近似計算初步

JINSI JISUAN CHUBU

俞咸宜編

1961年12月

目 錄

目　錄

第一章　近似計算的准確度

§1. 引　　言

我們已經看到，数学分析是研究科学技术問題的有力工具，因此在科学技术問題的研究中，就会碰到一些数和各式各样的数学公式，有关工程人員就必須能使用这些数，并会用公式求出所需要的数值，然后才能依据这些数来进行設計，制造，配量，檢驗等等具体工作，但是人們認識到而确定下来的数量都是近似的，例如雷达測得敌艦与我艦距离为21.5浬，实际上眞正的距离可能比这个数略大或略小。另一方面，解决問題时所使用的公式也未見得是准确公式，参与計算的某些常数也只可能用有限位的数近似表达，例如π就是这样的，因此在实际計算中以及最后所得的結果一般講来总是近似的。但是人們总可以想出各种的办法来減少計算过程的誤差来使計算的結果达到需要的准确度，近似計算研究的內容就是寻求近似公式和計算方法，使花費較少的劳动力和时間，来获得有一定准确度的結果。

§2. 誤差及其来源

若数a'是数a的近似值，那么就产生誤差$|a-a'|$，近似与誤差这两个概念是同时产生的矛盾的两方面，計算工作者想尽办法来促使这两方面的轉化，首先必需了解近似及誤差的特征及內在联系。

誤差的来源。

A. 数学描述和实际現象間的誤差：例如自由落体的运动方程是 $S=\frac{1}{2}gt^2$，但是实际現象的規律并非如此簡单，无論如何空气阻力总是要起作用的，不过我們忽略不計了，所以这个用公式描述自由落体的运动規律是有誤差的。

B. 观測誤差：当用仪表測量一个量时，所讀出的数只可能是近似的，因此也产生了誤差，这叫观測誤差，減少观測誤差应由改进仪表精密度的技术上来想办法。

C. 截断誤差：人們仅能进行四則运算，至于开方事实上也是通过四則运算来进行，数字計算机严格地說来只能进行加法运算，減法，乘法除法都是用加法来完成，而且不論是人或机器（机器是人造的）都只能作有限次的运算，当計算公式中包含超越函数时，就必須化为有限次四則运算来計算，例如取$\sin x$的代劳級数展式的有限項来計算 $\sin x$，

$$\sin x \approx x-\frac{x^3}{3!}+\frac{x^5}{5!}-+\cdots+(-1)^m\frac{x^{2m+1}}{(2m+1)!}.$$

这样就产生了截断誤差

$$|R_n(x)|=\frac{|\sin\theta x|}{(2m+3)!}|x|^{2m+3}.$$

D. 捨入誤差：参与計算的数常常不能用有限位数来表示，例如π就只能以有限位近似値3.1416或3.14159来参与計算，换句話說，把数的尾数去掉了，为了使計算过程中儘量減少产生誤差的机会，現在对四捨五入法作以下补充規定。

(1) 若捨去的尾数的第一位小于5，則将尾数抹去。

(2) 若捨去的尾数的第一位大于5，則将尾数抹去并在尾数前一位加上1。

(3) 若尾数的第一位恰巧是5，那么前一位是偶数，就将尾数抹去，若前一位是奇数，就将尾数抹去并在前一位上加1。总之讓数的最末位是偶数。

第(3)条規定的理由如下，因为数字出現奇偶数的机会是均等的，因此应讓捨和入的机会均等，这样可以在累次捨入时抵消誤差，又因为末位为偶数比末位为奇数容易被其它数除尽，因此可以減少一些捨入的机会。

关于誤差的积累：

在計算过程中，每一步所产生的誤差（不管是那一种誤差）都会遺留到下一步及以后所有的計算中去而产生积累誤差，当然前后誤差可能有机会互相抵消，但这很难估計，捨入誤差带有随机性，所以更难估計其积累，因此在实际計算中，多採用事后的实驗估計，即重复演算，逐次精密化，直到所要求的数位稳定为止，至于如何逐次精密化将在以下各章中根据具体計算方法来討論。

§ 3. 絶对誤差与相对誤差

若 x 是某量的眞值，$\bar{x}$ 是它的近似值，称

$$E_a=|x-\bar{x}|$$

为 $\bar{x}$ 的絶对誤差。

一般不能知道眞值，只能估計出絶对誤差不致超过某一正数，即估出最大絶对誤差，以下提到絶对誤差就是指最大絶对誤差，例如説經緯仪所測的角度准确度为5″，这就是說，測得的角度的誤差不致超过5″。

絶对誤差并不能充分表达度量的精确程度，例如若測得敌艦与我艦的距离为20浬或50浬时有相同絶对誤差，显然对后者来說就精确得多，因此度量的精确度必須与所量得的数相比較而得，称

$$E_r=\frac{|x-\bar{x}|}{|\bar{x}|}$$

为 $\bar{x}$ 的相对誤差，与絶对誤差一样，事实上只能估出一个最大相对誤差，以后簡称所估的最大相对誤差为相对誤差，显然絶对誤差与相对誤差有关系

$$E_a=|\bar{x}|\,E_r.$$

四則运算的誤差：

Ⅰ. 几个近似数的和及差的絶对誤差是它們的絶对誤差的和。

証 依两个数的情况来証明，

若$\bar{x}_1 \approx x_1$，$\bar{x}_2 \approx x_2$，其絶对誤差各为E_{a1}，E_{a2}，則

$$|(x_1 \pm x_2) - (\bar{x}_1 \pm \bar{x}_2)| = |(x_1 - \bar{x}_1) \pm (x_2 - \bar{x}_2)|$$
$$\leqslant |x_1 - \bar{x}_1| + |x_2 - \bar{x}_2| \leqslant E_{a1} + E_{a2}$$

这就証明了結論。

Ⅱ. 几个近似数（正数）的和的相对誤差是这些数的相对誤差中最大的那一个。

証 若$\bar{x}_1 \approx x_1$，$\bar{x}_2 \approx x_2$，相对誤差各为E_{r1}，E_{r2}其中E_{r1}較大，若記$\bar{x}_1 + \bar{x}_2$的相对誤差为E_r，則

$$E_r = \frac{|x_1 + x_2 - (\bar{x}_1 + \bar{x}_2)|}{|\bar{x}_1 + \bar{x}_2|} = \frac{|(x_1 - \bar{x}_1) + (x_2 - \bar{x}_2)|}{|\bar{x}_1 + \bar{x}_2|}$$
$$\leqslant \frac{\bar{x}_1}{\bar{x}_1 + \bar{x}_2} \frac{|x_1 - \bar{x}_1|}{\bar{x}_1} + \frac{\bar{x}_2}{\bar{x}_1 + \bar{x}_2} \frac{|x_2 - \bar{x}_2|}{\bar{x}_2}$$
$$\leqslant E_{r1} \frac{\bar{x}_1 + \bar{x}_2}{\bar{x}_1 + \bar{x}_2} = E_{r1}.$$

如果两个数是負数，可就絶对值来討論相对誤差。

Ⅲ. 相差悬殊的两个数（不妨設为正数）相減时，其中大数的相对誤差起决定性作用。

若$x_1 \gg x_2$，$\bar{x}_1 \approx x_1$，$\bar{x}_2 \approx x_2$，当然$\bar{x}_1 \gg \bar{x}_2$，因此

$$\frac{\bar{x}_1}{\bar{x}_1 - \bar{x}_2} \approx 1, \qquad \frac{\bar{x}_2}{\bar{x}_1 - \bar{x}_2} \approx 0$$

故 $\bar{x}_1 - \bar{x}_2$ 的相对誤差

$$E_r = \left| \frac{(x_1 - x_2) - (\bar{x}_1 - \bar{x}_2)}{\bar{x}_1 - \bar{x}_2} \right| = \left| \frac{(x_1 - \bar{x}_1) - (x_2 - \bar{x}_2)}{\bar{x}_1 - \bar{x}_2} \right|$$
$$= \left| \frac{x_1 - \bar{x}_1}{\bar{x}_1} \frac{\bar{x}_1}{\bar{x}_1 - \bar{x}_2} - \frac{x_2 - \bar{x}_2}{\bar{x}_2} \frac{\bar{x}_2}{\bar{x}_1 - \bar{x}_2} \right| \leqslant \left| \frac{x_1 - \bar{x}_1}{\bar{x}_1} \right|$$
$$= E_{r1}.$$

Ⅳ. 乘积的相对誤差約等于各因子相对誤差的和。

証 若$x_1 = \bar{x}_1 + \varepsilon_1$，$x_2 = \bar{x}_2 + \varepsilon_2$，当然$\bar{x}_1$，$\bar{x}_2$的相对誤差

$$E_{r1}=\left|\frac{\varepsilon_1}{\bar{x}_1}\right|,\qquad E_{r2}=\left|\frac{\varepsilon_2}{\bar{x}_2}\right|.$$

所以$\bar{x}_1\bar{x}_2$的相对誤差

$$E_r=\left|\frac{x_1x_2-\bar{x}_1\bar{x}_2}{x_1x_2}\right|=\left|\frac{(\bar{x}_1+\varepsilon_1)(\bar{x}_2+\varepsilon_2)-\bar{x}_1\bar{x}_2}{x_1x_2}\right|$$

$$=\left|\frac{\bar{x}_1\varepsilon_2+\bar{x}_2\varepsilon_1+\varepsilon_1\varepsilon_2}{x_1x_2}\right|\approx\left|\frac{\varepsilon_2}{\bar{x}_2}+\frac{\varepsilon_1}{\bar{x}_1}\right|\leqslant E_{r1}+E_{r2}.$$

Ⅴ. 商的相对誤差約等于相除二数相对誤差約和。

証 不妨設二数是正数$x=\bar{x}+d\bar{x}$，$y=\bar{y}+d\bar{y}$。若記$\bar{u}=\dfrac{\bar{x}}{\bar{y}}$，則由全微分知識可知

$$d\bar{u}=\frac{1}{\bar{y}}d\bar{x}-\frac{\bar{x}}{\bar{y}^2}d\bar{y}.$$

$\dfrac{\bar{x}}{\bar{y}}$的相对誤差

$$E_r=\left|\frac{d\bar{u}}{\bar{u}}\right|=\left|\frac{\frac{1}{\bar{y}}d\bar{x}-\frac{\bar{x}}{\bar{y}^2}d\bar{y}}{\frac{\bar{x}}{\bar{y}}}\right|=\left|\frac{d\bar{x}}{\bar{x}}-\frac{1}{\bar{y}}d\bar{y}\right|$$

$$\leqslant\left|\frac{d\bar{x}}{\bar{x}}\right|+\left|\frac{d\bar{y}}{\bar{y}}\right|=E_{rx}+E_{ry}.$$

其中E_{rx}，E_{ry}分别为$\bar{x}$及$\bar{y}$的相对誤差。

Ⅵ. 誤差的一般公式。

对函数

$$u=f(x_1,x_2,\cdots\cdots,x_n)$$

自变量$x_1,x_2,\cdots\cdots,x_n$的誤差将引起函数值u的誤差，若$\bar{x}_i\approx x_i$且$\bar{x}_i=x_i+dx_i$，$(i=1,2,3,\cdots\cdots,n)$，則$\bar{x}_i$的絶对誤差就是$|dx_i|$，分别記为$E_{a1},E_{a2},\cdots\cdots,E_{an}$。由全微分知識可知$\bar{u}=$ $=f(\bar{x}_1,\bar{x}_2,\cdots\cdots,\bar{x}_n)$的絶对誤差

$$E_a=\left|\frac{\partial f}{\partial x_1}dx_1+\frac{\partial f}{\partial x_2}dx_2+\cdots\cdots+\frac{\partial f}{\partial x_n}dx_n\right|$$

$$\leqslant\left|\frac{\partial f}{\partial x_1}\right|E_{a1}+\left|\frac{\partial f}{\partial x_2}\right|E_{a2}+\cdots\cdots+\left|\frac{\partial f}{\partial x_n}\right|E_{an}.$$

而 $\bar{u}$的相对誤差

$$E_r\leqslant\left|\frac{\partial f}{\partial x_1}\right|\frac{E_{a1}}{|\bar{u}|}+\left|\frac{\partial f}{\partial x_2}\right|\frac{E_{a2}}{|\bar{u}|}+\cdots\cdots+$$

$$+\left|\frac{\partial f}{\partial x_u}\right|\frac{E_{a2}}{|\bar{u}|}.$$

举 例

例1 求349.1×863.4的积，若已知这两个数都只准确到第四位数字；估計积的絶对誤差；并說明結果有几位数字可靠。

解 設 $\bar{x}=349.1$，$\bar{y}=863.4$。由于二数只准确到最末一位数字，則 $\bar{x}$，$\bar{y}$的絶对誤差分别为

$$\Delta x\leqslant 0.05,\qquad \Delta y\leqslant 0.05$$

因此积的相对誤差

$$E_r=\frac{0.05}{349.1}+\frac{0.05}{863.4}=0.000143+0.000057=0.00020.$$

若积取六位数字則

$$\bar{x}\bar{y}=349.1\times 863.4\approx 301413$$

积的絶对誤差

$$E_a=301413\times 0.00020=60$$

因此眞实結果在301473与301353之間，通常取这二数的平均值3014413作积的近似数，但第四位数字开始就是不可靠的。

例2 求$\frac{56.3}{\sqrt{2}}$的准确数字的位数，已知分子准确到末位数。

解 $\bar{x}=56.3$ 的絶对誤差是0.05，相对誤差是$\frac{0.05}{56.3}\leqslant 0.0009$.

今取$y=\sqrt{2}\approx 1.41421=\bar{y}$，使 $\bar{y}$ 的相对誤差与 $\bar{x}$ 的相对誤差比

起来可以略而不計，于是$\frac{56.3}{\sqrt{2}}$的相对誤差

$$E_r = 0.0009.$$

作除法得：

$$\frac{x}{y} = \frac{56.3}{1.41421} \approx 39.6,$$

只取小数点后一位，是因为分子已經只准确到这一位，不可能要求結果比原給数更精确些，故将商多取位数也是徒劳无益的，商39.6的絶对誤差

$$E_a \leqslant 39.6 \times 0.0009 < 0.036 < 0.05$$

由此可知，所得結果不影响小数点后第一位，39.6三位数字都是准确値。

§4. 有效数字及其与誤差的关系

因为作近似計算时，要估計誤差，所以在書写数时，应該把各近似数加以区别，就是說应該把各近似数的准确程度表示出来，因此建立有效数字这一概念。

有效数字

有效数字是指1，2，3，……，9中任何一个数字，至于0有时是有效数字，有时只起填空的作用，例如0.00263中有三位有效数字2，6，3，前面两个0是填空的，但3809中四位都是有效数字；至于1.300中除1，3均为有效数字外，最后两个0是否有效，那就要看这个数的絶对誤差是否不大于0.0005，也就是說，如果最后一位0是經过捨入后确定下来的，那它就是有效数字，否則它是被拿来填空的，对于二进位数来說，有效数字有1这个数，至于0是否有效，也要由絶对誤差来审核，当然如果数的首位和末位都是1的話，中間各位出現的0都是有效数字。

为了把数的近似程度写明白，在書写时应保証由左起第一个不为零的数开始到最后一位数都是有效数字。

有效数字說明了数的准确程度，工程技术部門提出的計算問題，对准确度的要求有时用有效数位来說明，有时用絕对誤差相对誤差来說明，所以計算工作者应該对精确度要求的两种不同表示法加以換算，事实上有效数字和誤差有下列相互关系。

定理1 若 $\bar{x}$的首位有效数字为x_1，且該数准确到n位有效数字，則 $\bar{x}$的相对誤差就不大于$\frac{1}{2x_1}\cdot\frac{1}{10^{n-1}}$.

証 数 $\bar{x}$总可写成

$$\bar{x}=\pm(x_1 10^m+x_2 10^{m-1}+\cdots\cdots+x_n 10^{m-n+1}),$$

其中m为整数，$x_i(i=1,2\cdots\cdots n)$是 0，1，2，…9中的一个数，那么

$$x_1 10^m\leqslant|\bar{x}|\leqslant(x_1+1)10^m,$$

又因为

$$|x-\bar{x}|\leqslant\frac{1}{2}\times10^{m-n+1}$$

所以 $\bar{x}$的相对誤差为

$$E_r=\frac{|x-\bar{x}|}{|\bar{x}|}\leqslant\frac{\frac{1}{2}10^{m-n+1}}{x_1 10^m}=\frac{1}{2x_1}\cdot\frac{1}{10^{n-1}}.$$

例如近似数864.32 的相对誤差就不超过$\frac{1}{2\times8}\cdot\frac{1}{10^4}$.

注意定理1的逆是不成立的，例如用0.484近似表示0.4900，($x_1=4$)有相对誤差

$$E_r=\frac{0.006}{0.484}=0.01239<0.0125<\frac{1}{2\times4}\cdot\frac{1}{10^{2-1}}$$

但事实上0.484对0.4900来講并沒有取2位有效数字。

定理2 如果近似数的相对誤差小于$\frac{1}{2(x_1+1)}\cdot\frac{1}{10^{n-1}}$，那么这个数便准确到$n$位有效数字，其中$x_1$是第一位有效数字。

証 由于

$$E_r = \frac{|x-\bar{x}|}{|\bar{x}|} \leqslant \frac{1}{2(x_1+1)} \cdot \frac{1}{10^{n-1}},$$

$$|\bar{x}| \leqslant (x_1+1)10^m$$

故得

$$E_a = |x-\bar{x}| \leqslant \frac{(x_1+1)10^m}{2(x_1+1)10^{n-1}} = \frac{1}{2}10^{m-n+1}$$

上式就說明了絕对誤差不大于近似数最末位的单位的一半，因此第 n 位是有效数字。

§ 5. 用复杂算式求值时的准确度估計

对函数作近似計算时，总是遇到下面两个基本問題。

(a) 給出各个自变量或近似数的誤差，要求出所得函数值的誤差。

(b) 要使求得函数值的誤差达到规定的准确度，求出各个自变量所容許的誤差。

由§3Ⅵ的誤差的一般公式来解决这两个問題，举例如下，

例3 由眞空彈道公式已知若炮彈以初速 v_0，发射仰角为 φ 时，可命中距炮位为:

$$x = \frac{1}{g} v_0^2 \sin 2\varphi$$

处的目的物，其中 g 为重力加速度，今以 $v_0 = 600$米/秒，$\varphi = 60°$ 发射，若 v_0 有絕对誤差 $|dv_0| = 0.1$米/秒，φ 有絕对誤差 $|d\varphi| = 1'$。問炮彈落地偏离多少米?

解

$$|dx| = \left| \frac{1}{g} v_0^2 2\cos 2\varphi d\varphi + \frac{1}{g} 2v_0 \sin 2\varphi dv_0 \right|$$

$$\leqslant \frac{2}{g} \{ v_0^2 |\cos 2\varphi| \, |d\varphi| + v_0 |\sin 2\varphi| \, |dv_0| \}$$

将 $\varphi=60°$，$v_0=600$米/秒及$|dv_0|=0.1$米/秒，$|d\varphi|=\frac{\pi}{180}\times\frac{1}{6}$

$=\frac{\pi}{10800}\approx0.000292$代入上式即得所求偏离

$$E_a=|dx|\leqslant\frac{2}{9.8}\left\{300^2\times\frac{1}{2}\times0.000292+600\times\frac{\sqrt{3}}{2}\times0.1\right\}$$

$$\approx\frac{1}{4.9}\{52.56+51.93\}\approx21.30\ (\text{米})$$

以上計算沒有估計捨入誤差。

例2 求函数 $6x^2(\lg x-\sin2y)$ 的值，須准确到两位小数。如果 x，y 的近似值各为15.2及57°，試求这些量能允許的誤差。

解 令 $z(x,y)=6x^2(\lg x-\sin2y)$，則

$z(15.2,57°)=6\times(15.2)^2(\lg15.2-\sin114°)=371.9$，

$\frac{\partial z}{\partial y}=12\times15.2(\lg15.2-\sin114°+6\times15.2\times0.4329)=88.54$，

$\frac{\partial z}{\partial y}=-12(15.2)^2\cos114°=1127.7$；

因此

$$dz=\frac{\partial z}{\partial x}dx+\frac{\partial z}{\partial y}\partial y=88.54dx+1127.7dy$$

若要使值 z 准确到小数点两位，就要求 $dz<0.005$，那么就要求

$$88.54dx<\frac{0.005}{2},\qquad 1127.7dy<\frac{0.005}{2}$$

故必須

$$dx<\frac{0.005}{2\times88.54}=0.000028\approx0.000003$$

$$dy<\frac{0.005}{2\times1127.7}=0.0000022\text{弧度}\approx0''.45.$$

由結果看出，若要求所得 z 值有两位小数准确，取 x，y 值时就

必須是有7位有效数字（整数二位小数五位）的数。

§6. 作近似計算的几項注意

一、两个或多个准确度不同的数相加时，应該将准确度較高的数先进行捨入，使比最不准确的那个数还多一位有效数字，然后再进行相加，全部保留各数有效数字，是徒劳无益的。

例如求

$$561.32+491.6+86.954+3.9462$$

第二数只准确到小数点一位，故将其余数捨入至保留小数点两位，即得

$$491.6+561.32+86.95+3.95=1143.8.$$

絶对誤差为0.05。

二、有效数字在減法中的丧失。

当两个相近的数相減时就会失去有效数字，例如两个有五位有效数字的数16950，16870相減后得差80，剩下两位有效数字，失去有效数字，在某些計算中将会引起极大誤差，例如

$$(1.2526-1.2448)\times 9800=76.44$$

但若对相減二数仅取三位有效数字，則有

$$(1.25-1.24)\times 9800=98$$

因此由失去重要有效数字引起了絶对誤差21.56，几乎有25%的相对誤差。

避免由失去必要的有效数字而产生的不准确性，可以經过事先估計，增加減数及被減数的有效数字位数。或者将計算公式变換一下，例如求$1-\cos 4°$可以換成

$$1-\cos 4°=2\sin^2 2°$$

来計算。

三、計算次序引起的誤差。

計算时应注意安排計算次序，否則也会引起极大的誤差，例如計算

$$D=\frac{0.0005\times 0.0143\times 0.0012}{0.0003\times 0.0125\times 0.0135},$$

若将分子分母分别計算，均取九位小数，则有

$$D\approx\frac{0.000000009}{0.000000051}\approx 0.17647$$

但眞值应为 0.16948……，因此只准确到小数点一位，如果分成三个因子分别計算，然后相乘，即得

$$D=\frac{0.0005}{0.0003}\cdot\frac{0.0143}{0.0125}\cdot\frac{0.0012}{0.0135}\approx 0.16948.$$

准确到小数点五位。

又例如

$$x=0.000030\times 0.000050\times 98000$$
$$=0.0000000015\times 98000=0.000147$$

如果一个計算机只有八位，那么按上面的次序計算时，就得結果为零，因此应将計算次序安排一下，如果計算

$$y=\frac{98000}{0.000030\times 0.000050}$$

则更要注意，因为分母在八位計算机上表現为零，用零作除数将会使計算机运轉发生故障。

习　题　一

1. 下列各数准确到末位数字，試求其和，并估出和的絕对誤差。136.421；28.3；321；68.243；17.482.

2. 数 48.932 及 6852.4 都是准确到末位的近似数，試求其差，并說出其結果有多少位可靠。

3. 求 $\sqrt{10}-\pi$ 的值准确到五位有效数字。

4. 由公式

$$R=\frac{r^2}{2h}+\frac{h}{2}$$