

# 作图法与线性回归法的分析与运用

冯元新

(浙江科技学院 理学院,浙江 杭州 310023)

**摘 要:** 主要比较和分析了作图法与线性回归法这两种实验中常用的数据处理方法,并阐述了如何利用这两种数据处理方法找出并删除错误或误差较大的实验数据,从而得到精确的实验结果。

**关键词:** 作图法;线性回归法;误差

**中图分类号:** O4-33

**文献标识码:** A

**文章编号:** 1671-8798(2004)01-0004-03

作图法与线性回归法在大学物理实验中是两种常用的数据处理方法。若自变量  $X$  与因变量  $Y$  呈线性关系即  $Y = aX + b$ ,则在实验过程中可测得一系列的测量点,其坐标为  $(X_i, Y_i)$ 。为求系数  $a$  与  $b$ ,用作图法处理数据来绘制直线,要求所绘的直线能使大多数的测量点都落在线上,或测量点能够尽量对称地散布于直线的两侧。也可用数学的方法,求出满足各测量点偏离直线的平方和为最小的一条直线,该直线就称为各测量点的回归直线<sup>[1]</sup>。由线性回归法可求得

$$a = \frac{\bar{x} \cdot \bar{y} - \overline{xy}}{\bar{x}^2 - \overline{x^2}} \quad (1)$$

$$b = \bar{y} - a\bar{x} \quad (2)$$

为了判断所得的结果是否合理,在待定常数确定以后,还需要计算一下相关系数  $r$ ,对于一元线性回归,  $r$  的定义为<sup>[2]</sup>:

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sqrt{(\bar{x}^2 - \overline{x^2})(\bar{y}^2 - \overline{y^2})}} \quad (3)$$

$r$  值越接近 1,说明实验数据点密集的分布在所求得的直线的近旁,用线性函数进行回归是合适的;相反,若  $r$  值远小于 1 而接近于零,说明实验数据对求得的直线很分散。

回归系数  $a$  的标准误差为:

$$S_a = \sqrt{\frac{(1-r^2)(\bar{y}^2 - \overline{y^2})}{(n-2)(\bar{x}^2 - \overline{x^2})}} \quad (4)$$

其相对误差为:

$$E_a = \frac{S_a}{a} \times 100\% = \sqrt{\frac{(1-r^2)(\bar{x}^2 - \overline{x^2})(\bar{y}^2 - \overline{y^2})}{(n-2)(\overline{xy} - \bar{x} \cdot \bar{y})^2}} \times 100\% = \sqrt{\frac{(1-r^2)}{(n-2)r^2}} \times 100\% \quad (5)$$

## 1 提出问题

由定义可以看出,作图法精度要求低,是半定量的求解方法,而线性回归法是定量的,并且有严格的数学依据,因此,在求解时会认为用线性回归法好。但问题是若实验数据中有个别数据是错误的或误差较大的,则

收稿日期: 2003-10-08

作者简介: (1976—),男,浙江兰溪人,工程师,主要从事物理实验教学。

即使用线性回归法按照严格的数学依据对已有的数据进行定量的数学分析处理,也会使实验结果错误或者误差较大。这时,如果采用作图法从图上来观察这些实验数据的分布情况,找出可疑点,并通过公式(5)来分析这些点的相对误差,则得到的实验结果会更精确些。

## 2 举 例

以刚体转动惯量的测量实验为例,分析如何处理实验数据。实验简图如图1所示。

调节轴线 $OO'$ 垂直,绕线保持水平(与 $OO'$ 垂直)时,使质量为 $m$ 的砝码从静止下落,而刚体绕轴 $OO'$ 转动,此时可推导出刚体转动惯量实验的测量式:

$$mgr = \frac{2hI}{rt^2} + M_\mu \quad (6)$$

其中 $mgr$ 为外力矩, $h$ 为下落高度, $I$ 为所要求的转动惯量, $t$ 为下落时间, $r$ 为绕线半径, $M_\mu$ 为总摩擦力矩。

实验中,一般保持 $r$ 与 $h$ 为定值, $I$ 与 $M_\mu$ 也为定量,因此,可以将(6)式转化为:

$$m = \frac{2hI}{gr^2} \cdot \frac{1}{t^2} + \frac{M_\mu}{gr} \quad (7)$$

由于 $\frac{2hI}{gr^2}$ 与 $\frac{M_\mu}{gr}$ 为不变量,因此,可以设

$$k = \frac{2hI}{gr^2} \quad (8)$$

$$c = \frac{M_\mu}{gr} \quad (9)$$

得

$$m = k \frac{1}{t^2} + c \quad (10)$$

从(7)式可以看出,质量 $m$ 与时间平方的倒数 $1/t^2$ 成线性关系,若用作图法与线性回归法得到斜率 $k$ 后,则由(8)式可得转动惯量 $I$ ,因此,如何求得斜率 $k$ 是处理数据的关键。

表1是实验中测得的一组数据。

## 3 分析说明

作图法中作直线的准则是:要使所作的直线尽可能地通过所测的点,或使测量点尽可能地均匀分布在所作直线的两侧。根据这条规则,可作出直线,如图2所示。当然还可以作出许多看似符合这条规则的直线,如图3所示等。因此,用作图法求斜率 $k$ 时,偶然误差很大。

表1 刚体转动惯量实验中测得的一组数据

SN	所挂砝码 质量 $m/g$	砝码下落高度 $h$ 所用时间 $t/s$
1	15	9"17
2	20	7"45
3	25	6"59
4	30	5"87
5	35	5"48
6	40	5"02
7	45	4"36

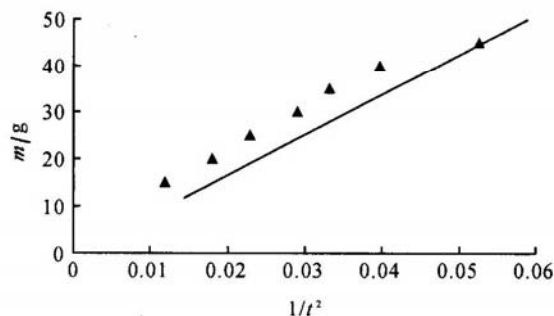


图2 由表1数据用作图法得到的直线1

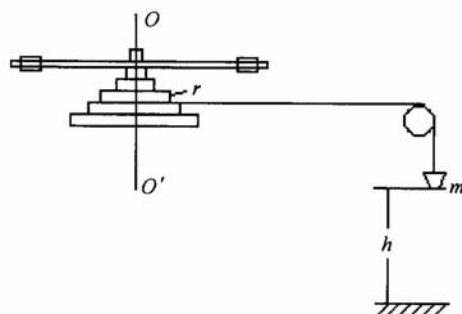


图1 刚体转动惯量实验简图

仔细分析图2,可以发现点(45,4"36)与所作直线偏差较大,用公式(3)和公式(5)来计算表1与除去可疑数据(45,4"36)后的相对系数 $r_1$ 、 $r_2$ 和斜率 $k$ 的相对误差 $E_{k1}$ 、 $E_{k2}$ ,公式(3)中的 $y$ 相当于 $m$ , $x$ 相当于 $1/t^2$ , $k$ 对应于公式(1)中的 $a$ ,得: $r_1 = 0.98544$ , $r_2 = 0.99906$ , $E_{k1} = 7.72\%$ , $E_{k2} = 2.17\%$ 。

为进一步分析,逐个删除实验数据点后,用公式(3)和公式(5)来计算相关系数 $r$ 和 $k$ 的相对误差 $E$ ,见表2。

表2 逐个删除各个实验数据点后所得的 $r$ 和 $E$

删除的数据点( $m, t$ )	相关系数 $r$	相对误差 $E/\%$
(15, 9"17)	0.979 85	10.19
(20, 7"45)	0.983 90	9.08
(25, 6"59)	0.984 85	8.80
(30, 5"87)	0.985 67	8.56
(35, 5"48)	0.989 06	7.46
(40, 5"02)	0.988 21	7.75

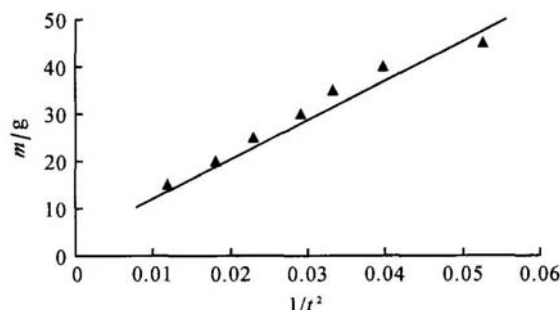


图3 由表1数据用作图法得到的直线2

#### 4 结 论

由结果可以看出,忽略数据(45,4"36)之后,得到 $k$ 的相对误差为2.17%,比原先未忽略此点前得到 $k$ 的相对误差为7.72%精确了三倍多。同时,从表2可以看出,若未忽略此点,而忽略其他实验数据点,得到 $k$ 的相对误差仍然很大。

从以上分析可以看出,作图法处理数据的优点是直观、简便、快捷,但偶然误差大,即对同一组数据多次作图处理时,可得到误差不同的实验结果。线性回归法处理数据的优点是偶然误差小,得到的结果唯一确定,但不能简单明了地发现实验数据中的错误数据,数据的计算量大。

因此,在处理数据时,为了得到更精确的实验结果,可先用作图法来观察测量点的分布,根据观察,剔除测量中的错误或误差较大的数据,然后用线性回归法来求得精确解。这样处理数据的秩序与方法,对于精度要求较高的实验来说,尤其重要。

#### 参考文献:

- [1] 吴乃爵. 大学物理实验教程[M]. 杭州: 杭州大学出版社, 1996. 31—35.
- [2] 陆廷济, 费定曜. 大学物理实验[M]. 上海: 同济大学出版社, 1996. 24—27.

## Analysis and application of mapping and linear regression

FENG Yuan-xin

(Dept. of Science, Zhejiang University of Science and Technology, Hangzhou 310023, China)

**Abstract:** This paper compares and analyses mapping and linear regression, the two common treatment of experimental data, and studies how to examine and delete the error data by this two treatments, so that the exact result of experiment can be obtained.

**Key words:** mapping; linear regression; error