

21世纪高等院校选用教材

工科类

数值分析原理

封建湖 车刚明 聂玉峰 编著

科学出版社

21 世纪高等院校选用教材(工科类)

数值分析原理

封建湖 车刚明 聂玉峰 编著

科学出版社

2001

内 容 简 介

本书系统地介绍了现代科学与工程计算中常用的数值计算方法及有关的理论和应用.全书共分9章,包括误差分析,函数插值,函数逼近,数值积分与数值微分、线性方程组的直接解法和迭代解法,非线性方程的数值解法,矩阵特征值与特征向量的计算,以及常微分方程初值问题的数值解法等.本书基本概念清晰准确,理论分析科学严谨,语言叙述通俗易懂,结构编排由浅入深,注重启发性.本书始终贯穿一个基本理念,即在数学理论上等价的方法在实际数值计算时往往是不等效的,因此,本书精选了大量的计算实例,用来说明各种数值方法的优劣与特点.各章末还有一定数量的习题供读者练习之用.

读者对象:高等院校工科研究生和数学系各专业本科生,从事科学与工程计算的科研工作者.

图书在版编目(CIP)数据

数值分析原理/封建湖,车刚明,聂玉峰 编著. -北京:
科学出版社,2001.9

[21世纪高等院校选用教材(工科类)]

ISBN 7-03-009730-0

I. 数… II. ①封…②车…③聂… III. 数值计算-计算方法-高等学校-教材 IV. O241

中国版本图书馆 CIP 数据核字(2001)第 058464 号

科学出版社 出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

新蕾印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2001年9月第 一 版 开本:720×1000 1/16

2001年9月第一次印刷 印张:21 1/4

印数:1—3 000 字数:378 000

定价:25.00 元

(如有印装质量问题,我社负责调换(新欣))

前 言

本书作为高等院校工科硕士研究生和数学系各专业本科生的“数值分析”(或“计算方法”)课程的教科书,系统地介绍了现代科学与工程计算中常用的数值计算方法、概念以及有关的理论分析和应用.全书共9章,包括误差分析,函数插值,函数逼近,数值积分与数值微分、线性方程组的直接解法,线性方程组的迭代解法,非线性方程的数值解法,矩阵特征值与特征向量的计算,以及常微分方程初值问题的数值解法等.考虑到不少工科院校还没有普及“计算方法”,数学系各专业本科学生也是初学本门课程,因此,本书从零开始讲起,只要具备高等数学、线性代数知识的学生就可以使用本教材.

本书主要特点如下:

第一,叙述简洁流畅,语言通俗易懂.本书编著者都是长期从事数值分析教学与研究工作的中青年骨干教师,具有丰富的教学经验和实际计算经验,对各层次的学生非常了解.因此,在编写本书的过程中,作者充分考虑了学生的知识水平,特别注重语言叙述的简洁流畅、通俗易懂;内容组织的由浅入深、过渡自然;理论分析科学严谨.另外,书中的概念也与实际背景紧密结合,可以激发学生的学习兴趣.

第二,取材合理,观点较新.考虑到本书的使用对象,在内容的选取上本书保留了一些经典的数值方法,充实了一些必要的理论知识,并增加了一些对方法进一步推广方面的分析与讨论,以适应不同层次学生的学习需要和今后研究工作的需要,如常微分方程数值解一章中的线性多步法的收敛性、相容性、稳定性以及预估-校正法等,求解非线性方程的不动点迭代法的斯蒂芬森加速收敛技术,对弦割法、抛物线法的非局部收敛性也给出了充分条件,等等.本书引入了较现代的数学工具,如不动点定理、压缩映射原理、赋范线性空间等,这些概念的引入不仅能使学员更深刻地理解本书内容与方法,而且对他们未来的学习与研究工作也是很有帮助的.

第三,加强了数值试验的内容.本书自始至终贯穿一个基本理念,即在数学理论上等价的方法在实际数值计算上往往是不等效的,因此特别注重数值计算的实践.其他同类书籍中对这点没有给予足够重视,但根据作者多年的计算经验证明这是非常有效的、适用于工程技术人员实际使用的方法,如方程求根的斯蒂芬森方法、微分方程数值解法中的预估-校正方法等,都给了足够的

重视. 书中精选了一些很有说服力的数值算例, 如线性方程组的迭代解法、非线性方程求根、矩阵特征值问题、常微分方程数值解等章中基本上是用不同的方法求解同一个问题, 以便更清楚地观察各种方法的优劣与特点. 对舍入误差的产生机理及算法的数值稳定性, 也都作了较详细的分析. 各章备有一定数量的习题供读者练习之用, 其中一些习题丰富, 补充了书中的相关内容.

讲授全书大约需要 80 学时, 去除打“*” (标有“*”的为选用内容) 的内容后, 约需 60 学时. 授课教师可根据实际学时数进行取舍.

第一、二、三章由聂玉峰执笔, 第四、五、六章由车刚明执笔, 第七、八、九章由封建湖执笔, 最后由封建湖统一定稿.

在本书编写过程中欧阳洁老师也多次参与了讨论, 周天孝教授仔细审阅了书稿, 提出了宝贵的意见, 西北工业大学教务处和科学出版社的有关同志对本书的出版给予了极大的帮助, 我们深表感谢. 限于作者的水平, 加之时间仓促, 缺点和错误在所难免, 敬请指正.

作 者

2001 年 2 月于西北工业大学

目 录

第一章 绪 论	(1)
§ 1.1 数值分析的对象与任务	(1)
§ 1.2 误差基础知识	(2)
1.2.1 误差来源	(2)
1.2.2 误差度量	(3)
1.2.3 初值误差传播	(6)
§ 1.3 舍入误差分析及数值稳定性	(11)
1.3.1 浮点数系及其运算的舍入误差	(11)
1.3.2 算法的数值稳定性	(14)
习题 1	(16)
第二章 函数插值	(17)
§ 2.1 插值问题	(17)
§ 2.2 插值多项式的构造方法	(19)
2.2.1 拉格朗日插值法	(19)
2.2.2 牛顿插值法	(22)
2.2.3 等距节点插值公式	(26)
2.2.4 带导数的插值问题	(30)
§ 2.3 分段插值法	(34)
2.3.1 高次插值的评述	(34)
2.3.2 分段插值	(37)
2.3.3 三次样条插值	(39)
*2.3.4 B样条插值	(46)
习题 2	(54)
第三章 函数逼近	(56)
§ 3.1 赋范线性空间与函数逼近问题	(56)
3.1.1 赋范线性空间	(56)
3.1.2 函数逼近问题	(57)
§ 3.2 内积空间与正交多项式	(58)
3.2.1 内积空间	(58)

3.2.2 正交多项式的性质	(61)
3.2.3 常用的正交多项式系	(65)
§ 3.3 最佳平方逼近与广义 Fourier 级数	(71)
3.3.1 最佳平方逼近问题的求解	(72)
3.3.2 基于正交函数基的最佳平方逼近	(76)
* 3.3.3 广义 Fourier 级数	(78)
§ 3.4 曲线拟合的最小二乘方法	(81)
3.4.1 曲线拟合模型及其求解	(81)
3.4.2 关于离散 Gram 矩阵的进一步讨论	(83)
3.4.3 用关于点集的正交函数系作最小二乘曲线拟合	(87)
* § 3.5 最佳一致逼近多项式	(89)
3.5.1 魏尔斯特拉斯定理	(89)
3.5.2 最佳一致逼近多项式的存在惟一性	(90)
3.5.3 最佳一致逼近多项式求法的讨论	(95)
习题 3	(99)
第四章 数值积分与数值微分	(101)
§ 4.1 数值积分概述	(101)
4.1.1 求积公式的代数精确度	(101)
4.1.2 收敛性与稳定性	(102)
§ 4.2 牛顿-柯特斯公式	(102)
4.2.1 插值型求积公式	(102)
4.2.2 牛顿-柯特斯公式	(103)
4.2.3 复化求积公式	(104)
4.2.4 截断误差	(106)
4.2.5 区间逐次分半求积法	(108)
§ 4.3 龙贝格求积算法	(109)
§ 4.4 高斯型求积公式	(111)
4.4.1 一般理论	(111)
4.4.2 高斯-勒让德求积公式	(115)
4.4.3 高斯-切比雪夫求积公式	(117)
4.4.4 高斯-拉盖尔求积公式	(117)
4.4.5 高斯-埃尔米特求积公式	(117)
* § 4.5 奇异积分与振荡函数积分的计算	(119)
4.5.1 无界函数积分的计算	(119)

4.5.2 无穷区间积分的计算	(121)
4.5.3 振荡函数积分的计算	(123)
* § 4.6 二重积分的计算	(124)
4.6.1 基本方法	(124)
4.6.2 复化求积公式	(125)
4.6.3 高斯型求积公式	(127)
§ 4.7 数值微分	(127)
4.7.1 插值法	(127)
4.7.2 泰勒展开法	(128)
习题 4	(129)
第五章 解线性代数方程组的直接法	(131)
§ 5.1 高斯消去法	(131)
5.1.1 高斯顺序消去法	(131)
5.1.2 高斯主元消去法	(135)
§ 5.2 矩阵三角分解法	(136)
5.2.1 直接三角分解法	(139)
5.2.2 列主元直接三角分解法	(141)
5.2.3 平方根法	(143)
5.2.4 三对角和块三对角方程组的追赶法	(146)
§ 5.3 矩阵的条件数和方程组的性态	(148)
5.3.1 向量和矩阵范数	(148)
5.3.2 扰动方程组解的误差界	(152)
5.3.3 矩阵的条件数和方程组的性态	(153)
5.3.4 关于病态方程组的求解	(155)
习题 5	(156)
第六章 解线性代数方程组的迭代法	(158)
§ 6.1 向量和矩阵序列的极限	(158)
6.1.1 极限概念	(158)
6.1.2 序列收敛的等价条件	(158)
§ 6.2 迭代法的基本理论	(160)
6.2.1 简单迭代法的构造	(161)
6.2.2 简单迭代法的收敛性和收敛速度	(161)
6.2.3 高斯-赛德尔迭代法及其收敛性	(164)
§ 6.3 几种常用的迭代法	(166)

6.3.1 雅可比迭代法	(166)
6.3.2 与雅可比法相应的高斯-赛德尔迭代法	(168)
6.3.3 逐次超松弛(SOR)迭代法	(170)
* § 6.4 最速下降法与共轭梯度法	(173)
6.4.1 最速下降法	(174)
6.4.2 共轭梯度法	(175)
习题 6	(179)
第七章 非线性方程求根	(182)
§ 7.1 二分法	(182)
§ 7.2 迭代法的算法和理论	(185)
7.2.1 不动点迭代法	(185)
7.2.2 不动点迭代法的一般理论	(187)
7.2.3 局部收敛性,收敛阶	(190)
§ 7.3 迭代的加速收敛方法	(193)
7.3.1 使用两个迭代值的组合方法	(193)
7.3.2 使用三个迭代值的组合方法	(195)
§ 7.4 牛顿迭代法	(198)
7.4.1 标准牛顿迭代法及其收敛阶	(199)
7.4.2 重根情形的牛顿迭代法	(203)
7.4.3 牛顿下山法	(205)
§ 7.5 弦割法和抛物线法	(208)
7.5.1 弦割法及其收敛性	(208)
7.5.2 抛物线法	(213)
* § 7.6 非线性方程组的迭代解法简介	(216)
7.6.1 一般概念	(216)
7.6.2 不动点迭代法	(218)
7.6.3 牛顿迭代法	(222)
习题 7	(224)
第八章 矩阵特征值与特征向量计算	(227)
§ 8.1 乘幂法与反幂法	(227)
8.1.1 乘幂法	(227)
8.1.2 乘幂法的加速技术	(233)
8.1.3 反幂法	(236)
§ 8.2 雅可比方法	(237)

8.2.1 古典雅可比方法	(238)
8.2.2 雅可比过关法	(243)
§ 8.3 QR 方法	(243)
8.3.1 反射矩阵与平面旋转矩阵	(244)
8.3.2 矩阵的 QR 分解	(248)
8.3.3 豪斯霍尔德方法	(250)
8.3.4 QR 方法的收敛性	(252)
* 8.3.5 带原点平移的 QR 方法	(254)
§ 8.4 求实对称三对角阵特征值的二分法	(255)
8.4.1 矩阵 A 的特征多项式序列及其性质	(256)
8.4.2 特征值的计算	(259)
习题 8	(261)
第九章 常微分方程初值问题的数值解法	(263)
§ 9.1 引言	(263)
§ 9.2 欧拉方法	(264)
9.2.1 显式欧拉方法	(264)
9.2.2 隐式欧拉方法和欧拉方法的改进	(266)
9.2.3 单步法的局部截断误差和阶	(268)
§ 9.3 龙格-库塔方法	(270)
9.3.1 泰勒方法	(270)
9.3.2 龙格-库塔方法	(271)
* 9.3.3 龙格-库塔方法的其他问题	(276)
§ 9.4 单步法的进一步讨论	(277)
9.4.1 收敛性	(277)
9.4.2 相容性	(279)
9.4.3 稳定性	(280)
§ 9.5 线性多步方法	(283)
9.5.1 线性多步方法的一般问题	(283)
9.5.2 线性多步方法的构造	(287)
* 9.5.3 预估-校正方法	(294)
* § 9.6 线性多步法的进一步讨论	(299)
9.6.1 线性多步法的相容性	(299)
9.6.2 线性多步法的收敛性	(300)
9.6.3 线性多步法的稳定性	(302)

9.6.4 预估-校正法的稳定性	(310)
* § 9.7 一阶方程组与刚性问题简介	(312)
9.7.1 一阶方程组	(312)
9.7.2 刚性问题简介	(314)
习题 9	(315)
参考文献	(318)
附录 关于线性常系数差分方程的几点知识	(319)
参考答案	(321)

第一章 绪 论

§ 1.1 数值分析的对象与任务

科学与工程领域中的问题求解一般需要经历如图 1.1 所示的过程. 某个领域的专家首先提出实际问题, 然后辨析其中的主要矛盾和次要矛盾, 并在合理假设的条件下, 运用各种数学理论、工具和方法, 建立起问题中不同量之间的联系, 进而得到完备的数学模型. 在模型正确建立的必要条件, 即解的存在性与惟一性得到论证后, 现实问题就是如何求得其解. 然而通常所建立的数学模型分析解是很难得到的, 于是退之局限于讨论该模型的各种特殊情形或简化之后模型的分析解, 但这样做不能满足精度要求, 甚至于较大地偏离实际问题. 随着计算机的迅猛发展, 特别是每秒数十亿次计算机系统的诞生, 它为用数值方法求解较少简化的数学模型提供了强大的工具保证.

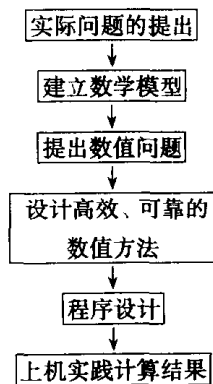


图 1.1

所谓数值问题是指有限个输入数据(问题的自变量、原始数据)与有限个输出数据(待求解数据)之间函数关系的一个明确无歧义的描述. 这正是数值分析所研究的对象.

需要注意的是数学模型不一定是数值问题, 如求解一阶常微分方程初值问题

$$\begin{cases} \frac{dy}{dx} = x + y^2 & x \in [0, 1], \\ y(0) = y_0, \end{cases}$$

要求得到定义于区间 $[0, 1]$ 的函数解析表达式 $y = y(x)$, 这实际上是要求无穷多个输出, 因而它不是数值问题. 但当我们要求得到 n 个点 $\{x_i\}_{i=1}^n \subset [0, 1]$ 处的函数值 $\{y(x_i)\}_{i=1}^n$ 的近似值时, 便成为一数值问题, 该数值问题可以通过欧拉(Euler)方法求得解. 数值分析的任务之一就是提供求得数值问题近似解的方法——算法.

从程序设计的角度来讲,所谓**算法**是由一个或多个进程组成;每个进程明确无歧义地描述由操作及操作对象合成的按一定顺序执行的有限序列;所有进程能够同时执行并且协调地在有限个操作步内完成一个给定问题的求解.这里操作可以是计算机能够完成的算术运算(加减乘除)、逻辑运算、字符运算等.

若算法包含有一个进程则称其为串行算法,否则为并行算法.从算法执行所花费的时间角度来讲,若算术运算占绝大多数时间,则称其为数值型算法(数值方法),否则为非数值型算法.本课程介绍数值型串行算法.

一个算法在保证可靠的大前提下再评价其优劣才是有价值的,所谓算法的可靠性包括如下几个方面:算法的收敛性、稳定性、误差估计等.这些是数值分析研究的第二个任务.

评价一个可靠算法的优劣,应该考虑其时间复杂度(计算机运行时间)、空间复杂度(占据计算机存储空间的多少)以及逻辑复杂度(影响程序开发的周期以及维护).这是数值分析研究的第三个任务.

由于数值分析研究对象以及解决问题方法的广泛适用性,现在流行的软件如 Maple、Matlab、Mathematica 等已将其绝大多数内容设计成简单的函数,简单调用之后便可以得到运行结果.但由于实际问题的具体特征、复杂性,以及算法自身的适用范围决定了应用中必须选择、设计适合于自己特定问题的算法,因而掌握数值方法思想和内容是至关重要的.

鉴于实际问题的复杂性,通常将其具体地分解为一系列子问题进行研究,本课程主要涉及如下几个方面的问题:

函数的插值和逼近、数值积分和数值微分、线性方程组求解、非线性方程(组)求解、代数特征值问题、常微分方程数值求解.

§ 1.2 误差基础知识

1.2.1 误差来源

在建立数学模型的过程中,不可避免地要忽略某些次要矛盾,因而数学模型往往是对实际问题的一种近似表达,我们将数学模型与实际问题的差异称为**模型误差**.同时,数学模型中常常还包含有一些参数,它们是通过仪表观测得到的,将其中包含的误差称之为**观测误差**.在数值分析中不研究这两类误差,总是假定数学模型是正确合理地反映了客观实际问题.

我们将数值问题的精确解与待求解模型的理论分析解之间的差异,称之为**截断误差**或**方法误差**,这是由于算法必须在有限步内执行结束而导致的,它

需要将无穷过程截断为有限过程. 我们知道 $e = 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots$, 如果以 $e_n = 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!}$ 作为 e 的近似值, 则 e 与 e_n 的差异是 e_n 近似 e 的截断误差.

在用计算机实现数值方法的过程中, 由于计算机表示浮点数采用的是固定有限字长, 因而仅能够区分有限个信息, 准确表示在某个有限范围内的某些有理数, 不能准确表示数学中的所有实数, 这样在计算机中表示的原始输入数据、中间计算数据、以及最终输出结果必然产生误差, 称此类误差为舍入误差. 如利用计算机计算 e 的近似值 e_n 时, 实际上得不到 e_n 的精确值, 只能得到 e_n 的近似 e^* . 这样由 $e^* - e = (e^* - e_n) + (e_n - e)$ 知 e^* 作为 e 的近似包含有舍入误差和截断误差两部分.

数值分析课程研究后两类误差的估计、传播和控制.

1.2.2 误差度量

1. 误差及误差限

定义1.1 设 x^* 是准确值 x 的一个近似, 称 $e(x^*) = x^* - x$ 为 x^* 近似 x 的绝对误差, 简称为误差. 在不引起混淆时, 简记符号 $e(x^*)$ 为 e^* .

误差有正有负, 当误差为正时, 近似值较真值偏大, 称此近似为“强近似”; 当误差为负时, 近似值较真值偏小, 称此近似为“弱近似”.

通常准确值 x 是不知道的, 故不能计算出绝对误差 e^* . 如果存在着的正数 $\epsilon^* = \epsilon(x^*)$, 使得绝对误差 $|e^*| = |x^* - x| \leq \epsilon^*$, 则称 ϵ^* 为 x^* 近似 x 的一个绝对误差限, 简称误差限.

此时有 $x \in [x^* - \epsilon^*, x^* + \epsilon^*]$, 工程上习惯用 $x = x^* \pm \epsilon^*$ 表示这一事实. 实际计算中所要求的绝对误差, 是指一个尽可能小的绝对误差限.

2. 相对误差及相对误差限

绝对误差限虽然能够刻画对同一真值不同近似的好坏, 但它不能刻画对不同真值近似程度的好坏. 例如, 对于测量结果 $x = 100 \pm 1$ 和 $y = 10000 \pm 5$, 尽管对 x 和 y 的测量绝对误差限满足 $\epsilon(x^*) = 1 < 5 = \epsilon(y^*)$, 但绝对误差限在真值中所占的比例却有不等关系:

$$\frac{\epsilon(y^*)}{y} \approx \frac{5}{10000} = 0.05\% < 1\% = \frac{1}{100} \approx \frac{\epsilon(x^*)}{x},$$

因此, 我们并不认为测量 $x = 100 \pm 1$ 比 $y = 10000 \pm 5$ 更精确.

定义 1.2 设 x^* 是准确值 $x (\neq 0)$ 的一个近似, 称 $e_r(x^*) = \frac{x^* - x}{x}$ 为 x^* 近似 x 的相对误差. 在不引起混淆时, 简记符号 $e_r(x^*)$ 为 e_r^* .

$$\text{因 } e_r^* - \frac{e^*}{x^*} = \frac{e^*}{x} - \frac{e^*}{x^*} = \frac{(e^*)^2}{x(x+e^*)} = \frac{1}{1+e^*/x} \left(\frac{e^*}{x} \right)^2 = O((e_r^*)^2),$$

即 e_r^* 与 $\frac{e^*}{x^*} = \frac{x^* - x}{x^*}$ 相差一个较 e_r^* 高一阶的无穷小量, 故时常也用后者来计算相对误差 e_r^* .

称数值 $|e_r^*|$ 的上界为相对误差限, 记为 ϵ_r^* , 也可以通过 $\epsilon_r^* = \epsilon^*/x^*$ 来计算. 类似地, 计算相对误差, 是指估计一个尽可能小的相对误差限.

上述示例是对不同测量量的近似, 由 $\epsilon_r(y^*) < \epsilon_r(x^*)$ 知, y^* 对 y 的近似较 x^* 对 x 的近似程度好.

3. 有效数字

为规定一种近似数的表示法, 使得用它表示的近似数自身就直接指示出其误差的大小. 为此需要引出有效数字和有效数的概念.

例 1.1 确定十进制数“四舍五入”方法的误差限.

解 设十进制数 x 有如下的标准形式:

$$x = \pm 10^m \times \underline{0.x_1x_2\cdots x_nx_{n+1}\cdots}, \quad (1.1)$$

其中 m 为整数, $|x_i| \in \{0, 1, 2, \cdots, 9\}$ 且 $x_1 \neq 0$. 对 x 四舍五入保留 n 位数字, 得到近似值 x^* :

$$x^* = \begin{cases} \pm 10^m \times \underline{0.x_1x_2\cdots x_n}, & \text{当 } x_{n+1} \leq 4, \\ \pm 10^m \times \underline{0.x_1x_2\cdots (x_n+1)}, & \text{当 } x_{n+1} \geq 5. \end{cases} \quad (1.2)$$

四舍情形下的误差限

$$|x^* - x| = 10^m \times \underline{\overset{n \uparrow 0}{0.00\cdots 0x_{n+1}\cdots}} \leq 10^m \times \overset{n \uparrow 0}{0.00\cdots 05} = \frac{1}{2} \times 10^{m-n}.$$

五入情形下的误差限

$$\begin{aligned} |x^* - x| &= 10^m \times \left| \overset{n-1 \uparrow 0}{0.00\cdots 01} - \overset{n \uparrow 0}{0.00\cdots 0x_{n+1}\cdots} \right| \\ &= 10^{m-n} \times |1 - \underline{0.x_{n+1}\cdots}| \leq \frac{1}{2} \times 10^{m-n}. \end{aligned}$$

综合以上两点,“四舍五入”法的误差限是 $\frac{1}{2} \times 10^{m-n}$. #

定义 1.3 设 x 的近似值 x^* 有如下标准形式

$$x = \pm 10^m \times 0.x_1x_2\cdots x_n\cdots x_p, \quad (1.3)$$

其中 m 为整数, $\{x_i\} \subset \{0, 1, 2, \cdots, 9\}$ 且 $x_1 \neq 0, p \geq n$. 如果有

$$|e^*| = |x^* - x| \leq \frac{1}{2} \times 10^{m-n},$$

则称 x^* 为 x 的具有 n 位有效数字的近似数, 或称 x^* 准确到 10^{m-n} 位, 其中数字 $x_1, x_2, \cdots, x_n$ 分别被称为 x^* 的第一、第二、 $\cdots$ 、第 n 个有效数字.

从如上定义可以看出, 近似同一真值的近似数的有效数字越多越精确.

当 x^* 准确到末位, 即有 $n=p$, 则称 x^* 为有效数. 综合例 1.1 和定义 1.3 知, 真值 x 通过四舍五入法得到的近似数都是有效数. 有效数的误差限是末位数单位的一半, 可见有效数本身就体现了误差界. 对于有效数 20.12 和 20.120 是不同的. 前者有 4 位有效数字, 绝对误差限是 0.005, 相对误差限是 0.00025; 后者有 5 位有效数字, 绝对误差限是 0.0005, 相对误差限是 0.000025. 可见有效数的末尾是不能随意添加零的.

本书约定, 凡没有标明误差界的近似数都是有效数.

例 1.2 取 $x = 12.49$, 问 x 的近似值 $x_1^* = 12.5, x_2^* = 12.4$, 和 $x_3^* = 12.48$ 分别有几位有效数字, 它们是否是有效数?

解 真值 $x = 12.49 = 10^2 \times 0.1249, m = 2$.

$$|e_1^*| = |x_1^* - x| = 0.01 \leq \frac{1}{2} \times 10^{-1} = \frac{1}{2} \times 10^{2-3},$$

$$|e_2^*| = |x_2^* - x| = 0.09 \leq \frac{1}{2} \times 10^0 = \frac{1}{2} \times 10^{2-2},$$

$$|e_3^*| = |x_3^* - x| = 0.01 \leq \frac{1}{2} \times 10^{-1} = \frac{1}{2} \times 10^{2-3},$$

故 x_1^*, x_2^* 和 x_3^* 近似 x 分别有 3 位、2 位、3 位有效数字, x_1^* 是有效数, x_2^* 和 x_3^* 不是有效数. #

4. 三种度量之间的关系

定义 1.3 表明: 对同一数的近似, 绝对误差越小, 有效数字不会减少; 有效数字增加, 绝对误差一定减少.

相对误差与有效数字之间的关系由如下定理表述:

定理 1.1 设 x 的近似数 x^* 具有形如式(1.3)所示的标准形式:

1° 若 x^* 具有 n 位有效数字, 则相对误差 $|e_r^*| \leq \frac{1}{2x_1} \times 10^{1-n}$;

2° 若相对误差 $|e_r^*| \leq \frac{1}{2(x_1+1)} \times 10^{1-n}$, 则 x^* 至少具有 n 位有效数字.

证明 1° 由 x^* 具有 n 位有效数字知绝对误差 $|e^*| \leq \frac{1}{2} \times 10^{m-n}$. 而相对误差

$$\begin{aligned} |e_r^*| &= \left| \frac{e^*}{x^*} \right| \leq \frac{1}{2|x^*|} \times 10^{m-n} \\ &\leq \frac{1}{2 \times 10^m \times 0.x_1} \times 10^{m-n} = \frac{1}{2x_1} \times 10^{1-n}. \end{aligned}$$

2° 绝对误差

$$\begin{aligned} |e^*| &= |e_r^*| \cdot |x^*| \leq \frac{1}{2(x_1+1)} \times 10^{1-n} \times 10^m \times 0.x_1x_2\cdots x_n \\ &\leq \frac{1}{2(x_1+1)} \times 10^{m+1-n} \times 0.(x_1+1) = \frac{1}{2} \times 10^{m-n}, \end{aligned}$$

由定义 1.3 知 x^* 至少具有 n 位有效数字. #

例 1.3 为使 $x = \sqrt{20}$ 的近似值 x^* 的相对误差不超过 $\frac{1}{2} \times 10^{-3}$, 问查开方表时至少要取几位有效数字?

解 设近似数 x^* 至少需保留 n 位有效数字可满足题设要求, 对于 $x = \sqrt{20}$, 有 $x_1 = 4$.

由定理 1.1 的第一个结论知, 此时有 x^* 的相对误差

$$|e_r^*| \leq \frac{1}{2x_1} \times 10^{1-n} = \frac{1}{8} \times 10^{1-n}.$$

令 $\frac{1}{8} \times 10^{1-n} \leq \frac{1}{2} \times 10^{-3}$, 解得 $n \geq 3.4$. 取 $n = 4$ 位有效数字. #

1.2.3 初值误差传播

近似数参加运算后所得之值一般也是近似值, 含有误差, 将这一现象称为误差传播. 数值运算中误差传播情况比较复杂, 主要表现在: 算法本身可能有