

第 1 章 误差理论与数据处理

本章介绍测量的概念，误差分析与数据处理的初步知识，给出一些结论和简化的计算方法，希望同学们结合每一个具体实验，通过运用加以掌握。

1.1 测量

1.1.1 测量

物理实验中为了找出有关物理量之间的定量关系，必须进行定量的测量。测量是物理实验中及其重要的一个组成部分。**测量就是把待测量直接或间接地与另外一个选作计量标准的同类物理量进行比较的过程。**

测量得到的实验数据应包含测量值的**大小和单位**，二者缺一不可。

1.1.2 测量的分类

在实验中会遇到各种类型的测量，可以从不同的角度对测量进行分类，按测量方法可分为直接测量和间接测量；按测量的条件可分为等精度测量和非等精度测量。

1. 直接测量和间接测量

(1) **直接测量**。用测量仪器或仪表与待测量进行比较，直接测出被测量结果的测量。例如用米尺测量物体长度，用天平测量物体的质量等都是直接测量。

(2) **间接测量**。利用几个直接测量的量按照一定的函数关系得到待测量的大小。例如通过测量体积和质量得到物体的密度，通过测量单摆的摆长和周期测定重力加速度等。

2. 等精度测量和非等精度测量

(1) **等精度测量**。是指在相同条件下对同一个物理量进行的多次测量。所得到的一组数据 $x_1, x_2, x_3 \cdots x_n$ 称为**测量列**。严格的等精度测量是不存在的，当某些条件的变化对测量结果影响不大或可以忽略时，可视为等精度测量。在物理实验中要求多次测量的均指等精度测量，对测量误差与数据处理的讨论，都是以等精度测量为前提的。

(2) **非等精度测量**。是指在测量过程中由于仪器的不同、方法的差异、测量条件的改变以及测量者的原因而造成测量结果的变化，这样的测量称为非等精度测量。非等精度测量通常用在科学研究实验中。

1.2 误差

1.2.1 误差的概念及表示方法

1. 真值

任何一个物理量在一定客观条件下（某一时刻、某一位置或某一状态），都存在着一个

不以人的意志为转移的客观值，这个客观值称为该物理量的**真值**。

被测量的真值是客观存在的，是一个理想的概念，一般是不可知的。在实际测量中常用被测量的实际值或已修正过的算术平均值来代替真值，称为约定真值。

2. 误差

误差就是测量结果与被测量的真值之间的差异。

物理实验是以测量为基础的，由于实验原理、测量装置、实验条件、观测者等种种因素的局限，任何测量结果总存在着误差。进行误差分析对科学实验有极其重要的指导意义：一是通过分析误差来源及其性质，采用合理的方法减少或消除误差，并对实验结果作出合理的评价；二是通过误差分析优化实验方法、选择测量仪器和测量条件、拟定实验步骤和数据处理方法等，获得合理的实验结果。

3. 误差的表示方法

误差的表示方法一般有两种，即绝对误差和相对误差。

(1) **绝对误差**。测量值与被测量的真值之间的差值。绝对误差用 Δx 表示。

$$\Delta x = x - x_0 \quad (1) \text{ 式}$$

中， x 为测量结果， x_0 为被测量的真值。误差的大小反映了测量结果的准确程度。

(2) **相对误差**。绝对误差与被测量的真值的比值。相对误差用 E 表示。

$$E = \frac{\Delta x}{x_0} \times 100\% \quad (2)$$

相对误差的大小反映了测量结果的优劣。

1.2.2 误差的分类

误差根据其来源和性质可分为系统误差和随机误差两大类。

1. 系统误差

系统误差是指在同一条件下（方法、仪器、环境、人员），多次测量同一被测量的过程中误差的大小和符号保持不变，或当条件改变时按某一规律变化的误差分量。

系统误差主要来源有以下几方面：

(1) **方法误差**。由于实验原理或方法的近似性带来的误差，如用伏安法测电阻没有考虑电表内阻的影响，用单摆测重力加速度时取 $\sin \theta \approx \theta$ 带来的误差等。

(2) **仪器误差**。由于仪器本身不完善而产生的误差，包括仪器的零值误差、示值误差、机构误差和测量附件误差等，如天平不等臂带来的误差。

(3) **环境误差**。由于实际环境条件与规定条件不一致引起的误差，如标准电池是以 20°C 时的电动势作为标称值的，若在 30°C 条件下使用时，如不加以修正就引入了系统误差。

(4) **人为误差**。由于测量人员主观因素和操作技术所引入的误差。

系统误差又可以分为已定系统误差和未定系统误差。已定系统误差的符号和绝对值可以确定。未定系统误差的符号和绝对值不能确定，实验中常用估计误差限的方法得出。

大学物理实验要重视对系统误差的分析，尽量减小它对测量结果的影响，一般采用的

方法是：① 对已定系统误差进行修正；②通过校准测量仪器、改进实验方案和实验装置、修正测量数据和采用适当的测量方法(如交换法、补偿法、替换法、异号法等)予以减小或消除；③合理评定系统误差分量大致对应的 B 类不确定度。

2. 随机误差

在多次测量同一被测量的过程中，绝对值和符号以不可预知的方式变化着的误差分量称为随机误差。在采取措施消除或修正一切明显的系统误差之后，对被测量进行多次测量时，测量值仍会出现一些无规律的起伏。随机误差是由实验中各种因素(如温度、湿度、气流、电源电压、杂散电磁场、震动等)的微小变动引起的，以及实验装置、测量机构在各次调整操作时的变动性，测量仪器示值的变动性，观察者本人在判断和估计读数上的变动性等等。随机误差，就某一测量而言是没有规律的，当测量次数足够多时，随机误差服从统计分布规律，可以用统计学方法估算随机误差。

3. 异常数据的剔除

剔除测量列中异常数据的标准有 3σ 准则、肖维准则、格拉布斯准则等。

统计理论表明，测量值的偏差超过 3σ 的概率已小于 1%。因此，可以认为偏差超过 3σ 的测量值是由于其它因素(实验装置故障、测量条件的意外变化、较强的外界干扰)或过失造成的异常数据，应当剔除。方法是用偏差 $\Delta x'_i = (x_i - \bar{x})$ 和 3σ 比较，若 $\Delta x'_i \geq 3\sigma$ ，则该测量值应该剔除掉。

1.2.3 仪器量程 精密度 准确度

量程是指仪器所能测量的范围。对量程的选择要适当，当被测量超过仪器的量程时会损坏仪器，但也不应一味选择大量程，因为如果仪器的量程比测量值大很多时，测量误差往往会比较大。

精密度是指仪器所能分辨物理量的最小值，一般与仪器的最小分度值一致，最小分度值越小，所测物理量的位数就越多，精密度越高。同时仪器精密度的大小反映了各次测量结果的离散程度。

准确度是表示测量结果与真值接近的程度，因而它是系统误差的反映。由于测量目的不同，对仪器准确程度的要求也不同。按国家规定，电气仪表的准确度等级 a 分为 0.1、0.2、0.5、1.0、1.5、2.5、5.0 共七级，在规定条件下使用时，其示值 x 的最大绝对误差为

$$\Delta x = \pm \text{量程} \times \text{准确度等级} \% \quad (3)$$

例如，0.5 级电压表量程为 3 V 时， $\Delta U = \pm 3 \times 0.5\% = \pm 0.015 \text{ V}$ 。

对仪器准确度的选择要适当，在满足测量要求的前提下尽量选择准确度等级较低的仪器。当待测物理量为间接测量时，各直接测量仪器准确度等级的选择，应根据误差合成和误差均分原理，视直接测量的误差对实验最终结果影响程度的大小而定，影响小的可选择准确度等级较低的仪器，否则应选择准确度等级较高的仪器。

1.3 测量结果的最佳值与随机误差的估算

随机误差与系统误差的来源和性质不同，所以处理的方法也不同。

1.3.1 随机误差的分布规律

实践证明，等精度测量中，当测量次数 n 很大时，测量列的随机误差多服从正态分布。正态分布的曲线如图 1 所示，图中横坐标表示随机误差 $\Delta x = (x_i - x_0)$ ，纵坐标为对应的误差出现的概率密度函数 $f(\Delta x)$ 。应用概率论方法可导出

$$f(\Delta x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\Delta x^2}{2\sigma^2}} \quad (4)$$

式中，特征量 σ 称为标准误差。

$$\sigma = \sqrt{\frac{\sum \Delta x_i^2}{n}} \quad (n \rightarrow \infty) \quad (5)$$

图 2 是不同 σ 值时的 $f(\Delta x)$ 曲线。 σ 小，曲线陡且峰值高，说明误差集中，小误差占优势，各测量值的离散性小，重复性好。反之， σ 大，曲线较平坦，各测量值的离散性大，重复性差。

随机误差落在 $[\Delta x, \Delta x + d(\Delta x)]$ 区间内的概率为 $f(\Delta x)d(\Delta x)$ ，显然误差出现在 $(-\infty, +\infty)$ 范围

内的概率为百分之百， $\int_{-\infty}^{+\infty} f(\Delta x)d(\Delta x) = 1$ 。

误差出现在 $(-\sigma, +\sigma)$ 内的概率 P 就是图 1.3.1 中该区间内 $f(\Delta x)$ 曲线下的面积，可以证

$$P = \int_{-\sigma}^{\sigma} f(x)dx = 68.3\%。$$

这说明任一次测量，随机误差落在 $(-\sigma, +\sigma)$ 区间的概率为 68.3%。区间 $(-\sigma, +\sigma)$ 称为置信区间，相应的概率称为置信概率。置信区间分别取 $(-2\sigma, +2\sigma)$ 、 $(-3\sigma, +3\sigma)$ 时，相应的置信概率为 $P(2\sigma) = 95.4\%$ 、 $P(3\sigma) = 99.7\%$ 。

服从正态分布的随机误差具有以下特征：

- ①单峰性。绝对值小的误差出现的概率比绝对值大的误差出现的概率大。
- ②对称性。绝对值相等的正、负误差出现的概率相等。
- ③有界性。绝对值很大的误差出现的概率很小，甚至趋近于零。
- ④抵偿性。随机误差的算术平均值随着测量次数的增加而越来越趋于零，即

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \Delta x_i = 0 \quad (6)$$

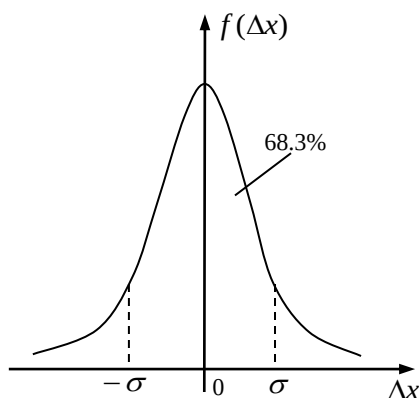


图 1 随机误差的正态分布

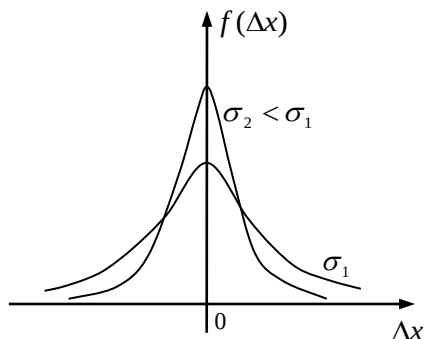


图 2 不同 σ 的概率密度曲线

1.3.2 测量结果最佳值——算术平均值

在测量不可避免地存在随机误差的情况下，每次测量都有差异，那么接近真值的最佳值是什么呢？

我们可以利用随机误差的统计特性来判断实验结果的最佳值。

设对某一物理量进行了 n 次等精度测量，所得测量列为： $x_1, x_2, x_3 \cdots x_n$ 。测量结果的算术平均值为

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

根据误差的定义有

$$\Delta x_i = x_i - x_0$$

$$\frac{1}{n} \sum_{i=1}^n \Delta x_i = \frac{1}{n} \sum_{i=1}^n x_i - x_0 = \bar{x} - x_0$$

随着测量次数的增加，测量列的算术平均值越来越趋近于真值。

随机误差的抵偿性，当 $n \rightarrow \infty$ 时 $\frac{1}{n} \sum \Delta x_i \rightarrow 0$ ，因此 $\bar{x} \rightarrow x_0$ 。

所以，测量列的算术平均值 $\bar{x}$ 是真值 x_0 的最佳估计值。

1.3.3 随机误差的估算——标准偏差

算术平均值作为真值的最佳估计值，在实际测量中，测量结果的随机误差究竟有多大？如何来估算呢？

各次测量值与算术平均值之差 $\Delta x'_i = (x_i - \bar{x})$ 称为**偏差（残差）**。

当测量次数有限时，随机误差引起测量值的离散性可用单次测量的标准偏差来表示，用 S_x 表示，它是 σ 的一个估算值，在有限次测量中可用由以下贝塞尔公式计算：

$$S_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} \quad (7)$$

S_x 的统计意义： S_x 小，说明随机误差的分布范围窄，小误差占优势，各测量值的离散性小，重复性好。反之， S_x 大，各测量值的离散性大，重复性差。

一般情况下，在多次测量后，是以算**术**平均值表达测量结果的，而算**术**平均值本身也是随机量，也有一定的分散性，可用平均值的标准偏差 $S_{\bar{x}}$ 来表征这一分散性：

$$S_x = \frac{S_x}{\sqrt{n}} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n(n-1)}} \quad (8)$$

可以看出，算术平均值的标准偏差要比单次测量的标准偏差小的多，这是因为算术平均值已经对单次测量的随机误差有一定的抵销，因而平均值会更接近真值。

1.4 测量结果的不确定度评定

1.4.1 不确定度的概念

由于测量误差的存在，测量结果只能得到一个真值的最佳估计值和用于表示该估计值近似程度的误差范围，这个用于定量评定测量结果质量的物理量就是不确定度。

不确定度 (Uncertainty) 是指由于测量误差的存在而对被测量值不能肯定的程度，用符号 U 表示。通过不确定度可以对被测量的真值所处的量值范围做出评定，而被测量的真值将以一定的概率（例对于标准不确定度 $P=68.3\%$ ）落在这个范围内；同时不确定度大小反映了测量结果可信程度的高低，不确定度越小，测量结果与被测量的真值越接近。

为了能更直观地反映测量结果的优劣，需要引入相对不确定度 E ，即

$$E = \frac{U}{X} \times 100\% \quad (9)$$

1.4.2 直接测量结果的不确定度估算

不确定度按其数值的评定方法可归并为两类分量：即多次测量用统计方法评定的 A 类分量 U_A ；用其它非统计方法评定的 B 类分量 U_B 。

1. A 类分量

对于多次重复测量，用算术平均值 $\bar{x}$ 表示测量结果，则可用算术平均值的标准偏差（式 8）来表征 A 类不确定度分量，即 $U_A = S_x$

实际中，在只进行有限次测量时，随机误差不完全服从正态分布规律，而是服从 t 分布（又称学生分布）规律。此时对随机误差的估计，要在（式 8）基础上乘上一个与 t 分布相关的修正因子，即

$$U_A = tS_x = t \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n(n-1)}} \quad (10)$$

式中， t 为与测量次数 n 和置信概率 P 有关的量，可从下面的数据表中查得。

表 1 t 因子表

测量次数 n	2	3	4	5	6	7	8	9	10	15	20	∞
$t_{0.683}$	1.84	1.32	1.20	1.14	1.11	1.09	1.08	1.07	1.06	1.04	1.03	1.00
$t_{0.954}$	12.71	4.30	3.18	2.78	2.57	2.45	2.36	2.31	2.26	2.15	2.09	1.96
$t_{0.997}$	63.66	9.93	5.58	4.60	4.03	3.71	3.50	3.36	3.25	2.98	2.86	2.58

从上表可以看出,当测量次数是 5 次以上时,对应置信概率是 68.3%的 t 因子, $t_{0.683} \approx 1$ 所以在实验中当测量次数在 5 次以上时,式(10)可简化为

$$U_A = S_x \quad (11)$$

2. B 类分量

B 类不确定度的来源一般应包含以下三种:仪器误差、估读误差和灵敏度误差,物理实验中一般只考虑仪器误差所带来的总不确定度的 B 类分量。仪器误差是指误差限,即在正确使用仪器的条件下,测量结果与真值之间可能产生的最大误差,用 $\Delta_{\text{仪}}$ 表示,物理实验常用仪器的仪器误差见表 2。在仅考虑仪器误差的情况下, B 类分量的 U_B 为

$$U_B = \frac{\Delta_{\text{仪}}}{C} \quad (12)$$

式中, C 为置信因子,是一个与仪器误差在 $[-\Delta_{\text{仪}}, \Delta_{\text{仪}}]$ 范围内的概率分布有关的常数。当置信概率 $P=68.3\%$ 时,对应的仪器误差如果服从正态分布、均匀分布和三角分布三者之一,相应的 C 分别取 3、 $\sqrt{3}$ 或 $\sqrt{6}$ 。若仪器说明书上未明确说明仪器误差的概率分布时,可按均匀分布处理,即

$$U_B = \frac{\Delta_{\text{仪}}}{\sqrt{3}} \quad (13)$$

表 2 物理实验常用仪器的仪器误差

仪 器 名 称	量 程	分度值(准确度等级)	仪 器 误 差
钢直尺	0 ~ 300 mm	1 mm	$\pm 0.1 \text{ mm}$
钢卷尺	0 ~ 1000 mm	1 mm	$\pm 0.5 \text{ mm}$
游标卡尺	0 ~ 300 mm	0.02, 0.05, 0.1 mm	分度值
螺旋测微计(一级)	0 ~ 100 mm	0.01 mm	$\pm 0.004 \text{ mm}$
TW-1 物理天平	1000 g	100 mg	$\pm 50 \text{ mg}$
WL-1 物理天平	1000 g	50 mg	$\pm 50 \text{ mg}$
TG928A 矿山天平	200 g	10 mg	$\pm 5 \text{ mg}$
水银温度计	$-30 \sim 300^\circ\text{C}$	0.2, 0.1 $^\circ\text{C}$	分度值
读数显微镜		0.01 mm	$\pm 0.004 \text{ mm}$

数字式测量仪器			最末一位的一个单位 或按仪器说明估算
指针式电表		$a = 0.1, 0.2, 0.5, 1.0, 1.5, 2.5, 5.0$	$\pm \text{量程} \times a\%$

3. 总不确定度的合成

总不确定度由 A 类分量和 B 类分量按“方、和、根”的方法合成，即

$$U = \sqrt{U_A^2 + U_B^2} \quad (14)$$

上式中，由于 A 类分量是由标准偏差表示的，所以上述不确定度称为合成标准不确定度，其置信概率为 68.3%。

4. 单次测量的不确定度

在实验中，有的被测量因为各种原因只能测量一次。例如，有些物理量是随时间变化的，无法进行重复测量。有些量因为对它的测量精度要求不高，没有必要进行重复测量，或因估算出的 U_A 对实验的最后结果影响甚小，这时的不确定度估算只能根据仪器误差、测量方法、实验条件以及操作者技术水平等实际情况，进行合理估计。

约定用仪器误差或其数倍作为单次直接测量的不确定度的估计值。当取 $U = \Delta_{\text{仪}}$ 时，并不意味着只测一次比多次测量时 U 的值小，只说明 $\Delta_{\text{仪}}$ 和用 $\sqrt{U_A^2 + U_B^2}$ 估算出的结果相差不多。

1.4.3 误差传递 间接测量结果的不确定度合成

直接测量的结果有误差，由直接测量值经过函数运算而得到的间接测量的结果也会有误差，这就是误差的传递。

设间接测量量 N 与各独立的直接测量量 $x, y, z, \dots$ 的函数关系为

$$N = f(x, y, z, \dots) \quad (15)$$

在对各直接测量 $x, y, z, \dots$ 进行有限次测量的情况下，将各直接测量的最佳值代入上式，即得到间接测量(最佳)值。

设 $x, y, z, \dots$ 的不确定度为 $U_x, U_y, U_z, \dots$ ，它们必然影响间接测量的结果，使 N 值也有相应的不确定度 U_N 。由于不确定度都是微小的量，相当于数学中的“增量”，只要用不确定度 U_x, U_y 等替代微分 dx, dy 等，再采用某种合成方法，就可得到不确定度传递公式。一般是用“方、和、根”形式合成。因此间接测量的不确定度计算公式与数学中的全微分公式基本相同。当函数表达式仅为“和差”形式，可用(16)式计算

$$U_N = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 U_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 U_y^2 + \left(\frac{\partial f}{\partial z}\right)^2 U_z^2 + \dots} \quad (16)$$

当函数表达式为“积和商（或积商和差混合）”形式，应先对间接测量量 $N = f(x, y, z, \dots)$ 函数式两边取自然对数，再求全微分可得到计算相对不确定度的公式如下

$$E_N = \frac{U_N}{N} = \sqrt{\left(\frac{\partial \ln f}{\partial x}\right)^2 U_x^2 + \left(\frac{\partial \ln f}{\partial y}\right)^2 U_y^2 + \left(\frac{\partial \ln f}{\partial z}\right)^2 U_z^2 + \dots} \quad (17)$$

已知 E_N 和 N 可以求出合成不确定度

$$U_N = N \cdot E_N \quad (18)$$

式(16)或(17)还常用来分析各直接测量量的误差对最后结果误差的影响大小,从而为设计或改进实验方案、选择测量仪器等提供重要依据。

1.4.4 测量结果的表示

一个完整的测量结果不仅要给出该量值的数值和单位,同时还要给出它的不确定度。

若被测量的量值为 X 、总不确定为 U , 测量结果应表示成

$$x = X \pm U \text{ (单位)} \quad (P=68.3\%) \quad (19)$$

该式表明被测量的真值将以 68.3% 的概率落在区间 $(X-U, X+U)$ 内。

不确定度一般取 1 位有效数字, 当不确定度值的首位数字是 1 和 2 时可取两位, 尾数采用“只进不舍”的原则。

【例 1】用天平 (仪器误差 $\Delta_{\text{仪}} = 0.02g$) 测量物体质量 m 九次, 测量数据如下:

i	1	2	3	4	5	6	7	8	9
m_i / g	18.79	18.72	18.75	18.71	18.74	18.73	18.78	18.76	18.77

求出测量结果。

解 1) 求平均值

$$\bar{m} = \frac{1}{9} \sum_{i=1}^9 m_i = \frac{1}{9} (18.79 + 18.72 + \dots + 18.77) = 18.75g$$

2) A 类不确定度 ($n=9$, 查表 1 得 $t=1.07$)

$$\begin{aligned} U_A &= t S_{\bar{m}} = t \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^9 (m_i - \bar{m})^2} \\ &= 1.07 \sqrt{\frac{1}{9 \times 8} [(18.79 - 18.75)^2 + (18.72 - 18.75)^2 + \dots + (18.77 - 18.75)^2]} = 0.010g \end{aligned}$$

3) B 类不确定度

$$U_B = \frac{\Delta_{\text{仪}}}{\sqrt{3}} = 0.012g$$

4) 合成不确定度

$$U_m = \sqrt{U_A^2 + U_B^2} = \sqrt{0.010^2 + 0.012^2} (g) = 0.016 (g)$$

5) 测量结果表示为

$$m = (18.75 \pm 0.02)g \quad (P=68.3\%)$$

计算结果表明, m 的真值以 68.3% 的置信概率落在 $[18.73g, 18.77g]$ 内。

【例 2】已知某铜环的外径 $D = (2.995 \pm 0.006)\text{cm}$, 内径 $d = (0.997 \pm 0.003)\text{cm}$,

高度 $H = (0.9516 \pm 0.0005)\text{cm}$, 求该铜环的体积及其不确定度, 并写出测量结果。

$$\text{解 } V = \frac{\pi}{4}(D^2 - d^2)H = \frac{3.1416}{4}(2.995^2 - 0.997^2) \times 0.9516 = 5.961(\text{cm}^3)$$

$$\ln V = \ln \frac{\pi}{4} + \ln(D^2 - d^2) + \ln H$$

$$\frac{\partial \ln V}{\partial D} = \frac{2D}{D^2 - d^2}, \quad \frac{\partial \ln V}{\partial d} = -\frac{2d}{D^2 - d^2}, \quad \frac{\partial \ln V}{\partial H} = \frac{1}{H}$$

$$\frac{U_V}{V} = \sqrt{\left(\frac{2D}{D^2 - d^2}\right)^2 U_D^2 + \left(-\frac{2d}{D^2 - d^2}\right)^2 U_d^2 + \left(\frac{1}{H}\right)^2 U_H^2}$$

$$= \sqrt{\left(\frac{2 \times 2.995 \times 0.006}{2.995^2 - 0.997^2}\right)^2 + \left(\frac{2 \times 0.997 \times 0.003}{2.995^2 - 0.997^2}\right)^2 + \left(\frac{0.0005}{0.9516}\right)^2}$$

$$= 0.0046$$

$$U_V = 0.0046 \times V = 0.0046 \times 5.961 = 0.027(\text{cm}^3)$$

$$\text{所以} \quad V = (5.961 \pm 0.027)\text{cm}^3$$

1.5 有效数字及运算规则

任何测量结果都有误差, 直接测量的读数在数据记录时应取几位数字? 间接测量中的数据运算又要保留几位数字? 这就是本节要介绍的内容。

1.5.1 有效数字的概念

为了理解有效数字的概念, 先举一个例子。用 300 mm 长、毫米分度钢直尺测量某物体的长度为 75.4 mm, 其中 75 是直接读出来的称为准确数字, 而最后一位 4 是估读出来的称为欠准确数字。准确数字和欠准确数字的全体称为**有效数字**。有效数字数字的个数叫做**有效数字的位数**。前述的 75.4 mm 为三位有效数字。

在进行测量时由于仪器多种多样，读数规则也略有区别，正确的读数方法可大致归纳如下：

(1) 一般读数应在最小分度值下再估读一位，但不一定估计到最小分度值的 $1/10$ ，当仪表的分度较窄、指针较粗或根据分度的数值可估读至 $1/5$ 或 $1/2$ 分度。

(2) 对于数字式仪表及步进读数仪器（例电阻箱）不需要进行估读，所显示的数字末位就是欠准确数字。

(3) 游标类量具只读到游标分度值，一般不估读。

(4) 有时读数的估读位取在最小分度位。如仪器的最小分度值为 0.5 ，则 0.1 、 0.2 、 0.3 、 0.4 及 0.6 、 0.7 、 0.8 、 0.9 都是估读的。这类情况不必估读到下位。

(5) 在读取整刻度值时，则必须补“0”。例如，用毫米分度钢直尺测出某物体的长度正好是 75 mm 整，应该记录为 75.0 mm ，不能写成 75 mm 。

(6) 间接测量结果的有效数字，应先算出不确定度，再由不确定度来决定。例由长度和直径测量值用计算器算出圆柱体体积为 $V=596.135\text{ mm}^3$ ，不确定度 $U_V=3\text{ mm}^3$ ，可知测量值的第三位已经是欠准确的，它后面的“135”已无保留的意义，所以测量结果应写为 $V=(596\pm 3)\text{ mm}^3$ 。

有效数字位数越多，测量的准确度就越高。容易证明，有效数字多一位，相对误差 E 差不多要小一个数量级。

有效数字书写时应注意以下几点：

(1) 有效数字的位数与小数点位置或采用的单位无关。如 75.4 mm 和 0.0754 m 都是三位有效数字，数字“7”左边的0只是表示小数点位置，而非有效数字。

(2) 为便于书写，对于较大或较小的数值，常用 $\times 10^{\pm n}$ 的形式(n 为正整数)来表示，通常在小数点前只写一位数字，例如 $6371\text{ km}=6.371\times 10^6\text{ m}$ ，都表示有四位有效数字，不能写成 $6371\text{ km}=6371000\text{ m}$ 。

(3) 表示测量最后结果的有效数字尾数与不确定度的尾数要对齐，**测量值的尾数采用“四舍六入五凑偶”规则。**

1.5.2 有效数字的运算规则

间接测量的计算过程即为有效数字的运算过程，严格说来，应根据测量误差或不确定度来确定测量值的有效数字位数。但是在不确定度估算之前，可根据下列运算规则进行粗略计算。当一位准确数字和一位欠准确数字做运算时其结果是欠准确数字。**有效数字运算取舍法则是：运算结果保留一位欠准确数字。**

1. 加减运算

几个数相加减时，计算结果的有效数字**末位**应与参与运算的各量中欠准确位数最高的**末位**对齐。

$$\text{例如：} \quad 21\bar{7} + 14.\bar{8} = 23\bar{2} \quad 116.\bar{9} - 1.65\bar{2} = 115.\bar{2} \quad 97.\bar{4} + 6.23\bar{8} = 103.\bar{6}$$

式中，在数字上方加一短线的为欠准确数字，第三个算式为有进位的情况。

2. 乘除运算

几个数相乘除时，计算结果的有效数字位数和参与运算的各量中有效数字位数最少的相同。

例如： $1.111\bar{1} \times 1.1\bar{1} = 1.2\bar{3}$

$$\begin{array}{r} 1.111\bar{1} \\ \times 1.1\bar{1} \\ \hline 1\bar{1}1\bar{1}1 \\ 1111\bar{1} \\ \hline 1.2\bar{3}3\bar{2}1 \end{array}$$

由此可知，如果间接测量是由几个直接测量值通过乘除运算得到，设计实验方案时应考虑各直接测量值的有效数字位数要基本相仿，否则精度过高的测量就失去意义了。

3. 乘方、立方、开方运算

运算结果的有效数字位数与底数的有效位数相同。

4. 对数和三角函数运算

对数、三角函数的运算，当直接测量量的不确定度给出时，可从不确定度传递公式计算出不确定度，根据不确定度确定其有效数字的位数。

例如： $a = 3068 \pm 2$ ，求 $y = \ln a$ 。按照不确定度传递公式

$$U_y = \frac{1}{a} U_a = \frac{1}{3068} \times 2 = 0.0007$$

所以 $y = \ln a = 8.0288$ 。

又如： $\theta = 60^\circ 0' \pm 3'$ ，求 $x = \sin \theta$ 。由不确定度传递公式

$$U_x = |\cos \theta| U_\theta = |\cos 60^\circ| \frac{3 \times \pi}{60 \times 180} = 0.0004$$

所以 $x = \sin 60^\circ 0' = 0.8660$ 。

当直接测量的不确定度未给出时，上述过程可简化为通过改变自变量末位的一个单位，观察函数运算结果的变化情况来确定其有效数字位数。例如 $\alpha = 20^\circ 6'$ 中的“6'”是欠准确数字，由计算器运算结果为 $\sin 20^\circ 6' = 0.343659695\cdots$ ， $\sin 20^\circ 7' = 0.343932851\cdots$ ，两种结果在小数点后面第四位出现了差异，所以 $\sin 20^\circ 6' = 0.3436$ 。

5. 常数

π 、 e 、 $\sqrt{2}$ 等常数的有效数字位数是无限的，应根据运算需要合理取值。一般情况下当常数参与加减运算时小数点后多取一位，当常数参与乘除运算时比参与运算的数据多取一位。例如：

$$S = \pi r^2, \quad r = 6.042 \text{ cm}, \quad \pi \text{ 取为 } 3.1416, \quad \therefore S = 3.1416 \times 6.042^2 = 114.7 \text{ cm}^2.$$

$$\theta = 129.3 + \pi, \quad \pi \text{ 取为 } 3.14, \quad \theta = 129.3 + 3.14 = 132.4 \text{ rad}.$$

采用有效数字运算规则，可以保证测量结果的准确度不因数字取舍不当而受到影响。

1.6 实验数据处理基本方法

数据处理是指从获得数据开始到得出最后结论的整个加工过程，包括数据记录、整理、计算、分析和绘制图表等。这里仅介绍一些基本的数据处理方法。

1.6.1 列表法

列表是有序记录原始数据的必要手段，也是用实验数据显示函数关系的原始方法。将数据按一定的规律列成表格使得数据表达清晰、条理化，易于检查数据和发现问题，有助于分析物理量之间的相互关系和规律。

在设计表格时要注意以下几点：

- (1) 表格的上方要写明表格的名称。
- (2) 各栏目均应注明所记录的物理量的名称(符号)、单位和量值的数量级。
- (3) 栏目的顺序应充分注意数据间的联系和计算顺序，力求简明、齐全、有条理。
- (4) 表中的原始测量数据应正确反映有效数字，数据不应随意涂改，确实要修改数据时，应将原来数据画条斜杠以供备查，把修正的数据写在旁边。

1.6.2 图示法和图解法

图线不仅能够直观地显示物理量之间相互的关系、变化趋势，而且能够从中找出变量的极值、转折点、周期性和某些奇异值等。如果通过内插法或外推法，可以从图线上直接读出没有进行观测的点的数值。

由于直线最易描绘，且直线方程的两个参数(斜率和截距)也较易算得，所以对于两个变量之间的函数关系是非线性的情形，在用图解法时应尽可能通过变量代换将非线性的函数转变为线性函数。例如， $y = ax^b$ (a 和 b 为常数)，等式两边取对数得 $\lg y = \lg a + b \lg x$ ，于是， $\lg y$ 与 $\lg x$ 为线性关系， b 为斜率， $\lg a$ 为截距。

图示和图解法的基本步骤和规则如下：

(1) 选择坐标纸。常用坐标纸有直角坐标纸(即毫米方格纸)、对数坐标纸、半对数坐标纸和极坐标纸等。一般图上最小分格对应测量数据的最后一位可靠数字，即坐标轴上的最小分度(1mm)对应于实验数据的最后一位准确数字

(2) 确定坐标轴、比例和分度。通常横坐标表示自变量，纵坐标表示因变量，用粗实线在坐标纸上描出坐标轴。一个坐标轴应包括四要素：物理量的名称(符号)、单位、轴的方向及等间隔标注的分度值。坐标轴的起点一般不从零开始，用略小于实验数据最小值的某一整齐数作为起点，略大于实验数据最大值的某一整齐数作为终点。

坐标比例是指坐标轴上单位长度(通常为1cm)所代表的物理量大小。为了便于读数和防止损失有效数字位数，应该选每厘米代表“1”、“2”、“5”及其倍率的比例，切勿采用“3”、“7”、“9”等的比例。通过选取合适的比例和坐标轴起点，使作出的曲线充满图纸。

比例确定后应对坐标轴进行分度，即在坐标轴上均匀地(一般每隔2cm)标出所代表物

理量的整齐数值，不要标注实验测量数据。

(3) 描点。用直尺和笔尖清楚地将实验数据点在图纸上准确地用“+”号标出。若在同一张图纸上同时作几条实验曲线，各条曲线的实验数据点应该用不同符号(如×、⊙等)标出，以示区别。

(4) 连线。使用曲线板或透明直尺将实验数据点拟合成光滑的曲线或直线(若是校准曲线应连成折线)。图线不一定要通过所有实验数据点，实验点应均匀分布在图线两侧，且离图线距离尽可能小。个别偏离曲线较远的点，应检查标点是否错误，若属错误数据，在连线时不予考虑。

(5) 图注与说明。在图纸的明显位置写明图线的名称、比例、必要的说明(主要指实验条件、数据来源)、作者及日期等。

(6) 图解法求经验公式。根据已作好的图线，用数学知识求出待定常数，得到曲线方程或经验公式即为图解法。当函数关系为线性关系时，步骤如下：

第一，取点。在直线上靠近实验数据两端点的内侧取两点 $A(x_1, y_1)$ 、 $B(x_2, y_2)$ ，并用不同于实验点的符号标明，注明其坐标值(注意有效数字)。

第二，求斜率和截距。设直线方程为 $y = a + bx$ ，则

$$b = \frac{y_2 - y_1}{x_2 - x_1} \quad (20)$$

$$a = \frac{x_2 y_1 - x_1 y_2}{x_2 - x_1} \quad (21)$$

注意：解析点不能采用测量数据点，斜率不能用纵坐标和横坐标的几何长度比值求出！

【例 3】金属电阻与温度的关系可近似表示为 $R = R_0(1 + \alpha t)$ ， R_0 为 $t = 0^\circ\text{C}$ 时的电阻， α 为电阻的温度系数。实验数据见表 3，试用图解法建立电阻与温度关系的经验公式。

表 3

i	1	2	3	4	5	6	7
$t/^\circ\text{C}$	10.5	26.0	38.3	51.0	62.8	75.5	85.7
R/Ω	10.423	10.892	11.201	11.586	12.025	12.344	12.679

解 比例选择。 $\frac{90.0 - 10.0}{17} = 4.7$ ，故取为 $5.0^\circ\text{C} \cdot \text{cm}^{-1}$ ； $\frac{12.800 - 10.400}{25} = 0.096$ ，故

取为 $0.100\Omega \cdot \text{cm}^{-1}$ 。所绘图线见图 3。

在图线上取两点 $A(13.0, 10.500)$ 和 $B(83.5, 12.600)$ ，斜率和截距计算如下：

$$b = \frac{y_B - y_A}{x_B - x_A} = \frac{12.600 - 10.500}{83.5 - 13.0} = \frac{2.100}{70.5} = 0.0298(\Omega/^\circ\text{C})$$

$$R_0 = R_1 - bt_1 = 10.500 - 0.0298 \times 13.0 = 10.500 - 0.387 = 10.113(\Omega)$$

$$\alpha = \frac{b}{R_0} = \frac{0.0298}{10.113} = 2.95 \times 10^{-3} / ^\circ\text{C}$$

所以，铜丝电阻与温度的经验公式为

$$R = 10.113(1 + 2.95 \times 10^{-3}t)\Omega$$

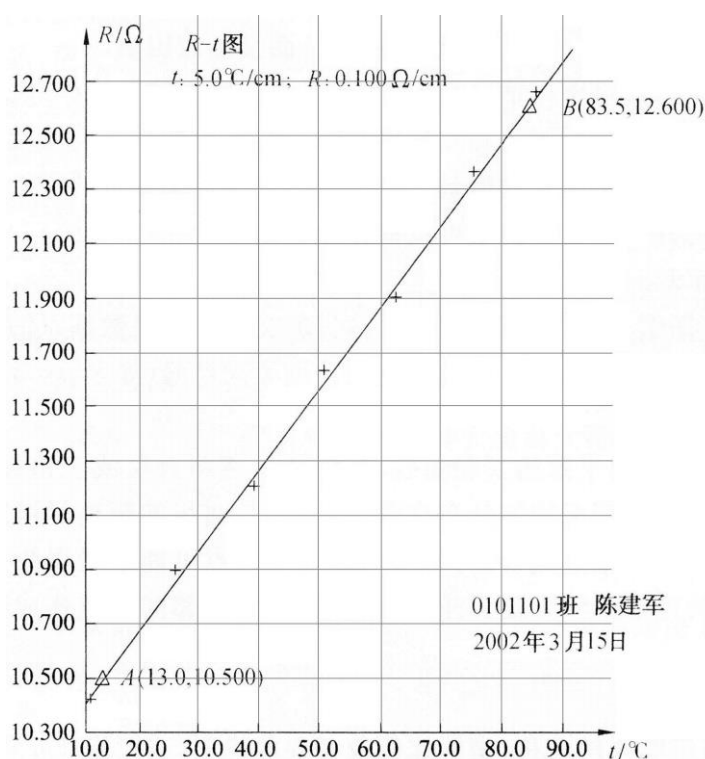


图3 铜丝电阻与温度关系曲线

1.6.3 逐差法

当因变量和自变量之间存在线性关系，且自变量为等间距变化的情况下，逐差法既能充分利用实验数据，又具有减小随机误差的效果。具体做法是将测量得到的偶数组数据分成前后两组，将对应项分别相减，然后再求平均值。举例说明如下：

在弹性限度内弹簧的伸长量 x 与所受的载荷(拉力) F 满足线性关系 $F = kx$ ，等差地改变载荷，测得一组实验数据，见表4。

表4

砝码质量/Kg	1.000	2.000	3.000	4.000	5.000	6.000	7.000	8.000
弹簧伸长位置/cm	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8

求每增加 1Kg 砝码弹簧的平均伸长量 Δx 。

将上述数据分成前后两组，前一组 (x_1, x_2, x_3, x_4) ，后一组 (x_5, x_6, x_7, x_8) ，然后对应项相减求平均，即

$$\Delta x = \frac{1}{4 \times 4} [(x_5 - x_1) + (x_6 - x_2) + (x_7 - x_3) + (x_8 - x_4)]$$

逐差法计算简便，能及时发现数据规律或错误数据。

1.6.4 最小二乘法

设物理量 y 和 x 之间满足线性关系 $y = a + bx$ ，由一组实验数据拟合出一条最佳直线，常用的方法是最小二乘法。最小二乘法的基本原理是：对于等精度测量，若存在一条最佳的拟合曲线，那么各测量值与这条曲线上对应点之差的平方和应取最小值。

假定 x 和 y 值中只有 y 有明显的随机误差，实验测量数据为 $(x_i, y_i; i=1, 2, \dots, n)$ ，其中 $x = x_i$ 时对应 $y = y_i$ 。由于测量总是有误差的，我们将这些误差归结为 y_i 的测量偏差，记为

$$\varepsilon_i = y_i - (a + bx_i),$$

根据最小二乘法的原理，偏差的平方和应最小，即

$$\sum_{i=1}^n \varepsilon_i^2 \rightarrow \min$$

令

$$S = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - a - bx_i)^2$$

使 S 为最小的条件是

$$\frac{\partial S}{\partial a} = 0, \quad \frac{\partial S}{\partial b} = 0, \quad \frac{\partial^2 S}{\partial a^2} > 0, \quad \frac{\partial^2 S}{\partial b^2} > 0$$

由一阶微商为零得

$$\left. \begin{aligned} \frac{\partial S}{\partial a} &= -2 \sum_{i=1}^n (y_i - a - bx_i) = 0 \\ \frac{\partial S}{\partial b} &= -2 \sum_{i=1}^n (y_i - a - bx_i) x_i = 0 \end{aligned} \right\}$$

解得

$$a = \frac{\sum_{i=1}^n x_i \sum_{i=1}^n (x_i y_i) - \sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i}{\left(\sum_{i=1}^n x_i \right)^2 - n \sum_{i=1}^n x_i^2} \quad (22)$$

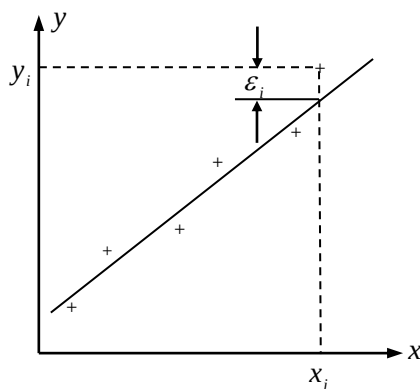


图 4 y_i 的测量偏差

$$b = \frac{\sum_{i=1}^n x_i \sum_{i=1}^n y_i - n \sum_{i=1}^n (x_i y_i)}{\left(\sum_{i=1}^n x_i \right)^2 - n \sum_{i=1}^n x_i^2} \quad (23)$$

令 $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$, $\bar{x}^2 = \left(\frac{1}{n} \sum_{i=1}^n x_i \right)^2$, $\overline{x^2} = \frac{1}{n} \sum_{i=1}^n x_i^2$, $\overline{xy} = \frac{1}{n} \sum_{i=1}^n (x_i y_i)$, 则

$$a = \bar{y} - b\bar{x} \quad (24)$$

$$b = \frac{\bar{x} \cdot \bar{y} - \overline{xy}}{\bar{x}^2 - \overline{x^2}} \quad (25)$$

如果实验是在已知 y 和 x 满足线性关系下进行的, 那么用上述最小二乘法线性拟合 (又称一元线性回归) 可解得斜率 b 和截距 a , 从而得出回归方程 $y = a + bx$ 。如果实验是要通过对 x 、 y 的测量来寻找经验公式, 则还应判断由上述一元线性拟合所确定的线性回归方程是否恰当, 可用下列相关系数 r 来判别

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sqrt{(\overline{x^2} - \bar{x}^2)(\overline{y^2} - \bar{y}^2)}} \quad (26)$$

其中, $\bar{y}^2 = \left(\frac{1}{n} \sum_{i=1}^n y_i \right)^2$, $\overline{y^2} = \frac{1}{n} \sum_{i=1}^n y_i^2$ 。

可以证明, $|r|$ 值总是在 0 和 1 之间。 $|r|$ 值越接近 1, 说明实验数据点密集地分布在所拟合直线的近旁, 用线性函数进行回归是合适的。 $|r|=1$ 表示变量 x 、 y 完全线性相关, 拟合直线通过全部实验数据点。 $|r|$ 值越小线性越差, 一般 $|r| \geq 0.9$ 时可认为两个物理量之间存在较密切的线性关系, 此时用最小二乘法直线拟合才有实际意义。

练 习 题

1. 试判断下列测量是直接测量还是间接测量? 你还能举出哪些例子?

- (1) 用弹簧测量力的大小; (2) 用天平称物体质量;
(3) 用伏安法测量电阻; (4) 用单摆测量重力加速度。

2. 试比较下列测量的优劣:

- (1) $x_1 = (55.98 \pm 0.03) \text{ mm}$; (2) $x_2 = (0.488 \pm 0.004) \text{ mm}$;
(3) $x_3 = (0.0098 \pm 0.0012) \text{ mm}$; (4) $x_4 = (1.98 \pm 0.05) \text{ mm}$ 。

3. 用电子秒表 ($\Delta_{\text{仪}} = 0.01 \text{ s}$) 测量单摆摆动 20 个周期的时间 t , 测量数据如表 5 所示。

表 5

i	1	2	3	4	5	6	7
-----	---	---	---	---	---	---	---

t/s	20.12	20.19	20.11	20.13	20.14	20.12	20.17
-------	-------	-------	-------	-------	-------	-------	-------

试求周期 T 及测量不确定度，并写出测量结果。

4. 利用单摆测重力加速度 g ，当摆角 $\theta < 5^\circ$ 时有 $T \approx 2\pi\sqrt{l/g}$ 的关系。式中， l 为摆长， T 为摆动周期，它们的测量结果分别为 $l = (97.69 \pm 0.03)\text{cm}$ ， $T = (1.984 \pm 0.023)\text{s}$ ，试求重力加速度 g 的测量值及其不确定度，并写出测量结果。

5. 试推导下列间接测量的不确定度合成公式。

$$(1) f = \frac{uv}{u+v}; \quad (2) f = \frac{D^2 - L^2}{4D}; \quad (3) n = \frac{\sin \frac{1}{2}(\alpha + \delta)}{\sin \frac{\alpha}{2}}.$$

6. 已知某圆柱体的质量 $m = (236.12 \pm 0.05)\text{g}$ ，直径 $d = (2.345 \pm 0.005)\text{cm}$ ，高 $h = (8.21 \pm 0.05)\text{cm}$ 。求圆柱体的密度及不确定度，并分析直接测量值 m 、 d 和 h 的不确定度对间接测量值 ρ 的影响程度的大小。

7. 改正下列错误，写出正确答案。

- (1) 0.10860 的有效数字为六位；
- (2) $P = (31690 \pm 300)\text{kg}$ ；
- (3) $d = (10.8135 \pm 0.0176)\text{cm}$ ；
- (4) $E = (1.98 \times 10^{11} \pm 3.27 \times 10^9)\text{N/m}^2$ ；
- (5) $g = (9.795 \pm 0.0036)\text{m/s}^2$ ；
- (6) $R = 6371\text{km} = 6371000\text{m} = 637100000\text{cm}$ 。

8. 根据有效数字运算规则计算下列各题：

$$(1) \frac{76.013}{40.03 - 2.0} = ; \quad (2) \frac{50.00 \times (18.30 - 16.3)}{(103 - 3.0)(1.00 + 0.001)} = ;$$

$$(3) \frac{25^2 + 493.0}{\ln 406.0} = ; \quad (4) \frac{\sin \frac{1}{2}(60^\circ 2' + 51^\circ 20')}{\sin 30^\circ 1'} = .$$

9. 一定质量的气体，当体积一定时压强与温度的关系为

$$p = p_0(1 + \beta t) \quad (\text{SI 单位})$$

通过实验测得一组数据如表 6 所示。

表 6

i	1	2	3	4	5	6	7
$t/^\circ\text{C}$	7.5	16.0	23.5	30.5	38.0	47.0	54.5
p/Pa	0.98×10^5	1.02×10^5	1.03×10^5	1.07×10^5	1.09×10^5	1.12×10^5	1.15×10^5

试用作图法求出 p_0 、 β ，并写出经验公式。

10. 试用最小二乘法对习题 9 的数据进行直线拟合，求出 p_0 和 β 。

11. 用伏安法测量电阻的实验数据如表 7 所示。

表 7

i	1	2	3	4	5	6	7	8
I/mA	2.00	3.00	4.00	5.00	6.00	7.00	8.00	9.00
U/V	0.540	0.782	1.025	1.265	1.510	1.750	1.995	2.240

试用逐差法求电阻 R 。