工科研究生教材

数值分析

SHUZHI FENXI

沐定夷　胡鸿钊　编著

上海交通大学出版社

数值分析

沐定夷　胡鸿钊等编

上海交通大学出版社

（沪）新登字 205 号

内 容 提 要

本书共分八章，内容包括：误差，解线性方程组的直接法，解线性方程组的迭代法，非线性方程的数值解法，矩阵的特征值与特征向量的计算，插值与逼近，数值积分，微分方程的数值解法。

本书论述清晰流畅，深入浅出，分析问题运用启发式。介绍方法与阐明原理并重，传授知识与培养能力兼顾。

本书供工科与邻近学科各专业硕士生选用，也可作工程技术人员与自学者参考用书或复习应考用书。

责任编辑：冯　愈
封面设计：陈玉兰

数 值 分 析

出版：上海交通大学出版社
（上海市华山路1954号　邮政编码：200030）

发行：新华书店上海发行所　　印刷：上海交通大学印刷厂
开本：850×1168（毫米）1/32　　印张：8.875 字数：229000
版次：1994年9月　第1版　　印次：1994年9月　第1次
印数：1-1000　　科目：328－296

ISBN 7-313-01344-2/O·241　　定　价：5.90元

序

计算数学源远流长，是数学的同龄人。

纵观计算数学的历史可见其特点之一是：计算方法与计算工具之间存在着密切的联系，从数数、作加法需用手指开始，伴随着四则运算、代数运算、超越运算等的发展出现了算筹、算盘、算尺、计算器等工具。

只有方法与工具能紧密配合，我们才如虎添翼、如鱼得水，更有效地把计算数字应用于科学技术以及日常生活。当年数值天气预报的方法被找到时，由于计算工具落后，预报明天的天气，远远不止一两天才能完成，又若只有Cramer公式而没有Gauss消去法，则即使用最先进的电算机来解1000阶的线性方程组，所需的时间也大大超过地球的寿命。这些都是方法与工具不匹配的典型例子。

在计算工具的发展史上，值得大书特书的是40年代由v.Neumann领衔设计的电算机的出现，这件事被誉为科技史上的第四个里程碑，它为计算方法的研究带来了一个黄金时期，目前常用的一些卓越的算法都是在这个时期获得的。现在电算机技术已渗透到人类活动的各个领域，也是世所公认的科学研究的第三种手段。

特别是70年代兴起的并行机，冲破了v.Neumann局限于串行的构想，开创了电算机技术的新局面。发展至今并行机的浮点运算速度已达百亿、千亿次/秒。适应这种算机的算法是崭新的并行算法，它尚处于方兴未艾阶段。据美、日等计算机大国报道，并行机通常实际运行速度离机器峰值有很大距离，这为计算方法的研究注入了新的生命力。

“数值分析”是工科硕士生的基础课程，其重要性从近年电算机技术在经济建设、科学技术等各方面日益广泛深入的应用可见一斑。

本书在介绍方法、阐明原理的同时，希望能培养学生处理问题解决问题的能力，因此在编写中，注意运用启发式，讲清来龙去脉，交待思考方法，说明理论源于实际又用于实际，对推导作了某些简化以突出本质、便于接受，也不予证明地介绍一些常用的结论。

本书的前身是1987年由胡鸿钊、吴紫电、倪弘杰与我合作撰写的讲义，经讲授积累了经验、发现了问题后于1989年由胡鸿钊与我作了修订（吴紫电、倪弘杰已出国）。

这次成书，内容有了较大的增删，主要依据是国家教委工科研究生数学课程指导委员会制定的工科硕士生数值分析教学基本要求。由胡鸿钊加写了偏微分方程数值解、吴忠英与杨升荣参加了习题与例题的准备、我对全书作了大修改，这次改写加入了在教学实践中考虑到的一些处理方式，也吸收了国内外同类教材的优点。

徐桂芳老教授拨冗审阅了书稿并提出宝贵意见，冯愈小姐对本书做了出色的编辑，颜钰芬副教授对我的工作始终给予配合与支持，对他们敬表谢忱。

限于见闻、水平，书中舛误尚祈读者指正。

沐定夷

1993.12于上海交通大学

目　录

第一章 误　　差

第一节 误差的来源

用数值分析的方法求一个实际问题的数值解往往会产生误差，这些误差主要是：

（1）模型误差：从实际问题归结为数学问题时，通常要略去某些次要的因素，或再加以简化，由此而产生的误差称之为模型误差。

（2）观测误差：在归结所得的数学问题中常含有一些数据需要观测。当用各种工具观察、测量这些数据时，人的观测与测量工具的使用都会给观测数据带来误差，这就是观测误差。

（3）方法误差：当对数学问题求精确解有困难时，常用数值方法求其近似解，因此引起的误差称之为方法误差。

（4）舍入误差：由于计算机只能将数表示成有限位进行运算，所以对超过位数的数字要按一定的规则作舍入，由此产生的误差称之为舍入误差。

对模型误差的分析属于应用数学的范围，观测误差也不在数值分析中讨论，以后将只讨论方法误差与舍入误差。

例 1.1 研究一个单摆的摆角 θ 的变化规律（如图1-1）。为了归结为数学问题，先略去空气阻力、支点的摩擦力、摆线的质量等，仅考虑摆锤所受的重力作用，由 Newton 第二定律得出

$$ml\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2}=-mg\sin\theta,$$

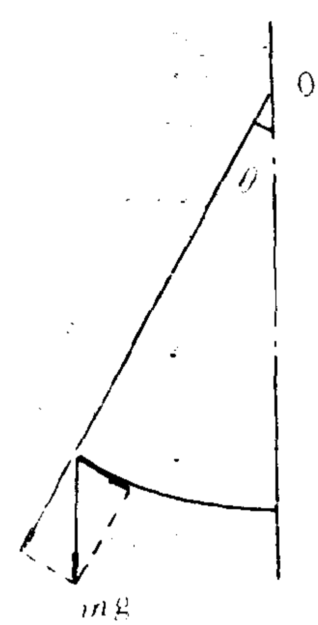

图 1-1

其中m是摆锤的质量，l是摆线的长度。

上式是比较复杂的非线性微分方程，在$|\theta|$较小时，$\sin\theta\approx\theta$，于是可以进一步把它简化为

$$\frac{d^2\theta}{dt^2}=-\omega^2\theta,\quad \omega^2=\frac{g}{l}。\qquad (1.1)$$

经过上述略去次要因素与简化得出的式（1.1）已含有模型误差。

解式(1.1)得

$$\theta(t)=A\sin\omega t+B\cos\omega t,$$

由此得出单摆的周期$T=\frac{2\pi}{\omega}=2\pi\sqrt{\frac{l}{g}}$。

在实际计算T时，需要对l与g作观察与测量，这时不可避免会带来观测误差；当使用各种方法求$2\pi\sqrt{\frac{l}{g}}$时，所用的近似方法求得的解与$2\pi\sqrt{\frac{l}{g}}$之间还有方法误差，最后，按算式所作运算都用有限位数进行，于是又产生了舍入误差，利用初始条件确定A、B时也有类似的误差。

第二节　绝对误差　相对误差　有效数字

一、绝对误差　相对误差

设x为某一准确数，x^*为x的近似数，则称

$$e=x-x^*$$

为x^*对x的绝对误差。

如$x=\sqrt{2}$，$x^*=1.4142$，则$e=\sqrt{2}-1.4142=0.00001356\cdots$。

一般地说 e 是未知的，往往只能根据实际情况得出 $|e|$ 的一个上界 ε：$|e|\leqslant\varepsilon$，并把 ε 称为 x^* 对 x 的绝对误差限。有了 ε，可由 x^* 与 ε 得出 x 的范围：

$$x^*-\varepsilon\leqslant x\leqslant x^*+\varepsilon。$$

在工程技术上将上述不等式表示成

$$x=x^*\pm\varepsilon。$$

如 $l=10\pm0.01(\mathrm{cm})$，即 $l^*=10$ 是 l 的一个近似值，其绝对误差限 $\varepsilon=0.01$，$|e|\leqslant0.01$。

例 1.2　考虑“四舍五入”的绝对误差限。

设有实数 x，它的十进制标准表示式是

$$x=\pm0.a_1a_2\cdots a_na_{n+1}\cdots\times10^m,$$

其中 m 为整数，$a_i\in\{0,1,2,\cdots,9\}$，$i=1,2,\cdots$，$a_1\neq0$。若对 a_{n+1} 作四舍五入，得 x 的近似值 x^*：

$$x^*=\begin{cases}\pm0.a_1a_2\cdots a_n\times10^m, & \text{当 } a_{n+1}\leqslant4 \text{ 时（四舍）},\\ \pm0.a_1a_2\cdots(a_n+1)\times10^m, & \text{当 } a_{n+1}\geqslant5 \text{ 时（五入）},\end{cases}$$

则四舍时

$$\begin{aligned}|e|=|x-x^*|&=(0.a_1a_2\cdots a_na_{n+1}\cdots-0.a_1a_2\cdots a_n)\times10^m\\&\leqslant0.\underbrace{00\cdots0}_{n}5\times10^m=\frac{1}{2}\times10^{m-n};\end{aligned}$$

五入时

$$\begin{aligned}|e|=|x-x^*|&=(0.a_1a_2\cdots(a_n+1)-0.a_1a_2\cdots a_na_{n+1}\cdots)\times10^m\\&\leqslant0.\underbrace{00\cdots0}_{n}5\times10^m=\frac{1}{2}\times10^{m-n}。\end{aligned}$$

即对 x 作四舍五入后，其绝对误差限为被保留的最后数位上的半个单位。

有时候，单用绝对误差的概念还不足以完整地反映误差的情况，如测量长为 100(m)的甲物时，有绝对误差 1(cm)，测量长为 1(m)的乙物时，绝对误差也是 1(cm)，虽然两个测量的绝对误差相同，显然测量甲物比测量乙物精确。

由此可见，要全面地反映近似值的精度，还需要考虑绝对误差与所讨论的数量的比。

把近似数 x^* 的绝对误差 e 与准确数 x 之比

$$\frac{e}{x}=\frac{x-x^*}{x}$$

称为 x^* 对 x 的相对误差。实际上，x 是未知的，所以通常改用

$$\frac{e}{x^*}=\frac{x-x^*}{x^*}$$

作为 x^* 的**相对误差**，并记作 e_r：

$$e_r=\frac{e}{x^*}。$$

当 e_r 很小时，$\frac{e}{x^*}$与$\frac{e}{x}$的差

$$\frac{e}{x^*}-\frac{e}{x}=\frac{e^2}{x^*x}=\frac{(e/x^*)^2}{1+e/x^*}\approx e_r^2$$

可以忽略不计。

$|e_r|$的上界称之为 x^* 对 x 的**相对误差限**。记作 ε_r：

$$|e_r|\leqslant\varepsilon_r,$$

显然$\frac{\varepsilon}{|x^*|}$是$x^*$的一个相对误差限。

如国际大地测量会议建议光速

$$c=299792458\pm1.2(\mathrm{m/s}),$$

其绝对误差限

$$\varepsilon=1.2(\mathrm{m/s}),$$

相对误差限

$$\frac{\varepsilon}{|c^*|}=\frac{1.2}{299792458}\leqslant0.0000000041。$$

由此可见，$c^*=299792458(\mathrm{m/s})$的相对误差是很小的。

当 x 是有单位的量值时，其近似值 x^* 对 x 的绝对误差、绝

对误差限有与 x 相同的量纲；x^*对 x 的相对误差、相对误差限是无量纲的。

二、有效数字

设 x 的近似数

$$x^* = \pm 0.a_1a_2\cdots a_n \times 10^m,$$

如果

$$|x-x^*| \leqslant \frac{1}{2}\times 10^{m-n}, \tag{1.2}$$

那末称 x^* 作为 x 的近似数具有 n 位有效数字。

例如，在用某种计算方法求 x 时，所得的 x^* 往往具有若干位有效数字；又如用四舍五入取准确数x的前 n 位作近似数 x^*，则 x^* 具有 n 位有效数字（见例1.2）。在计算机中表示与处理数时，都是有限位的，因此用（二进制下的）舍入法产生 x 的近似数x^*是常用的产生有效数字的方法。

例 1.3 对某地重力常数 g（以 m/s^2 为单位）作四舍五入得 $g\approx g^*=9.81(m/s^2)$，

$$|g-g^*| \leqslant \frac{1}{2}\times 10^{-2} = \frac{1}{2}\times 10^{1-3},$$

由 $m=1$ 知它有 3 位有效数字。若以 km/s^2 为单位，$g\approx g^*=0.00981(km/s^2)$：

$$|g-g^*| \leqslant \frac{1}{2}\times 10^{-5} = \frac{1}{2}\times 10^{-2-3},$$

由 $m=-2$ 知它也有 3 位有效数字。

由此可见，有效位数与近似数采用的单位无关。

如上所述，绝对误差 $x-x^*$ 与相对误差 $\dfrac{x-x^*}{x^*}$ 是衡量x^*近似 x 时近似程度好坏的标志；x 与它的有 n 位有效数字的近似数 x^* 之间满足式(1.2)，于是自然猜测误差与有效数字这两个概念

之间有一定的联系。事实上，有

定理 1.1 （有效数字与相对误差的关系）

(1)若
$$x^*=\pm 0.a_1a_2\cdots a_n\times 10^m\ (a_1\neq 0)$$
是 x 的有 n 位有效数字的近似数，则
$$\varepsilon_r\leqslant\frac{1}{2a_1}\times 10^{1-n},$$

(2) 若
$$x^*=\pm 0.a_1a_2\cdots a_n\times 10^m\ (a_1\neq 0)$$
是 x 的近似数，且
$$\varepsilon_r\leqslant\frac{1}{2(a_1+1)}\times 10^{1-n},$$
则 x^* 具有 n 位有效数字。

证 (1) 由 $|x^*|\geqslant a_1\times 10^{m-1}$ 及 $|x-x^*|\leqslant\frac{1}{2}\times 10^{m-n}$ 得
$$|e_r|=\left|\frac{e}{x^*}\right|=\frac{|x-x^*|}{|x^*|}\leqslant\frac{1}{2}10^{m-n}\times\frac{10^{1-m}}{a_1}=\frac{1}{2a_1}\times 10^{1-n},$$
故 $\varepsilon_r\leqslant\frac{1}{2a_1}\times 10^{1-n}$。

(2) 由 $|x^*|\leqslant(a_1+1)\times 10^{m-1}$ 及 $\varepsilon_r\leqslant\frac{1}{2(a_1+1)}\times 10^{1-n}$ 得
$$|x-x^*|\leqslant|x^*|\varepsilon_r\leqslant(a_1+1)\times 10^{m-1}\times\frac{1}{2(a_1+1)}\times 10^{1-n}$$
$$=\frac{1}{2}\times 10^{m-n},$$
故 x^* 有 n 位有效数字。 证毕

例 1.4 求 $\sqrt[3]{100}$ 的近似数，使其相对误差小于 0.1%，问应取几位有效数字？

解 由定理1.1的(1)，令

$$\varepsilon_r \leqslant \frac{1}{2a_1} \times 10^{1-n} < 0.1\%,$$

而 $4 < \sqrt[3]{100} < 5$，故 $a_1 = 4$，代入上述不等式得

$$10^n > 10^4/8,$$

故可取 $n=4$，即取4位有效数字可使其相对误差小于0.1%。

三、误差的传播

实际问题中欲求的量，大都不是由直接测量求得，而是测量与之有关的量再经过运算得出的。如求长方体的体积，通常由测量其长、宽、高再相乘得出。因此怎样从长、宽、高的误差来了解所算得长方体体积的误差是一个有意义的问题。

一般地说，设有多元函数

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

x_k^* 为 $x_k = x_k^{(0)}$ 的近似数，试用 $e_k = x_k^{(0)} - x_k^*$，$k = 1 \sim n$，估计，

$$\begin{aligned} e &= y^{(0)} - y^* \\ &= f(x_1^{(0)}, x_2^{(0)}, \cdots, x_n^{(0)}) - f(x_1^*, x_2^*, \cdots, x_n^*)。 \end{aligned}$$

设 f 对所有 x_k 有连续偏导数，由一阶 Taylor 公式得

$$e = y^{(0)} - y^* \approx \sum_{k=1}^{n} \left(\frac{\partial f}{\partial x_k}\right)_* (x_k^{(0)} - x_k^*) = \sum_{k=1}^{n} \left(\frac{\partial f}{\partial x_k}\right)_* e_k, \tag{1.3}$$

其中 $\left(\frac{\partial f}{\partial x_k}\right)_* = \frac{\partial f}{\partial x_k}(x_1^*, x_2^*, \cdots, x_n^*)$。由 $\frac{\partial f}{\partial x_k}$，$k = 1 \sim n$，的连续性可知，当 $|e_k|$，$k = 1 \sim n$，很小时，上式的近似程度一般很好。

相对误差

$$e_r = \frac{e}{y^*} \approx \sum_{k=1}^{n} \left(\frac{\partial f}{\partial x_k}\right)_* \frac{x_k^*}{y^*} (e_k)_r。 \tag{1.4}$$

由(1,3)，(1.4)两式可见：下列两组数

$$\left\{\left(\frac{\partial f}{\partial x_k}\right)_*\right\}_1^n, \quad \left\{\left(\frac{\partial f}{\partial x_k}\right)_* \frac{x_k^*}{y^*}\right\}_1^n \tag{1.5}$$

表示原始的带误差的数经计算后得出结果的误差将原始误差的放缩倍数，这些重要的数称之为求 y 的**条件数**。当条件数的绝对值很大时，e、e_r 就可能很大（即使 e_k，$(e_k)_r$，$k=1\sim n$，都很小），这时候求 y 的问题称之为**病态问题**或**坏条件问题**。

例 1.5 考虑多项式方程

$$p(x)=\sum_{k=0}^{n}a_kx^k=0$$

的求根问题的条件数。

解 为简单起见，仅考虑某根 α 关于某个系数 a_i 的条件数：$\dfrac{d\alpha}{da_i}$。

将恒等式

$$p(\alpha)=\sum_{k=0}^{n}a_k\alpha^k\equiv 0$$

关于 a_i 求导，注意到 $\alpha=\alpha(a_i)$，得

$$\alpha^i+p'(\alpha)\frac{d\alpha}{da_i}=0,$$

当 $p'(\alpha)\neq 0$ 时

$$\frac{d\alpha}{da_i}=-\frac{\alpha^i}{p'(\alpha)}$$

由此可见，当 $p'(\alpha)\approx 0$ 时，问题是病态的。

第三节 减少运算误差的几个原则

现有的计算机只能对有限位数作运算，从而在运算中产生误差是不可避免的。而解实际问题时，数值运算的次数成千上万是屡见不鲜的，因此必须努力防止误差的产生、传播与扩大。

下面给出几个原则是防患于未然的积极做法。

一、避免两个相近数相减

在数值计算中，直接将两个相近数相减会导致有效数字的严重损失。如 $a=128.375$，$b=128.314$ 都有 6 位有效数字，而其差 $a-b=0.061$ 至多只有两位有效数字。除了上述这种硬性而孤立的减法外，我们努力选择计算方法，避免有效数字的大量损失。

如当 x 是很大的正数，需要计算 $\sqrt{x+1}-\sqrt{x}$ 时，若利用恒等式

$$\sqrt{x+1}-\sqrt{x}=\frac{1}{\sqrt{x+1}+\sqrt{x}}$$

的右端代替左端作计算，就能保持较多的有效数字，下列各例也是类似的。如

当 $ac\approx 0$，$b>0$ 时，可利用

$$\frac{1}{2a}(-b+\sqrt{b^2-4ac})=\frac{-2c}{b+\sqrt{b^2-4ac}}$$

作计算；

当 $x\approx 0$ 时，可利用

$$\frac{1-\cos x}{\sin x}=\frac{\sin x}{1+\cos x},$$

$$\operatorname{arc\,tg}(2+x)-\operatorname{arc\,tg}2=\operatorname{arc\,tg}\frac{x}{1+2(2+x)}\approx\frac{x}{1+2(2+x)}$$

作计算。

一般地，当 x 与 x^* 很接近时，若有 Taylor 展开式

$$f(x)-f(x^*)=f'(x^*)(x-x^*)+\frac{f''(x^*)}{2!}(x-x^*)^2+\cdots,$$

则常取右端有限项作左端的近似。如 $x\approx 0$ 时，有

$$1-\cos x\approx\frac{x^2}{2!}-\frac{x^4}{4!}+\frac{x^6}{6!},$$

$$e^x - 1 \approx x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^{10}}{10!}。$$

当找不到其他算法避免相近数相减时，可以多取几位有效数字进行计算，以保持结果的精确性。

二、避免大小悬殊的数相加

计算机在作运算前先要对阶，因此大、小数（指绝对值）相加时，会出现大数"掩盖"小数的现象。如

$$a=61025, \quad b=0.9,$$

则在5位表示的计算机上作截尾时，a、b分别被表示成

$$0.61025\times 10^5, \quad 0.00000\times 10^5,$$

将它们相加得

$$0.61025\times 10^5 = a,$$

即 b 在与 a 对阶时被"掩盖"了，在和中就体现不出 b 的存在。

若求

$$S = a + \sum_{k=1}^{1000} b_k, \quad 0.1\leqslant b_k \leqslant 0.9, \; k=1\sim 1000,$$

则加法次序大有讲究。若将 $a=61025$ 与每个 b_k 逐一相加，则如上述其结果仍为 a；这时应该先求出 $\sum_{k=1}^{1000} b_k$ 再与 a 相加，才能得出较精确的结果。

三、减少运算次数　简化计算步骤

减少运算次数不仅减少了误差产生的机会，还可以节约机时，是一举两得的事。

如计算多项式

$$p(x) = \sum_{k=0}^{n} a_k x^k$$

的值。若直接计算 $a_k x^k$ 再求和，共需 $\sum_{k=0}^{n} k = \frac{n}{2}(n+1)$ 次乘法与 n 次加法；如果改为