们所采用。其优点是直观且不需要特殊的设备。缺点则是劳动强度较大,其准确度受坐标分度、描点以及绘曲线等诸因素的影响,容易引入主观误差。

根据回归分析的原理得知:具有函数关系的变量可以用最小二乘法确定的回归方程来表达它们之间的关系,且该方程所表达的数值与所有观测点数据误差的平方和是最小的^[2]。分析实验室中常用的分光光度和原子吸收等分光方法中吸收值A与浓度(ppm)之间的关系,在一定条件下,从一定的统计意义上来看,它们是存在着某种确定的函数关系的,这种函数关系可以用二次函数式 $y = ax^2 + bx + c$ 来表达^[1],这样只需要根据一组标准数据通过解方程组

$$a \sum x^4 + b \sum x^3 + c \sum x^2 = \sum x^2 y$$

$$a \sum x^3 + b \sum x^2 + c \sum x = \sum xy$$

$$a \sum x^2 + b \sum x + c \sum N = \sum y$$

求出其中的a、b、c三个系数,确定了函数式后就可以由吸收值A计算出浓度ppm(或微克)数,直接用计算的方法代替作图法进行结果计算,这就是通常所说的用最小二乘法拟合检量线。

Sharp PC-1500计算机体积小,对环境无特殊要求,使用灵便,是分析实验室中进行数据处理的较好工具。本文介绍的主要程序包括:

1. 用最小二乘法拟合检量线(可以计算、打印结果并绘制检量线,包括用离子选择电极法测定氟的计算)

2. 在直角坐标平面上绘制一根或数根曲线(绘折线或回归线可以任选)

3. 计算标准偏差

4. 分析结果检查(根据“地质实验工作技术

管理制度”中有关的允许误差范围可对部份分析项目结果进行检查。)

同时对其中的第一种程序进行了较深入的讨论;提出了用“浮动值” K_F 作为参比量计算误差,自动舍弃标准数据中的可疑值,使计算结果更为可靠。实践证明,方法操作简便有较大的实用意义。

结果和讨论

1. 函数式的选择

一般说来根据几对数据 $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ 能够确定一个 $(n-1)$ 次多项式

$$y = a_0 + a_1x + a_2x^2 + \dots + a_{n-1}x^{n-1}$$

来表达函数 $y = f(x)$,但这样的表达式往往并不一定就是最合适的,因为我们并不要求找到这样一条曲线,使得它通过每一点 (x_i, y_i) ($i = 1, 2, \dots, n$) 这是由于完全符合观测数据的函数式不可避免地包含了观测结果的随机测量误差,反而不能真实地表达变量 x 与 y 之间的函数关系。

我们通常遇到的检量线极少是一条理想的直线,而往往是一条近似直线的抛物线。为了表达这些检量线中的函数 $y = f(x)$,显然用一次函数式是满足不了要求的。而二次函数式的图象正是一条抛物线^[3],事实证明二次函数式 $y = ax^2 + bx + c$ 是能够较理想地跟踪一条抛物线,因而它可以作为检量线的函数表达式。

2. 可疑值的舍弃

资料[3]介绍了经简化后的Chauvenet数值舍弃标准后指出:若可疑值与平均值相差大于4 δ (四倍算术平均误差)时则弃去此观测值。能否用这种方法来解决用计算机处理检量线数据时可疑观测值的舍弃问题呢?下面结合例子加以讨论。首先应该指出的是,由于问题所涉及的是一条曲线上的动点,若只考虑数值 x 、 y (点的座标即吸收值和浓度)本身,是很不够的,而必需同时讨论每一点的斜率(y/x)才具有实际意义。以下是一条测 Fe_2O_3 用的检量线数据:

序号	y_i	x_i	$K_i(y_i/x_i)$	$d_i(\bar{K} - K_i)$
1	0	0		
2	93	50	$K_1 \approx 1.86$	0.01
3	183	100	$K_2 \approx 1.83$	0.04
4	279	150	$K_3 \approx 1.86$	0.01
5	420	200	$K_4 \approx 2.10$	0.23
6	459	250	$K_5 \approx 1.836$	0.03
7	533	300	$K_6 \approx 1.78$	0.09
8	649	350	$K_7 \approx 1.85$	0.02
9	747	400	$K_8 \approx 1.87$	0

$$\delta = \frac{\sum |d_i|}{n} = 0.05375$$

$$\bar{K} = \frac{\sum K_i}{n} = 1.87$$

$$y_i = A \times 1000$$

其中 $(K_4 - \bar{K}) > 4\delta$, 可以断定第五点是应该舍弃的。这说明对于线性关系比较好的检量线来说, 可以用 $\bar{K}$ 来作为参比量计算各点的斜率误差以决定取舍。

下面这组测定的数据:

序号	y_i	x_i	$K_i(y_i/x_i)$	$d_i(\bar{K} - K_i)$
1	0	0		
2	21	5	$K_1 = 4.2$	2.1
3	46	10	$K_2 = 4.6$	2.0
4	113	20	$K_3 = 5.65$	0.95
5	260	40	$K_4 = 6.5$	0.10
6	457	60	$K_5 = 7.62$	1.02
7	617	80	$K_6 = 7.71$	1.11
8	814	100	$K_7 = 8.14$	1.54
9	1008	120	$K_8 = 8.40$	1.80

$$\bar{K} = \frac{\sum K_i}{n} = 6.6$$

$$\delta = \frac{\sum d_i}{n} = 1.365$$

$$y_i = A \times 1000$$

用作图法处理时很容易发现第六点是应该舍弃的。若用 Chauvenet 数值舍弃标准计算则 $\bar{K} = 6.6$, $K_5 = 7.62$, $d_5 = 1.02$, $\delta = 1.365$, 因为 $1.02 < 1.365$ 所以即使将允许误差范围控制为 δ (Chauvenet 法提出 4δ) 也无法得出应该舍弃第六点的结论。这说明不能直接用 Chauvenet 法来解决其中的可疑值舍弃问题。其原因主要是: 这类检量线上各点斜率变化较大 ($4.2 \rightarrow 8.4$), 这就不能只用一个固定不变的平

均斜率 $\bar{K}$ 作为参比量来计算每一点的斜率误差, 否则计算将毫无意义。

为了解决上述问题, 本文提出了用一个“浮动值” K_F 作为参比量来计算各点的斜率误差, 最后决定取舍。 K_F 的数值是在 $\bar{K}$ 的基础上, 依据检量线的线性程度按一定的规律随着 y_i 值的变化而变化的。在上例中 K_F 的计算方法如下:

$$\bar{K} = \frac{\sum K_i}{8} = 6.6 \quad \bar{y} = \frac{\sum y_i}{8} = 417$$

$$K_E = (K_7 + K_8) / 2 (8.27)$$

$$y_E = (y_7 + y_8) / 2 (911)$$

$$K_S = (K_1 + K_2) / 2 (4.4)$$

$$y_S = (y_1 + y_2) / 2 (33.5)$$

(检量线少于 6 个点时 $K_E = K_n$, $K_S = K_1$,

$y_E = y_n$, $y_S = y_1$)

$\Delta K_1 = (\bar{K} - K_S) / (\bar{y} - y_S)$ ($y_i < \bar{y}$ 时斜率的增值)

$\Delta K_2 = (K_E - \bar{K}) / (y_E - \bar{y})$ ($y_i > \bar{y}$ 时斜率的增值)

$$K_F = \bar{K} + \Delta K_1 \times (y_i - \bar{y}) \quad (y_i < \bar{y} \text{ 时})$$

$$K_F = \bar{K} + \Delta K_2 \times (y_i - \bar{y}) \quad (y_i > \bar{y} \text{ 时})$$

不难看出 K_F 是在 $\bar{K}$ 的基础上加了一个斜率的增值 ΔK (ΔK 本身决定于检量线的线性程度) 与 $(y_i - \bar{y})$ 的乘积。 K_F 是由 $\bar{K}$ 、 ΔK 和 y_i 所决定的, 它是一个“浮动值”。在程序中设定了若干检量线多于 6

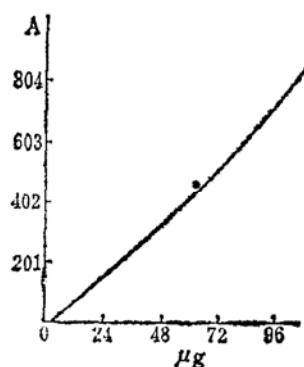


图 1 硼的校正曲线
(舍弃第六点)

Fig 1 Calibration Curve of B
(reject sixth point)

个点时把它分为四段来处理。对于曲线上的每一个点 (x_i, y_i) (零除外) 都可以求出相应的 K_i 和 K_F 。实践证明: 对于大多数检量线来说 K_i 与 K_F 之间的允许误差控制在 2.5δ 内是比较合适的。考虑到某些线性较好的检量线虽然个别点 $(K_i - K_F) > 2.5\delta$, 但 K_i

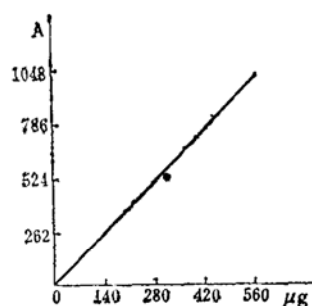


图2 铁的校正曲线
(舍弃第七点)

Fig 2 Calibration curve of Fe_2O_3
(reject seventh point)

与 K_F 之间的相对误差仍然不大, 为使这样的点免遭舍弃, 程序中设定了只有 $(K_I - K_F) > 2.5\delta$ 同时 $K_I > 103 \times K_F$ 或 $K_I < 0.97 \times K_F$ (相对误差3%) 的点才被舍弃。被舍弃的数据计算机将用红色字体打印出来, 此时该数据将不参加回归。见图1、图2。图1为测定硼的检量线, 其中的第六点被计算机自动舍弃。图2为测定铁的检量线其中的第七点被计算机舍弃。经过这样处理所得的结果更接近于用作图法求得平均值。(见“结果比较”)

3. 结果比较

下表所列是一组用磺基水杨酸比色法测定铁的数据。试样的 $\bar{r}$ 及 r^* 分别是五位工程师分析者用作图法处理所得的结果及它们的平均值。

结果比较

Comparison of results (Fe_2O_3)

标准系列		试							样		
A	r	A	r					$\bar{r}$	r PC-1500	r* PC-1500	
0	0										
0.093	50	0.342	184	184	185	184	181.5	184.3	186.20	184.76	
0.183	100	0.204	110	110	111	111	111.5	110.7	111.48	110.59	
0.279	150	0.109	59	60	59	61	60	59.8	59.76	59.38	
0.370	200	0.870	467	467	466.5	466	465.5	466.4	467.7	466.24	
0.459	250	0.851	456.5	457	456	456	455	451.1	457.7	456.17	
0.533	300	1.222	652	655	654	654	653	653.6	651.41	651.87	
0.649	350	1.228	655	657	656.5	657	656	656.3	654.51	655.01	
0.747	400	1.184	631	635	633	633	632.5	632.9	631.73	631.90	
0.839	450	0.792	425	425	425	425	424	424.8	426.57	424.88	
0.935	500	0.763	410	410	410	409	408.5	409.5	411.23	409.49	
1.030	550	0.836	449	447	448.5	448.5	447.5	448.1	449.79	448.20	
1.130	600	1.140	607	609	609	610	614	609.8	608.9	608.76	
1.216	650	0.160	86.5	86	86.5	88	87.5	86.9	87.56	86.89	
1.310	700	0.081	43.5	44	43	46	45.5	44.4	44.47	44.27	

打印格式及绘图举例

Data,

15.79
16.14
15.98
15.91
16.34
16.27
16.12
16.12
16.39
16.78
15.85
16.2

Mean = 16.1575

d = 0.19875

S.D = 0.272167

R.S.D = 1.68446232

N = 12

A. 计算标准偏差

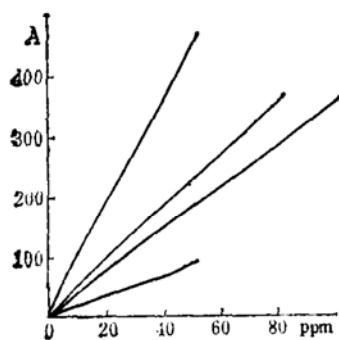
A. Standard deviation

数据表明由计算机处理所得的结果与 $\bar{r}$ 相当吻合, 尤以 r^* (计算机自动舍弃标准系列的第七点数据后计算所得的结果) 与 $\bar{r}$ 更为接近。(作图法数据由地矿部岩矿测试技术研究所秦大章、邵济新、张宗纯、杨桂芳等同志提供)

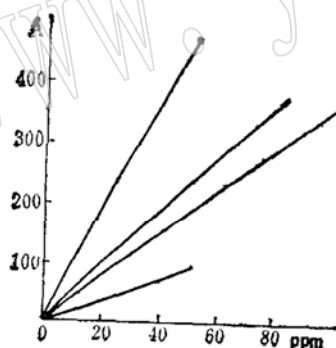
程序及操作(略)

参 考 资 料

1. 沈恒范,《概率论讲义》,人民教育出版社,1966。
2. 中国科学院数学研究所数理统计组,《回归分析方法》,科学出版社,1975。
3. 冯师颜,《误差理论与实验数据处理》,科学出版社,1963。
4. W.希尔 G.洛夫,《应用数学基础》上册,周焕山译,科学出版社,1974。



B. 绘图(回归线)
B. Regressive curves



C. 折线
C. Broken lines