第 1 章　误差理论和数据处理方法

1.1　误差的来源及分类

1.1.1　误差理论中的基本概念

在对某物理量进行多次测量时，经过大量的实践证明，无论仪器多么精确，测量者怎么仔细，各测量值之间总会存在着差异，而且被测对象的真值与测量值之间也存在一定的差异，这种差异通常称为测量误差。

一种材料的抗拉强度、弹性模量、泊松比等，一根试样的尺寸，一个砝码的质量都存在一个客观的、真正的值，称为真值。对某种材料的抗拉强度、弹性模量、泊松比、试样尺寸的测量和砝码的称重都会得到一个实际测定的数值，称为测量值。真误差 Δ_i 的定义为被测对象的测量值 l_i 与真值 x 所得的差值，也称为绝对误差，即

$$\Delta_i = l_i - x \tag{1.1}$$

一般情况下，被测对象的真值难以获取，为了进行绝对误差的计算，可以采用被测对象的多次测量的平均值来代替真值。绝对误差反映测量值相对于真值的偏差大小，其单位与给出值单位相同。由于绝对误差有时难以反映不同被测对象的测量精度，如测量两种不同规格试样的尺寸分别为 $l_1=100\text{mm}$ 和 $l_2=50\text{mm}$，如果测量绝对误差 $\Delta_1=\Delta_2=0.2\text{mm}$，则这两根试样的测量精度显然不同，而绝对误差并不能反映这种差别。因此，在这种情况下，工程上一般采用相对误差 k 来进行精度评定，相对误差是用绝对误差与被测对象的测量真值之比来描述的，即

$$k = \frac{\Delta}{x} \times 100\% \tag{1.2}$$

1.1.2　误差的来源

产生测量误差的主要原因如下：一是仪器误差，由于测量仪器的构造不可能十分完善；二是测量误差，由于测量者的感觉器官的鉴别能力和技术水平与经验的限制；三是环境误差，由于测量需要在一定的外界条件下进行，所以测量结果必然会含有误差。将仪器条件、测量条件、外界条件称为测量的三大客观（测量）条件。三大客观条件相同的测量称为等精度测量；三大客观条件不同的测量称为不等精度测量。

1. 仪器误差

该误差通常包括实验设备、测量仪器及仪表带来的误差，如安装调试不准确、刻度不准确、设备加工粗糙、仪表非线性及元器件之间的间隙造成的误差。

2. 测量误差

该误差通常包括测量方法不准确而引起的误差，以及测量者的视觉分辨能力、熟练程

度和精神状态等引起的误差。

3. 环境误差

由外界环境引起的测量误差主要指测量环境的温度、气压、湿度、电场、磁场等与要求的标准状态不一致引起的误差。

1.1.3 误差的分类

误差按其对测量结果影响的性质分为系统误差和偶然(随机)误差两大类。另外,在测量结果中有时还会出现测量错误,也称粗差,如读错、记错、测错等均属于粗差。粗差在测量结果中是不允许出现的,它不属于误差的范畴。为了防止粗差,通常在测量中除仔细认真地工作外,还要采取一定的检核措施,以发现是否有粗差存在。

1. 系统误差

在相同的测量条件下,对某量进行一系列的测量,若误差出现的符号、数值的大小均一致,或者按一定的规律变化,则称这种误差为系统误差。例如,用名义长度为 10cm,而实际长度为 10.05cm 的游标卡尺测量某一试样的直径,其测量结果必然会含有系统误差。

2. 偶然误差

在相同的测量条件下,对同一对象进行一系列的测量,若误差出现的大小和符号均不一致,且从表面上看没有任何规律性,则称这种误差为偶然误差。如读数时的估读误差。对于单个的偶然误差,测量前无法预料其出现的符号和大小,但就大量的偶然误差来研究,它具有一定的规律性,并且随着测量次数的增多,这种统计规律越是明显。认识这种规律,可以更好地指导测量实践。偶然误差是误差理论研究的主要内容。

1.2 偶然误差的性质

偶然误差表面上无规律可寻,但受其内部必然规律的支配。实践表明:对某量进行多次等精度的重复测量,得到一系列不同的测量值,在只含有偶然误差的情况下,其误差出现统计学上的规律性。测量次数越多,规律性越明显。如掷硬币,出现正反面的机会随次数的增多而趋于相等。

1.2.1 偶然误差的特性

例如,在相同的测量条件下,对某一截面为三角形的试样的截面内角独立地进行了 360 次测量,每个三角形的内角和与其理论值 180°之差,即为该三角形内角和的真误差 Δ_i

$$\Delta_i = a_i + b_i + c_i - 180^\circ$$

取误差区间间隔 $d\Delta = 0.20''$,并将 360 个真误差按其符号和大小排列,列于表 1.1 中。

表 1.1 误差分布表

误差所在区间	正误差个数	负误差个数	总数	误差所在区间	正误差个数	负误差个数	总数
0.0″~0.2″	48	49	97	1.0″~1.2″	12	12	24
0.2″~0.4″	39	38	77	1.2″~1.4″	5	6	11
0.4″~0.6″	32	32	64	1.4″~1.6″	3	4	7
0.6″~0.8″	22	25	47	1.6″以上	0	0	0
0.8″~1.0″	16	17	33	$\sum$	177	183	360

从表 1.1 中的统计结果可以得出偶然误差具有如下特性。

1. 有界性

在一定测量条件下的有限次测量中，偶然误差的绝对值不会超过一定限度。

2. 范围性

在一定测量条件下，绝对值较小的偶然误差出现的概率大，绝对值较大的偶然误差出现的概率小。

3. 对称性

在一定测量条件下，绝对值相等的正、负偶然误差出现的概率相等。

4. 抵偿性

在一定测量条件下，对同一量进行等精度观测，随着测量次数的增加，其偶然误差的代数和趋于 0，即

$$\lim_{n\to\infty}(\Delta_1+\Delta_2+\cdots+\Delta_n)=0 \tag{1.3}$$

1.2.2 偶然误差的分布密度函数

德国科学家高斯根据偶然误差的特性，总结出偶然误差服从正态分布，正态分布的偶然误差值与其出现概率之间的函数关系为

$$f(\Delta)=\frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\frac{\Delta 2}{2\sigma 2}} \tag{1.4}$$

式中：σ 为均方根差或标准差，是与测量条件有关的参数；f 为偶然误差 Δ 出现的概率密度。

将偶然误差正态分布密度函数绘成曲线，如图 1.1 所示。

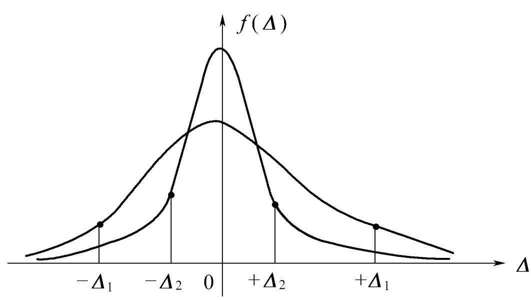

图 1.1　误差分布密度曲线

当 $\Delta\to\pm\infty$ 时，$f(\Delta)\to 0$。由于 $f(\Delta)$ 随 Δ 的增大而迅速减小，故当 Δ 达到足够大时，$f(\Delta)$ 已很小，实际上可视为零，此时的 Δ 值可以作为误差的限值。这是偶然误差的第一特性，即有界性。

若 $|\Delta_2|<|\Delta_1|$，则 $f(\Delta_2)>f(\Delta_1)$，这说明 $f(\Delta)$ 随 Δ 绝对值的增大而减小。$f(\Delta)$ 为降函数，这是偶然误差的第二特性，即范围性。

若 $\Delta_2=-\Delta_1$，则 $f(\Delta_2)=f(\Delta_1)$，这说明 $f(\Delta)$ 为偶函数，这是偶然误差的第三特性，即对称性。

积分 $\int_{-\infty}^{+\infty} f(\Delta)\mathrm{d}\Delta = 1$，表明误差落在全部区间内这一事件为必然事件。这也说明了测量结果中必然含有误差。

从图 1.1 中的两条误差分布密度曲线可以看出，曲线中部升得越高，曲线形状越陡峻，说明零附近的小误差出现的机会越多，表明测量质量好；相反，曲线中部升得越低，曲线形状越平缓，说明零附近的小误差出现的机会越少，表明测量质量差。测量质量的优劣取决于测量条件的好坏，而测量条件的好坏在误差分布密度曲线的形态上得到充分的反映。

1.3 偶然误差的精度评价指标

在一定的测量条件下，对某一量进行一系列的测量，对应着一种确定的误差分布。如果误差分布比较密集，即小误差出现的个数较多，则表示测量质量较好，也即测量精度高；反之，如果误差分布比较离散，则表示测量质量较差，也即测量精度低。反映精度高低的具体数字，称为精度评价指标。

1.3.1 算术平均值与改正数

在等精度测量条件下，对某量进行多次测量，通常取其平均值作为最后结果，认为是最可靠的。例如，对某试样直径进行 n 次测量，测量值为 $l_1, l_2, \cdots, l_n$，则该试样直径的算术平均值为

$$L = \frac{l_1 + l_2 + \cdots + l_4}{n} = \frac{1}{n}\sum_{i=1}^{n} l_i \tag{1.5}$$

由于测量仪器、测量方法、环境、人的观察力等条件的影响，物理量的真值无法测得，通常采用多次测量的平均值近似为真值。

设某量的真值为 x，测量值为 $l_i(i=1,2,\cdots,n)$，其真误差为 $\Delta_i(i=1,2,\cdots,n)$，则

$$\Delta_i = l_i - x \tag{1.6}$$

上式两端取和，得

$$\sum_{i=1}^{n}\Delta_i = \sum_{i=1}^{n} l_i - n \cdot x \tag{1.7}$$

两端除以 n，得

$$\frac{\sum_{i=1}^{n}\Delta_i}{n} = \frac{\sum_{i=1}^{n} l_i}{n} - x \tag{1.8}$$

由偶然误差的抵偿性可知

$$\lim_{n\to\infty}\frac{\sum_{i=1}^{n}\Delta_i}{n} = 0 \tag{1.9}$$

即

$$\lim_{n \to \infty} \frac{\sum_{i=1}^{n} l_i}{n} = x \tag{1.10}$$

当测量次数 n 无限增加时，测量值的算术平均值就是该未知量的真值；当测量次数 n 有限时，采用算术平均值来近似真值。

在一般情况下，被测量的真值为未知的，这时可用算术平均值代替测量的真值进行计算，测量值与算术平均值之差称为剩余误差，也叫改正数，用 v_i 表示。

$$v_i = l_i - L \tag{1.11}$$

对于等精度多次重复测量，测量值的改正数的代数和等于零，剩余误差的平方和为最小。

若将上式两端取和，有

$$\sum_{i=1}^{n} v_i = \sum_{i=1}^{n} l_i - n \cdot L = 0 \tag{1.12}$$

1.3.2 精度评价指标

1. 中误差

在一定测量条件下测量结果的精度可用标准差来衡量，标准差是偶然误差分散性的一个重要特征。但在实际测量工作中，不可能对某一量做无穷多次观测，因此，定义有限次观测的真误差来求标准差，也称为中误差。

1）根据真误差计算中误差

设在同精度测量条件下，对某量进行了 n 次测量，得测量值为 $l_1, l_2, \cdots, l_n$，相应的真误差分别为 $\Delta_1, \Delta_2, \cdots, \Delta_n$，则定义该组测量值的方差为

$$\sigma^2 = \lim_{n \to \infty} \frac{\sum_{i=1}^{n} \Delta_i^2}{n} \tag{1.13}$$

当测量次数 n 有限时，则采用中误差 m 来近似计算标准差为

$$m = \pm \sqrt{\frac{\sum_{i=1}^{n} \Delta_i^2}{n}} \tag{1.14}$$

式(1.14)也称菲列罗公式。其中，m 代表一组测量值的精度，即这组测量值中的每一个测量值都具有这样的精度，或者说同精度测量值具有相同的精度。而 Δ_i 彼此并不相同，这是由于偶然误差的性质所决定的。

中误差 m 的大小反映了一系列测量值的精度。不同的系列测量中，标准差越小，测量精度越高。若两列测量值的中误差相同，则表示二者的精度相同。

2）根据改正数计算中差

通常测量值的真值是不知道的，因此，无法计算真误差 Δ_i，因此也就不能利用菲列罗公式计算一组测量值的中误差。但是观测量的算术平均值 $L = \frac{\sum_{i=1}^{n} l_i}{n}$ 是可求的，这时可用

测量值的改正数 v_i 来计算这组测量值的中误差，从而衡量这组测量值的精度，即用贝塞尔公式计算：

$$m=\pm\sqrt{\frac{\sum_{i=1}^{n}v_i^2}{n-1}} \tag{1.15}$$

所以，在已知测量值真值时，用菲列罗公式求测量值的中误差；未知测量值真值时，用贝塞尔公式求测量值的中误差。

例 1.1 在同一观测条件下，对某试样长度进行了 5 次观测，其 5 次观测值分别为 10.234cm，10.238cm，10.236cm，10.240cm，10.242cm，试求：

（1）1 次观测的中误差；

（2）5 次观测平均值的中误差。

解：求平均值：

$$\overline{D}=\frac{10.234+10.238+10.236+10.240+10.242}{5}=10.238\text{cm}$$

各次观测的改正数：

$$v_1=D_1-\overline{D}=-0.004\text{cm}$$

$$v_2=D_2-\overline{D}=0.000\text{cm}$$

$$v_3=D_3-\overline{D}=-0.002\text{cm}$$

$$v_4=D_4-\overline{D}=0.002\text{cm}$$

$$v_5=D_5-\overline{D}=0.004\text{cm}$$

1 次观测的中误差：

$$m=\pm\sqrt{\frac{\sum_{i=1}^{5}v_i^2}{n-1}}=\pm 0.0032\text{cm}$$

5 次观测平均值的中误差：

$$m=\pm\sqrt{\frac{\sum_{i=1}^{5}v_i^2}{n(n-1)}}=\pm 0.0014\text{cm}$$

2. 相对误差

在某些测量工作中，用绝对误差还不能反映出测量质量的高低，如测量 100mm 和 50mm 的两根试样的长度尺寸，其中误差都是±0.02mm，但不能简单地认为二者的精度一样。因此，采用相对误差可以很好地衡量它们精度的高低，相对误差 k 为中误差的绝对值与被测量真值之比，一般表示为分子是 1 的分数形式，可写成

$$k=\frac{|m|}{x}\times 100\%=\frac{|m|}{x}=\frac{1}{x/|m|} \tag{1.16}$$

上例中，长度为 100mm 的试样相对误差为 $k_1=0.02/100=1/5000$，长度为 50mm 的试样相对误差为 $k_2=0.02/50=1/2500$，所以前者精度高于后者。

3. 极限误差

根据偶然误差的有界性可知，在一定的观测条件下偶然误差的绝对值不会超过一定的极限值，这个极限值就是极限误差。中误差只能代表一组观测值的精度，而不能代表某一个观测值的真误差大小，但二者之间有一定的统计学上的联系。在一系列等精度测量误差中，真误差与中误差之间具有如下概率统计规律：

$|\Delta|>|m|$ 的偶然误差出现的概率约为 30%；

$|\Delta|>2|m|$ 的偶然误差出现的概率约为 5%；

$|\Delta|>3|m|$ 的偶然误差出现的概率约为 0.3%。

故一般认为大于 $3|m|$ 的偶然误差是不可能的，所以一般取 $3|m|$ 为偶然误差的极限误差，即

$$-3|m|<\Delta_{极}<3|m| \tag{1.17}$$

测量中，取 $2m$ 为 Δ 的容许值 $\Delta_{容}$，即

$$-2|m|<\Delta_{容}<2|m| \tag{1.18}$$

若观测值的偶然误差 $2|m|<|\Delta|\leqslant 3|m|$，则认为该值不可靠(但不是错误的)，应舍去不用。

1.4 误差传播定律及其应用

在实际工作中，有些物理量是可以直接测量的，如试样的直径和长度，有些物理量是不能直接测量的，如屈服极限、强度极限、延伸率和断面收缩率等。对于这些不能直接测量的物理量必须通过一些直接测量得到的数值，按一定的公式或函数去计算而间接得到。由于各直接测定的数值都含有误差，因此，由计算得到的间接量中也必然含有误差，阐述测量值标准差与测量值函数标准差之间关系的定律，称为误差传播定律。

1.4.1 非线性函数误差传播定律及其应用

设有独立测量值 $x_1,x_2,\cdots,x_n$，其中误差分别为 $m_{x_1},m_{x_2},\cdots,m_{x_n}$，现有函数

$$z=f(x_1,x_2,\cdots,x_n) \tag{1.19}$$

求函数值 z 的中误差 m_z。

在函数 z 中，由于独立测量值 $x_1,x_2,\cdots,x_n$ 存在真误差 $\Delta_1,\Delta_2,\cdots,\Delta_n$，所以必然会引起未知量 z 产生真误差 Δ_z。

对式(1.19)进行全微分有

$$\mathrm{d}z=\frac{\partial f}{\partial x_1}\mathrm{d}x_1+\frac{\partial f}{\partial x_2}\mathrm{d}x_2+\cdots+\frac{\partial f}{\partial x_n}\mathrm{d}x_n$$

因 Δ_i、Δ_z 均很小，由数学分析可知，可用 Δ_i、Δ_z 代替全微分 $\mathrm{d}x_i$、$\mathrm{d}z$，从而有真误差关系式：

$$\Delta_z=\frac{\partial f}{\partial x_1}\Delta_1+\frac{\partial f}{\partial x_2}\Delta_2+\cdots+\frac{\partial f}{\partial x_n}\Delta_n \tag{1.20}$$

式(1.20)中，$\dfrac{\partial f}{\partial x_i}(i=1,2,\cdots,n)$是函数对各个变量所取的偏导数，以测量值代入计算得到，它们是一常数。设

$$k_i = \frac{\partial f}{\partial x_i}(i=1,2,\cdots,n)$$

则式(1.20)可写为

$$\Delta_z = k_1 \cdot \Delta_1 + k_2 \cdot \Delta_2 + \cdots + k_n \cdot \Delta_n \tag{1.21}$$

为了建立函数值与观测量之间的中误差关系式，设想对各 $x_i(i=1,2,\cdots,n)$ 均测量 N 次，则可以写出 N 个类似式(1.21)的关系式

$$\begin{cases} \Delta z_1 = k_1 \cdot \Delta_{11} + k_2 \cdot \Delta_{21} + \cdots + k_n \cdot \Delta_{n1} \\ \Delta z_2 = k_1 \cdot \Delta_{12} + k_2 \cdot \Delta_{22} + \cdots + k_n \cdot \Delta_{n2} \\ \vdots \\ \Delta z_N = k_1 \cdot \Delta_{1N} + k_2 \cdot \Delta_{2N} + \cdots + k_n \cdot \Delta_{nN} \end{cases} \tag{1.22}$$

上式两端平方后再求和，得

$$\sum_{i=1}^{N} \Delta z_i^2 = k_1^2 \cdot \sum_{i=1}^{N} \Delta_{1i}^2 + k_2^2 \cdot \sum_{i=1}^{N} \Delta_{2i}^2 + \cdots + k_n^2 \cdot \sum_{i=1}^{N} \Delta_{ni}^2 + 2\sum_{p=1}^{N}\Big(\sum_{\substack{i=1,j=1\\ i\neq j}}^{n} k_i k_j \Delta_{ip}\Delta_{jp}\Big) \tag{1.23}$$

根据偶然误差的特性，当 $i \neq j$ 时，独立测量量 x_i、x_j 的偶然误差 Δ_i、Δ_j 之乘积 $\Delta_i \cdot \Delta_j$ 也表现为偶然误差。依据偶然误差的抵偿性，有

$$\lim_{N\to\infty} \sum_{p=1}^{N}\Big(\sum_{\substack{i=1,j=1\\ i\neq j}}^{n} k_i k_j \Delta_{ip}\Delta_{jp}\Big) = 0 \tag{1.24}$$

故式(1.23)可写为

$$\frac{\sum_{i=1}^{N} \Delta z_i^2}{N} = k_1^2 \cdot \frac{\sum_{i=1}^{N} \Delta_{1i}^2}{N} + k_2^2 \cdot \frac{\sum_{i=1}^{N} \Delta_{2i}^2}{N} + \cdots + k_n^2 \cdot \frac{\sum_{i=1}^{N} \Delta_{ni}^2}{N} \tag{1.25}$$

由中误差定义式(1.14)，有

$$m_z^2 = k_1^2 \cdot m_{x_1}^2 + k_2^2 \cdot m_{x_2}^2 + \cdots + k_n^2 \cdot m_{x_n}^2 \tag{1.26}$$

即

$$m_z = \pm\sqrt{k_1^2 \cdot m_{x_1}^2 + k_2^2 \cdot m_{x_2}^2 + \cdots + k_n^2 \cdot m_{x_n}^2} \tag{1.27}$$

式(1.27)即为计算函数中误差的一般形式。

例 1.2 测一矩形截面试样的截面积，a 边的边长及其中误差分别为 $a=5\text{cm}$，$m_a=\pm 0.02\text{cm}$，b 边的边长及其中误差分别为 $b=2\text{cm}$，$m_b=\pm 0.01\text{cm}$，求矩形截面的面积 S 及其中误差 m_S。

解：(1)列函数式 $S=a \cdot b=5\times 2=10\text{cm}^2$

(2) 取偏导数

$$\frac{\partial S}{\partial a} = b = 2\text{cm}$$

$$\frac{\partial S}{\partial b} = a = 5\text{cm}$$

(3) 求中误差

$$m_S^2 = \left(\frac{\partial S}{\partial a}\right)^2 m_a^2 + \left(\frac{\partial S}{\partial b}\right)^2 m_b^2$$

$$=(2\times0.02)^2+(5\times0.01)^2$$
$$=0.0041$$

故 $m_S=\pm0.064\text{cm}^2$。

1.4.2 线性函数误差传播定律及其应用

对于一般线性函数而言，未知量 z 与独立直接观测量 $x_1,x_2,\cdots,x_n$ 的函数关系式为

$$z=k_1\cdot x_1+k_2\cdot x_2+\cdots+k_n\cdot x_n \tag{1.28}$$

式中：$k_1,k_2,\cdots,k_n$ 均为常数。

当独立直接观测量 $x_1,x_2,\cdots,x_n$ 的中误差为 $m_{x_1},m_{x_2},\cdots,m_{x_n}$ 时，则函数值 z 的中误差表达形式为

$$m_z=\pm\sqrt{k_1^2\cdot m_{x_1^2}+k_2^2\cdot m_{x_2^2}+\cdots+k_n^2 m_{x_n^2}} \tag{1.29}$$

例 1.3 对某试样在线弹性范围内进行应力应变测量，已知其弹性模量 $E=210\text{GPa}$，某一时刻在测量点处测得的应变值及其中误差分别为 $\varepsilon=2\times10^{-5}$，$m_\varepsilon=\pm5\times10^{-7}$，试求该时刻测量点处的应力 σ 及其中误差 m_σ。

解：(1)列函数式 $\sigma=E\cdot\varepsilon=210\times10^9\times2\times10^{-5}=4.2\text{MPa}$

(2) 求中误差

$$m_\sigma^2=E^2m_\varepsilon^2$$
$$=(210\times10^9)^2\times(5\times10^{-7})^2$$
$$=1.1025\times10^{10}$$

故 $m_\sigma=\pm0.105\text{MPa}$。

1.5 实验数据处理方法

在生产和科学实验中，一方面要研究被测量的最佳值和精度问题，另一方面要研究变量之间的内在关系。表达变量之间的内在关系的方法有很多种，如表格、数学表达式、散点图、曲线等，其中数学表达式更有利于从理论上做进一步分析研究，其形式紧凑，能客观地反映事物的内在规律性。对研究不同物理量之间的关系具有重要意义。

1.5.1 一元线性回归

为建立变量之间相互关系的数学模型，回归分析法是一种有效工具，如果两个随机变量 x 和 y 之间存在一定关系，并通过实验测量得到一系列 x 和 y 的数据，则可通过一元线性回归得出两个变量之间的关系式。

现对两随机变量进行一系列的测量，并得到一组样本数据(x_1,y_1)，(x_2,y_2)，$\cdots$，(x_n,y_n)。

为了正确反映 x 和 y 之间的关系，将样本数据在直角坐标系中描出相应点的位置，画出数据散点图，如图 1.2 所示，能够直观地反映出随机变量之间是否存在相关关系和函数形式。从图 1.2 中的点位分布情况看，所有数据点表现为线性分布趋势，因此，可以假定 x 和 y 之间存在线性相关关系。设两个变量 x 和 y 之间关系式的数学模型用一元线性方程表示为

$$y=f(x)=ax+b \tag{1.30}$$

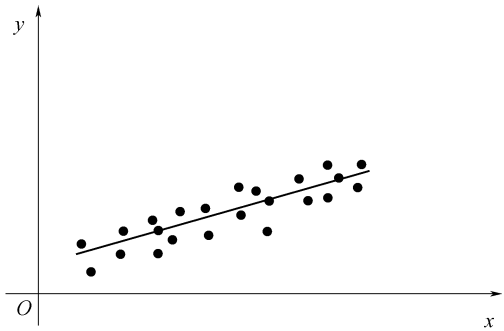

图 1.2　散点图

1. 一元线性回归方程的求解

一元线性回归就是根据一系列的测量数据，通过最小二乘法来求得式(1.30)中的系数 a 和 b。

根据最小二乘法原理，有

$$v = \sum_{i=1}^{n}[y_i - (ax_i + b)]^2 = \min \tag{1.31}$$

将式(1.31)分别对 a 和 b 求偏导数，得

$$\frac{\partial v}{\partial a} = -2\sum_{i=1}^{n}(y_i - ax_i - b)x_i$$

$$\frac{\partial v}{\partial b} = -2\sum_{i=1}^{n}(y_i - ax_i - b)$$

要使 v 为最小，则必要条件为

$$\frac{\partial v}{\partial a} = 0;\ \frac{\partial v}{\partial b} = 0$$

整理，得

$$\sum_{i=1}^{n}x_iy_i - a\sum_{i=1}^{n}x_i^2 - b\sum_{i=1}^{n}x_i = 0$$

$$\sum_{i=1}^{n}y_i - a\sum_{i=1}^{n}x_i - nb = 0$$

所以，得到系数 a 和 b 的解为

$$\begin{cases} a = \dfrac{n\sum\limits_{i=1}^{n}x_iy_i - \sum\limits_{i=1}^{n}x_i\sum\limits_{i=1}^{n}y_i}{n\sum\limits_{i=1}^{n}x_i^2 - \left(\sum\limits_{i=1}^{n}x_i\right)^2} \\ b = \dfrac{\sum\limits_{i=1}^{n}y_i\sum\limits_{i=1}^{n}x_i^2 - \sum\limits_{i=1}^{n}x_i\sum\limits_{i=1}^{n}x_iy_i}{n\sum\limits_{i=1}^{n}x_i^2 - \left(\sum\limits_{i=1}^{n}x_i\right)^2} \end{cases} \tag{1.32}$$

将得出的 a 和 b 代入式(1.30),就可以得出 x 和 y 的一元线性回归方程式。a 和 b 的值与样本数据的多少、实验数据的精度有关。

2. 相关系数的检验

随机变量 x 和 y 之间具有线性变化的趋势,求出的线性回归方程才具有重要意义,通常采用相关系数 ρ 来描述两个变量之间的线性关系密切程度。ρ 可以描述为

$$\rho = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{1.33}$$

式中:$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$;$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$。

相关系数 ρ 是描述两个变量 x 和 y 之间具有线性关系密切程度的数量指标,其取值范围为 $0 \leqslant |\rho| \leqslant 1$,$|\rho|$ 越接近于 1,x 和 y 之间的线性关系越密切。

表 1.2 是 $a=0.05$ 和 $a=0.01$ 两种显著水平相关系数检验表,表中的数值是相关系数的临界值,记为 ρ_a,如果相关系数在一定显著性水平下超过表中的数值,就认为 ρ 在某一水平上显著,即此时按线性回归处理才有意义。

表 1.2 相关系数显著性检验表

$n-2$ \ a	0.05	0.01	$n-2$ \ a	0.05	0.01
1	0.997	1.000	11	0.553	0.684
2	0.950	0.990	12	0.532	0.661
3	0.878	0.959	13	0.514	0.641
4	0.811	0.917	14	0.497	0.623
5	0.754	0.874	15	0.482	0.606
6	0.707	0.834	16	0.468	0.590
7	0.666	0.798	17	0.456	0.575
8	0.632	0.765	18	0.444	0.561
9	0.602	0.735	19	0.433	0.549
10	0.576	0.708	20	0.423	0.537

1.5.2 逐级加载法中的数据处理

在许多实验中,为了提高测量精度,经常采用等间距逐级加载法进行被测对象的测量工作,如在结构的应力分析中,对各点的应变测量工作是在线性范围内完成的,而且采用等间距逐级加载的方法来测量各级载荷的应变值,即对应的加载次数和应变值为

$$
\begin{array}{ll}
P_1 & \varepsilon_1 \\
P_2 = P_1 + \Delta P & \varepsilon_2 \\
P_3 = P_2 + \Delta P & \varepsilon_3 \\
\vdots & \vdots \\
P_n = P_{n-1} + \Delta P & \varepsilon_n
\end{array}
$$

求出相应的应变增量 $\Delta\varepsilon_i$ 和应变增量平均值 $\Delta\bar{\varepsilon}$，即

$$\Delta\varepsilon_i = \varepsilon_{i+1} - \varepsilon_i \quad i = 1,2,\cdots,n-1$$

$$\Delta\bar{\varepsilon} = \frac{\Delta\varepsilon_1 + \Delta\varepsilon_2 + \cdots + \Delta\varepsilon_{n-1}}{n-1}$$

通常将应变增量的算术平均值 $\Delta\bar{\varepsilon}$ 作为 ΔP 所对应的应变增量的最佳值。

1.5.3 测量数据的修约

测量是以确定测量值为目的的一组操作。测量值是由一个数(值)乘以测量单位所表示的特定量的大小。测量有间接测量和直接测量之分:直接测量的结果可直接测到而不必通过函数计算;而间接测量的结果需将直接测量的结果代入函数计算才能得到。对某一表示测量结果的数值(拟修约数),根据保留位数的要求,将多余的数字进行取舍,按照一定的规则,选取一个近似数(修约数)来代替原来的数,这一过程称为数值修约。有效数字是数据修约的基础。

1. 有效数字

有效数字是指在表达一个数量时,其中的每一个数字都是准确、可靠的,而只允许保留最后一位估计数字,这个数量的每一个数字为有效数字。一般情况下,在处理有效数字时,数字 0 需要区别看待。例如,用 0.02 精度的卡尺测量试样直径时,得到 10.06mm 和 10.20mm 两个数字,这里的 0 都是有效数字。当测量某一构件长度时得到 0.00530m 的数字,这里前面 3 个 0 都不是有效数字,它们只与所取的单位有关,而与测量的精度无关,当采用 mm 为单位时,则前面的 3 个 0 就不存在了,变为 5.30mm,其有效数字是 3 位。对于 32000m 和 25000Pa,很难肯定其中的 0 是否为有效数字,这种情况下采用指数的表示法。如 32000m 写为 3.2×10^4 m,则表示有效数字是 2 位;如果把它写为 3.20×10^4 m,则表示有效数字是 3 位。

2. 数值修约规则概述

测量结果及其不确定度同所有数据一样都只取有限位数,多余的位数应该按照相关规程进行修约。修约采用国家标准规定的数值修约规则。修约规则与修约间隔有关系。修约间隔又称修约区间或化整间隔,是确定修约保留位数的一种方式。根据金属拉伸实验方法标准 GB228—87,测量得到的力学性能数值可按照表 1.3 进行修约。

表 1.3　按照 GB228—87 标准得到的力学性能数值修约表

测量项目	范围	修约值
屈服极限 强度极限	≤200MPa	1MPa
	200～1000MPa	5MPa
	>1000MPa	10MPa
断后伸长率	≤10%	0.5%
	>10%	1%
断面收缩率	≤25%	0.5%
	>25%	1%

修约方法按照 GB8170—87 和 GB3101—93 执行。如果应力为 200～1000MPa，应力计算的尾数小于 2.5，则舍去；计算的尾数大于或等于 2.5 且小于 7.5，则取 5；计算的尾数大于或等于 7.5，则取 10。