袁慰平 孙志忠 吴宏伟 闻震初 编

计算方法与实习

Computing Methods and Laboratory

南大学出版社

计 算 方 法 与 实 习

（第3版）

袁慰平　孙志忠
吴宏伟　闻震初　编

东 南 大 学 出 版 社

内 容 提 要

全书分两篇。第1篇为计算方法,包括误差分析、方程求根、线性方程组求解、函数插值、曲线拟合、数值微积分、常微分方程数值解法及矩阵特征值计算等8章,各章末有应用实例、内容小结、复习思考题和习题;第2篇为计算实习,用于指导学生上机实习和供学生自学,与第1篇各章相应共有8个实习,每一实习均给出了该实习的目的与要求、算法概要、用C语言编写并在Turbo C 2.0上调试通过的程序、实例及上机实习题。

本书取材适当,思路清晰,富有启发性,便于教学,可作为高等工科院校非数学专业学生的教材,也可作同等程度的自学教材,或科技人员的参考书。

图书在版编目(CIP)数据

计算方法与实习/袁慰平等编. —3版. —南京:东南大学出版社,2000.7
ISBN 7-81050-828-8

Ⅰ.计… Ⅱ.袁… Ⅲ.数值计算-计算方法 Ⅳ.O241

中国版本图书馆CIP数据核字(2000)第35473

东南大学出版社出版发行
(南京四牌楼2号 邮编210096)
出版人:宋增民
江苏省新华书店经销 华东有色地研所印刷厂印刷
开本:787mm×1092mm 1/16 印张:15 字数:371千字
2000年7月第3版 2003年2月第5次印刷
印数:30001-40000 定价:19.50元

第 3 版说明

本书自 1991 年第 2 版出版以来已有近 10 年，在此期间受到兄弟院校同行老师和读者的关心与支持，给了我们很大鼓舞，在此表示深深的谢意。

在第 3 版的编写过程中吸取了广大读者的宝贵意见和我们在教学实践中的经验与体会，对内容主要作了如下修改：

1. 在第 2 章中增加了收敛阶的概念，用以刻画收敛速度；并给出了迭代法发散的充分条件；对二分法、割线法及劈因子法进行了改写；删去了简化牛顿法和牛顿下山法。

2. 在第 3 章中简单介绍了向量范数和矩阵范数的概念及计算，使得讲述线性方程组近似解的误差及迭代法收敛的概念能更为深刻些。

3. 在第 4 章中删去了逐步线性插值和埃尔米特插值两节；为了第 6 章辛卜生求积公式误差估计的需要，在第 2 节加了一类带导数条件的插值；另外还充实了分段插值的一些内容。

4. 在第 5 章中删去了原第 2 节正交多项式的曲线拟合，充实了原第 1 节的内容并将其分为两节以突出数据的曲线拟合。

5. 在第 6 章中增加了高斯求积公式简介一节，不作理论推导，只介绍其思想方法和算法。

6. 在第 7 章中统一地定义了单步显式公式、单步隐式公式及线性多步公式的局部截断误差和整体截断误差；删去了步长的自适应一小段。

7. 在第 8 章中对特征值 QR 方法采用施密特正交化算法，使之与线性代数课程的相关内容衔接得更为密切。

8. 不加证明但指明出处的共列出 5 个重要定理：方程求根埃特金加速算法的阶数；牛顿法的大范围收敛条件；线性方程组迭代法收敛充分必要条件；三次样条插值函数的误差估计以及求积节点为高斯点的充分必要条件。

9. 在第 2,4,5,6,7 等章中均增加了应用实例一节，使学生知道计算方法在实际问题中的具体应用，这些内容可由学生自学。

10. 第 2 篇的内容仍与第 1 篇平行，用以指导学生能及时将所学方法在计算机上实习，所有程序改用 C 语言编写并在 Turbo C 2.0 上调试通过，学生编写程序时，可用自己所熟悉的其它语言改写，以便达到自己动手做实习的目的。

11. 更换并调整了部分例题和习题。本书另配有习题解答(待出版)。

在本书第 3 版编写过程中得到东南大学教务处、应用数学系及东南大学出版社等领导的关心和支持，在此一并表示感谢。

本版编写力求准确、精练，但限于水平可能还会有疏漏和不妥之处，恳请读者批评指正。

编者

2000 年 4 月

前　言

随着电子计算机的迅速发展和广泛应用，在众多的领域内，科学计算已逐渐超过和代替了实验方法。人们愈来愈认识到科学计算是科学研究的第3种方法，特别是工科大学的学生，应具备这方面的知识与能力。现在，很多工科专业已经开设“计算方法”课，并列为大学生的必修课程。

本书是在多年讲授该课程的讲义基础上修订而成的。初稿自1982年以来在本校及兄弟院校中使用过多遍，反映良好。在内容取舍上，本书力求精简，并重视理论联系实际，在讲授计算方法的基本内容及计算机上的常用算法的同时，专门安排了计算实习的有关内容，包括主要算法的程序框图及用BASIC语言编写的程序与例题。我们认为，只要牢固地掌握了这些基本方法，便不难进一步学习计算方法中其它更为深入的内容。

全书分两篇，第1篇为计算方法，用于课堂讲授，一般需40～45学时，第2篇为计算实习，用于学生上机演算，应与第1篇并行使用。读者最好在读通例题后，转换为其它语言(如FORTRAN或PASCAL语言)，再上机演算。

本书第1篇由袁慰平、张令敏编写，第2篇由黄新芹、闻震初编写。由于水平有限，缺点与错误在所难免，恳请读者批评指正。

南京大学数学系何旭初教授和王嘉松副教授仔细地审阅了全部书稿，提出了不少宝贵意见，我们作了修改和补充，最后又由何旭初教授审定。在本书使用、修改和出版过程中得到本教研室有关同志及东南大学出版社的支持与帮助，在此一并表示衷心感谢。

编者

1988年6月

目　录

第1篇　计算方法

第2篇　计算实习

第1篇　计算方法

1　绪论

1.1　计算方法的对象与特点

计算方法是研究数学问题的数值解及其理论的一个数学分支，它涉及面很广，如：代数、微积分、微分方程等都有数值解的问题。自电子计算机成为数值计算的主要工具以来，计算方法主要研究适合于在计算机上使用的数值计算方法及与此相关的理论，包括方法的收敛性、稳定性以及误差分析，还要根据计算机的特点研究计算时间最短、需要计算机内存最少的计算方法。某些在理论上虽然不够严格，但通过实际计算、对比分析等手段，被证明是行之有效的方法也可采用。因此计算方法除具有数学的抽象性与严格性外，还具有应用的广泛性与实际试验的技术性等，是一门与计算机密切结合的实用性很强的课程。

1.2　误差的来源及误差的基本概念

1.2.1　误差的来源

一个物理量的真实值和我们算出的值往往不相等，其差称为误差。引起误差的原因是多方面的：

1）从实际问题转化为数学问题，即建立数学模型时，对被描述的实际问题进行了抽象和简化，忽略了一些次要因素，这样建立的数学模型虽然具有“精确”、“完美”的外衣，其实只是客观现象的一种近似。这种数学模型与实际问题之间出现的误差称为**模型误差**。

2）在给出的数学模型中往往涉及一些根据观测得到的物理量，如电压、电流、温度、长度等，而观测不可避免会带误差，这种误差称为**观测误差**。

3）在计算中常常遇到只有通过无限过程才能得到的结果，但实际计算时，只能用有限过程来计算。如无穷级数求和，只能取前面有限项求和来近似代替，于是产生了有限过程代替无限过程的误差，称为**截断误差**，这是计算方法本身出现的误差，所以也称为方法误差，这种误差是本课程中需要特别重视的。

4）在计算中遇到的数据可能位数很多，也可能是无穷小数，如$\sqrt{2}$、1/3等，但计算时只能对有限位数进行运算，因而往往进行四舍五入，这样产生的误差称为**舍入误差**。少量舍入误差是微不足道的，但在电子计算机上完成了千百万次运算后，舍入误差的积累有时可能是十分惊人的。

由以上误差来源的分析可以看到：误差是不可避免的，要求绝对准确，绝对严格实际上

是办不到的。既然描述问题的方法都是近似的，那么要求解的绝对准确也就没有意义了。因此在计算方法里讨论的都是近似解，那种认为近似解是不可靠的、不准确的看法是错误的，应该认为求近似解是正常的，问题是怎样尽量设法减少误差，提高精度。在 4 种误差来源的分析中，前 2 种误差是客观存在的，后 2 种是由计算方法所引起的。本课程是研究数学问题的数值解法，因此只涉及后 2 种误差。

1.2.2 绝对误差与绝对误差限

定义 1 设 x^* 为准确值，x 是 x^* 的一个近似值，称 $e = x^* - x$ 为近似值 x 的**绝对误差**，简称**误差**。

这样定义的误差 e 可正可负，所以绝对误差不是误差绝对值。通常我们不能算出准确值 x^*，也不能算出误差 e 的准确值，因为这个值虽然客观存在，但实际计算中是得不到的，得到的只能是误差的某个范围，即根据测量工具或计算情况估计出误差的绝对值不超过某正数 ε，即

$$|e| = |x^* - x| \leqslant \varepsilon$$

称 ε 为近似值 x 的**绝对误差限**，简称**误差限**，有时也可以表示成 $x^* = x \pm \varepsilon$。

例如，用毫米刻度的直尺测量一长度为 x^* 的物体，测得其长度的近似值为 $x = 123$ mm，由于直尺以毫米为刻度，所以其误差不超过 0.5 mm，即

$$|x^* - 123| \leqslant 0.5$$

从这个不等式我们不能得出准确值 x^*，但却知道 x^* 的范围

$$122.5 \leqslant x^* \leqslant 123.5$$

对于给定的正数 ε，若近似值 x 满足

$$|x^* - x| \leqslant \varepsilon$$

则在允许误差 ε 范围内认为 x 就是 x^*，也即近似值 x 和真值 x^* 关于允许误差 ε 可以看成是“重合”的，或者说值 x 关于允许误差 ε 是“准确”的。

1.2.3 相对误差与相对误差限

误差限的大小还不能完全表示出近似值的好坏。例如，测得光速的近似值为 $x = 299\,796$ km/s，误差限为 4 km/s，约为光速本身的十万分之一，显然测量是非常准确的；如果测量运动员的跑速，误差限是 0.01 km/s，即 10 m/s，接近运动员的真正跑速，显然这是十分粗糙的测量。为了较好地反映近似值的精确程度，必须考虑误差与真值的比值，即相对误差。

定义 2 设 x^* 为准确值，x 是 x^* 的一个近似值，则称 $(x^* - x)/x^* = e/x^*$ 为近似值 x 的**相对误差**，记作 e_r。

在实际计算中，通常真值 x^* 总是难以求得的。人们常以

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

作为相对误差。事实上，

$$\bar{e}_r - e_r = \frac{\bar{e}_r^2}{1 + \bar{e}_r} = \frac{e_r^2}{1 - e_r}$$

因而当 $\bar{e}_r$ 和 e_r 有一为小量时，$\bar{e}_r - e_r$ 是该小量的二阶小量。

计算相对误差与计算绝对误差具有相同的困难，因此通常也只能考虑相对误差限，即如果有正数 ε_r，使

$$|e_r| \leqslant \varepsilon_r \text{ 或 } |\bar{e}_r| \leqslant \varepsilon_r$$

则称 ε_r 为 x 的**相对误差限**。

1.2.4 有效数字

在工程上对于测量得到的数经常表示成 $x \pm \varepsilon$，它虽然表示了近似值 x 的准确程度，但用这个量进行数值计算太麻烦，因此希望所写出的数本身就能表示它的准确程度，于是需要引进有效数字的概念。另外，当准确值 x^* 有很多位数时，常常按四舍五入原则得到 x^* 的前几位近似值 x。例如：

$$x^* = \sqrt{3} = 1.732\,050\,808\cdots$$

取 3 位，$x_1 = 1.73$，$\varepsilon_1 < 0.005$；取 5 位，$x_2 = 1.732\,1$，$\varepsilon_2 < 0.000\,05$。它们的误差都不超过末位的半个单位，即

$$|\sqrt{3} - 1.73| < \frac{1}{2} \times 10^{-2}, \qquad |\sqrt{3} - 1.732\,1| < \frac{1}{2} \times 10^{-4}$$

定义 3 如果近似值 x 的误差限是其某一位上的半个单位，且该位直到 x 的第 1 位非零数字一共有 n 位，则称近似值 x 有 n 位**有效数字**（见图 1-2-1）。

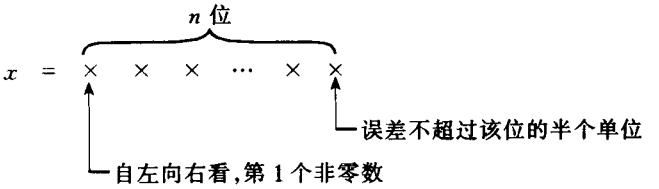

图 1-2-1 有效位数

如$\sqrt{3}$ 的近似值取 $x_1 = 1.73$，则 x_1 有 3 位有效数字；取 $x_2 = 1.732\,1$，则 x_2 有 5 位有效数字；若取 $x_3 = 1.732\,0$，则 x_3 只有 4 位有效数字，因为它的误差限已超过$\frac{1}{2} \times 10^{-4}$。

在讲了有效数字之后，我们规定今后所写出的数都应该是有效数字，如$\sqrt{3}$ 的近似值根据所需要的不同位数的有效数字应是 1.73 或 1.732 或 1.7321，而不能是 1.7320。同时，在同一问题中，参加运算的数，都应尽可能有相同位数的有效数字。

例 1 对下列各数写出具有 5 位有效数字的近似值：

$$236.478, \quad 0.002\,347\,11, \quad 9.000\,024, \quad 9.000\,034 \times 10^3$$

按定义，上述各数具有 5 位有效数字的近似值分别是

$$236.48, \quad 0.002\,347\,1, \quad 9.000\,0, \quad 9.000\,0 \times 10^3$$

注意 $x^* = 9.000\,024$ 的 5 位有效数字近似值是 9.000 0，而不是 9，因为 9 只有 1 位有效数字。

例 2 指出下列各数有几位有效数字：

$$2.000\,4, \quad -0.002\,00, \quad -9\,000, \quad 9 \times 10^3, \quad 2 \times 10^{-3}$$

按定义，上述各数的有效位数分别是 5，3，4，1，1。

1.2.5 数据误差的影响

数值运算中由于所给数据的误差必然引起函数值的误差，这种数据误差的影响较为复杂，一般采用泰勒级数展开的方法来估计。如计算

$$y = f(x_1, x_2)$$

的值，设给定值(即数据)x_1, x_2 是近似值，则由此计算得到的 y 也只能是近似值，现在来研究 y 的绝对误差与相对误差。

设 x_1^*, x_2^* 为准确值，其函数准确值为 $y^* = f(x_1^*, x_2^*)$，于是函数值 y 的误差是

$$e(y) = y^* - y = f(x_1^*, x_2^*) - f(x_1, x_2)$$

将 $f(x_1^*, x_2^*)$ 在 (x_1, x_2) 处作泰勒展开，并取一阶泰勒多项式，则得 $e(y)$ 的近似表示式为

$$\begin{aligned} e(y) = y^* - y &\approx \frac{\partial f(x_1, x_2)}{\partial x_1}(x_1^* - x_1) + \frac{\partial f(x_1, x_2)}{\partial x_2}(x_2^* - x_2) \\ &= \frac{\partial f(x_1, x_2)}{\partial x_1} e(x_1) + \frac{\partial f(x_1, x_2)}{\partial x_2} e(x_2) \end{aligned} \tag{2.1}$$

式中，$e(x_1) = x_1^* - x_1$；$e(x_2) = x_2^* - x_2$。

式(2.1)的左端实际上就是函数 $y = f(x_1, x_2)$ 在 (x_1, x_2) 处分别有增量 $\Delta x_1 = x_1^* - x_1 = e(x_1)$，$\Delta x_2 = x_2^* - x_2 = e(x_2)$ 时，函数的全增量 Δy，因此 $e(y)$ 的近似表达式实质上就是 y 的全微分 $\mathrm{d}y$，即

$$e(y) = \Delta y \approx \mathrm{d}y = \frac{\partial f(x_1, x_2)}{\partial x_1}\mathrm{d}x_1 + \frac{\partial f(x_1, x_2)}{\partial x_2}\mathrm{d}x_2$$

函数的相对误差

$$\begin{aligned} e_r(y) = \frac{e(y)}{y} &\approx \frac{\partial f(x_1, x_2)}{\partial x_1}\frac{x_1}{y}\frac{e_1}{x_1} + \frac{\partial f(x_1, x_2)}{\partial x_2}\frac{x_2}{y}\frac{e_2}{x_2} \\ &= \frac{\partial f(x_1, x_2)}{\partial x_1}\frac{x_1}{y} e_r(x_1) + \frac{\partial f(x_1, x_2)}{\partial x_2}\frac{x_2}{y} e_r(x_2) \end{aligned} \tag{2.2}$$

式中，$e_r(x_1), e_r(x_2)$ 分别是 x_1, x_2 的相对误差。

利用函数值的误差估计式(2.1)和(2.2)，可以得到两数和、差、积、商的误差估计：

$$e(x_1 + x_2) \approx e(x_1) + e(x_2)$$

$$e(x_1 - x_2) \approx e(x_1) - e(x_2)$$

$$e(x_1 x_2) \approx x_2 e(x_1) + x_1 e(x_2)$$

$$e\left(\frac{x_1}{x_2}\right) \approx \frac{1}{x_2} e(x_1) - \frac{x_1}{x_2^2} e(x_2) \qquad x_2 \neq 0$$

$$e_r(x_1 + x_2) \approx \frac{x_1}{x_1 + x_2} e_r(x_1) + \frac{x_2}{x_1 + x_2} e_r(x_2)$$

$$e_r(x_1 - x_2) \approx \frac{x_1}{x_1 - x_2} e_r(x_1) - \frac{x_2}{x_1 - x_2} e_r(x_2)$$

$$e_r(x_1 x_2) \approx e_r(x_1) + e_r(x_2)$$

$$e_r\left(\frac{x_1}{x_2}\right) \approx e_r(x_1) - e_r(x_2) \qquad x_2 \neq 0$$

例 3 已测得某物体行程 s^* 的近似值 $s = 800\ \text{m}$，所需时间 t^* 的近似值 $t = 35\ \text{s}$。若已知 $|t^* - t| \leqslant 0.05\ \text{s}$，$|s^* - s| \leqslant 0.5\ \text{m}$，试求平均速度 v 的绝对误差限和相对误差限。

解 因为 $v = \dfrac{s}{t}$，由商的误差估计式有

$$e(v) = e\left(\frac{s}{t}\right) \approx \frac{1}{t}e(s) - \frac{s}{t^2}e(t), \quad e_r(v) = e_r\left(\frac{s}{t}\right) \approx e_r(s) - e_r(t)$$

得

$$\begin{aligned} |e(v)| &\approx \left|\frac{1}{t}e(s) - \frac{s}{t^2}e(t)\right| \\ &\leqslant \frac{1}{t}|e(s)| + \frac{s}{t^2}|e(t)| \\ &\leqslant \frac{1}{35} \times 0.5 + \frac{800}{35^2} \times 0.05 \approx 0.046\,9 \leqslant 0.05 \\ |e_r(v)| &\approx |e_r(s) - e_r(t)| \\ &\leqslant |e_r(s)| + |e_r(t)| \\ &\leqslant \frac{0.5}{800} + \frac{0.05}{35} \approx 0.002\,05 \end{aligned}$$

所以平均速度 v 的绝对误差限和相对误差限分别为 0.05 和 0.002 05。

应该指出，在由误差估计式得出绝对误差限和相对误差限的估计时，由于取了绝对值并用三角不等式放大，因此是按最坏情形得出的，所以由此得出的结果是很保守的。事实上，出现最坏情形的可能性是很小的。因此近年来出现了一系列关于误差的概率估计。一般说来为了保证运算结果的精确度，只要根据运算量的大小，比结果中所要求的有效数字的位数多取 1 位或 2 位进行计算就可以了。

1.3 机器数系

1.3.1 数的浮点表示

一个实数在科学计算中常常被表示成浮点形式。例如 456.789，-6.473，0.005 67，0.321 等被分别表示成 $0.456\,789 \times 10^3$，$-0.647\,3 \times 10^1$，0.567×10^{-2}，0.321×10^0，其中 0.456 789，$-0.647\,3$，0.567，0.321 等称为浮点表示的**尾数部**，10^3，10^1，10^{-2}，10^0 等称为浮点表示的**定位部**，这种表示形式可以使得一个数的数量级一目了然，更重要的是它可以扩大计算机表示数的范围。

一个**基数**为 β 的 t 位数字的浮点表示形式为

$$x = (\pm 0.a_1a_2\cdots a_t)\beta^p \tag{3.1}$$

式中，$\beta \geqslant 2$ 是整数，通常取 $\beta = 2,8,10,16$；每个 a_i 都是整数，且 $0 \leqslant a_i \leqslant \beta - 1$；$t$ 是计算机的**字长**；带有符号的整数 p 称为指数，也称为计算机的**阶码**，它有固定的**下限** L 和**上限** U，即 $L \leqslant p \leqslant U$，$L$、$U$ 和 t 是由该计算机的硬件所决定的某些常数。尾数部

$$s = \pm 0.a_1a_2\cdots a_t = \pm\left(\frac{a_1}{\beta} + \frac{a_2}{\beta^2} + \cdots + \frac{a_t}{\beta^t}\right) \tag{3.2}$$

而

$$x = s \times \beta^p \tag{3.3}$$

若规定 $a_1 \neq 0$，则 $\beta^{-1} \leqslant |s| < 1$，此时 x 称为**规格化浮点数**。今后除特别指出外，都认为浮点数是规格化表示的。

1.3.2 机器数系

上述数的浮点表示，几乎是当今所有计算机都采用的表示法。把计算机中浮点数所组成的集合加上“机器零”记为 F，则 F 被以下4个参数所描述：基数 β、字长 t、阶码范围 $[L, U]$，我们称这个集合 F 为**机器数系**，需要指出的是机器数系 F 是一个离散的分布不均匀的有限集。

例如，设有一个二进制的2位字长的计算机，即 $\beta = 2, t = 2$，其指数 $p \in [-1, 1]$，则它所能表示的数只有如下几个：

当 $p = -1$ 时，有：$\pm (0.10 \times 2^{-1})_2 = \pm (0.25)_{10}$，$\pm (0.11 \times 2^{-1})_2 = \pm (0.375)_{10}$；

当 $p = 0$ 时，有：$\pm (0.10 \times 2^{0})_2 = \pm (0.5)_{10}$，$\pm (0.11 \times 2^{0})_2 = \pm (0.75)_{10}$；

当 $p = 1$ 时，有：$\pm (0.10 \times 2^{1})_2 = \pm (1)_{10}$，$\pm (0.11 \times 2^{1})_2 = \pm (1.5)_{10}$。加上机器零，共13个数，它们在数轴上的表示见图 1-3-1。

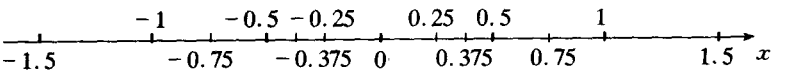

图 1-3-1 机器数的表示

不难证明集合 F 仅含有

$$2(\beta - 1)\beta^{t-1}(U - L + 1) + 1 \tag{3.4}$$

个数，而且这些数不是等间隔地分布在数轴上。

当 $\beta = 10, t = 4, -L = U = 99$ 时，此计算机的机器数系 F 仅含有 3 582 001 个数，-0.1000×10^{-99} 和 0.1000×10^{-99} 是该数系 F 中绝对值最小的非零数，而 -0.9999×10^{99} 和 0.9999×10^{99} 分别是此数系 F 中的最小数和最大数，若计算的中间结果超出了上述范围，则称为**溢出**。

由于机器数系有上述特性，因此一个实数 x 进入计算机后，成为计算机里的数，称它为 x 的机器数，用 $\mathrm{fl}(x)$ 表示，一般讲 $\mathrm{fl}(x)$ 只是 x 的一个近似值。

例如，对于数 $x = 0.6$，在上述二进制2位字长计算机的机器数系 F 里找不到这个数，通常取与 x 最靠近的数 0.5 作为 x 的近似值，即它的机器数 $\mathrm{fl}(0.6) = (0.1\upsilon \times 2^0)_2$。

目前的计算机分截断机和舍入机两种。对于截断机，$\mathrm{fl}(x)$ 取 x 的前 t 位数字；对于舍入机，$\mathrm{fl}(x)$ 按四舍五入原则取 x 的前 t 位数字。

例 4 假设具有十进制、3位字长、$-L = U = 5$ 的2台计算机，一台是截断机，另一台是舍入机，则它们对下述实数的规格化浮点数如表 1-3-1 所示。

表 1-3-1　实数的浮点表示

实　　数	截断机浮点数	舍入机浮点数
1278	0.127×10^4	0.128×10^4
$-43\frac{1}{3}$	-0.433×10^2	-0.433×10^2
0.005 669	0.566×10^{-2}	0.567×10^{-2}
123 456	溢出	溢出

这 2 台计算机能表示的最大数和最小数分别是 0.999×10^5、-0.999×10^5，因此数 123 456 超出了它所能表示的范围。

1.3.3　机器数的相对误差限

设 $x = (\pm 0.b_1 b_2 \cdots b_t b_{t+1} \cdots)\beta^p$，$b_1 \neq 0$。对于舍入机，当 $| b_{t+1} | \geqslant \frac{\beta}{2}$ 时，$\mathrm{fl}(x) = (\pm 0.b_1 b_2 \cdots \overline{b_t + 1})\beta^p$；当 $| b_{t+1} | < \frac{\beta}{2}$ 时，$\mathrm{fl}(x) = (\pm 0.b_1 b_2 \cdots b_t)\beta^p$。无论哪种情形均有 $| x - \mathrm{fl}(x) | \leqslant \frac{1}{2}\beta^{-t}\beta^p$。因而

$$\left| \frac{x - \mathrm{fl}(x)}{x} \right| \leqslant \frac{\frac{1}{2}\beta^{-t}\beta^p}{\beta^{-1}\beta^p} = \frac{1}{2}\beta^{1-t}$$

对于截断机，$| x - \mathrm{fl}(x) | \leqslant \beta^{-t}\beta^p$，因而 $\left| \frac{x - \mathrm{fl}(x)}{x} \right| \leqslant \beta^{1-t}$。

综上所述，我们有如下结论：在浮点数范围内，每个非零数 x，其机器数 $\mathrm{fl}(x)$ 的相对误差限为

$$\frac{| x - \mathrm{fl}(x) |}{| x |} \leqslant \begin{cases} \frac{1}{2}\beta^{1-t} & \text{舍入机} \\ \beta^{1-t} & \text{截断机} \end{cases} \tag{3.5}$$

所以当使用的计算机确定后，相应的机器数的相对误差限也就确定了，此相对误差限通常称为计算机的精度。如通常用的 8 位字长的十进制计算机，其机器数的相对误差限为

$$\begin{cases} \frac{1}{2} \times 10^{1-8} = \frac{1}{2} \times 10^{-7} & \text{舍入机} \\ 10^{1-8} = 10^{-7} & \text{截断机} \end{cases}$$

1.4　误差危害的防止

误差分析在数值运算中是一个重要而又复杂的问题，因为每步运算都有可能产生误差，而一个工程或科学计算问题往往要算千万次，如果每步运算都分析误差，这是不可能的，也是不必要的。这里提出的若干原则，就是为了鉴别计算结果的可靠性和防止误差危害现象的产生。