

○○○

建筑 给水排水及采暖工程 施工质量问答

○○○

宋波◎主编

中国建筑工业出版社

TU82/15

2008

建筑给水排水及采暖工程施工质量问答

宋 波 主编

中国建筑工程出版社

图书在版编目 (CIP) 数据

建筑给水排水及采暖工程施工质量问答/宋波主编.

北京: 中国建筑工业出版社, 2007

ISBN 978-7-112-09713-5

I. 建… II. 宋… III. ①建筑-给水工程-工程质量-问答②建筑-排水工程-工程质量-问答③房屋建筑设备: 采暖设备-建筑安装工程-工程质量-问答 IV. TU8-44

中国版本图书馆 CIP 数据核字 (2007) 第 159609 号

本书以问答的形式对涉及建筑给水排水及采暖工程施工质量的知识进行详解, 包括基本规定、室内给水排水系统、室外给水排水系统、卫生器具、建筑中水系统、游泳池水系统、采暖系统、锅炉等系统的安装和验收, 还针对相关质量验收的表格资料的填写作专门说明。本书设置的问题针对性和操作性强, 是给水排水及采暖工程从业人员必备参考资料。

* * *

责任编辑: 丁洪良

责任设计: 赵明霞

责任校对: 孟楠 陈晶晶

建筑给水排水及采暖工程施工质量问答

宋波 主编

*

中国建筑工业出版社出版、发行 (北京西郊百万庄)

各地新华书店、建筑书店经销

北京密云红光制版公司制版

北京市安泰印刷厂印刷

*

开本: 787×1092 毫米 1/16 印张: 17³/₄ 字数: 429 千字

2008 年 1 月第一版 2008 年 1 月第一次印刷

印数: 1-4000 册 定价: 30.00 元

ISBN 978-7-112-09713-5

(16377)

版权所有 翻印必究

如有印装质量问题, 可寄本社退换

(邮政编码 100037)

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理论、方法及有关应用,内容包括误差分析、线性方程组的直接解法与迭代解法、非线性方程的数值解法、插值法与曲线拟合、数值积分与数值微分、常微分方程的数值解法等.为了加强对学生基本知识的训练与综合能力的培养,每章末都配备了练习题,以便读者巩固、复习及应用所学知识.

本书可作为工科院校各专业硕士研究生“数值计算方法”课程的教材或教学参考书,也可供从事数值计算的科技工作者参考.

前 言

本书是为工科院校硕士研究生编写的教材.它是在原来所用教材的基础上,结合多年教学经验和科研实践修订而成的.本着重概念、重方法、重应用和重能力培养的原则,从构造算法、分析算法和使用算法三方面组织教材内容.在构造算法上,除阐明算法的构造思想和原理外,通过进一步的归纳和整理,我们尽量使同类算法都由某一基本原理或某一基本方法导出,以便读者易于领会和掌握同类算法的共同特征以及同类算法中不同方法之间的相异特征.在分析算法的有关理论推导中,我们力求深入浅出、通俗易懂,并补充少量基础知识,便于阅读和教学;在算法设计与理论分析中,对每种算法均十分关注其应用条件及使用中的问题.

每类算法都配以例题与习题,以助理解和练习.学习本书所需的数学基础是微积分、线性代数以及常微分方程的基本概念,可针对工科硕士研究生所要求的内容进行选材.本书中也包含一部分适合高水平学生深入理解的内容,可供选学.全书共六章,约需60~80学时,对不同专业,其具体内容和学时数可作适当增减.

本教程增加了目前工程和科学计算中的实际常用的数值计算方法及有关理论,如蒙特卡罗数值积分方法,大型稀疏线性方程组数值解,非线性方程组的数值解等.这些内容部分已超出目前出版的众多《数值计算方法》教材的内容.我们希望本教材出版有助于促进工科研究生《数值计算方法》课程水平并推动课程改革,希望有更多具有不同风格和特点的《数值计算方法》教材面世.

本书作者不仅长期从事本门学科的教学,而且具有长期从事科研项目计算的经历.这种实践形成了本书朴素、求实的风格.希望通过本书的介绍,使读者在较短的时间内比较顺利地掌握这些数值方法的要领和基本技巧,为今后从事科学计算打下牢固的基础.

李铁军老师编写第一、三章,肖辞源老师编写第二章,刘志斌、杨雁老师编写第四、五章,刘小华老师编写第六章,全书由李铁军老师统稿.限于水平,书中疏漏和缺陷之处难免,敬请读者批评指正.

西南石油大学应用数学教研室

目 录

1 绪论	1
1.1 现代科学技术研究的一般过程	1
1.1.1 工程问题数学化(数学建模)	1
1.1.2 数学问题数值化(算法与分析)	1
1.1.3 数值问题机器化(程序设计)	2
1.1.4 科学实验	2
1.2 数值计算研究的主要问题	3
1.2.1 线性和非线性方程组的数值解法	3
1.2.2 数值逼近	3
1.2.3 微分方程数值解	4
1.3 误差与数值计算的误差估计	4
1.3.1 误差的来源与分类	4
1.3.2 误差与有效数字	5
1.3.3 数值计算的误差估计	8
1.3.4 选用和设计算法时应遵循的原则	10
习题一	13
2 线性方程组的数值解法	15
2.1 消元法	16
2.1.1 三角形方程组的解	16
2.1.2 高斯消元法与列主元消元法	17
2.1.3 高斯-若当(Gauss-Jordan)消元法	21
2.2 直接分解法	21
2.2.1 多利特尔分解法	23
2.2.2 库朗分解	26
2.2.3 追赶法	27
2.2.4 对称矩阵的 LDL^T 分解	28
2.3 向量和矩阵的范数	30
2.3.1 向量范数	30
2.3.2 矩阵的范数	32
2.3.3 矩阵的条件数	33
2.3.4 误差分析	36
2.4 雅可比迭代	37

2.4.1	雅可比迭代格式	37
2.4.2	雅可比迭代的收敛性	38
2.5	高斯 - 塞德尔 (Gauss-Seidel) 迭代	41
2.5.1	高斯 - 塞德尔迭代格式	41
2.5.2	高斯 - 塞德尔迭代的收敛性	41
2.6	松弛迭代	43
2.6.1	松弛迭代格式	43
2.6.2	松弛迭代的收敛性	44
2.7	大型稀疏线性方程组数值解	44
2.7.1	大型稀疏矩阵的压缩存储	44
2.7.2	解大型稀疏线性方程组的共轭斜量法	47
	习题二	48
3	方程 (组) 的迭代解法	51
3.1	引言	51
3.2	迭代解法	51
3.2.1	根的初值确定方法	52
3.2.2	迭代法的求解过程	54
3.2.3	迭代法的收敛性	57
3.2.4	迭代序列的误差估计	58
3.3	迭代公式的改进	60
3.3.1	改变方程式法之一	61
3.3.2	改变方程式法之二	64
3.3.3	牛顿迭代法	65
3.3.4	弦截法	68
3.3.5	$ \varphi'(x) > 1$ 的处理方法	71
3.3.6	高阶迭代函数的构造方法	72
3.4	联立方程组的迭代解法	75
3.4.1	简单迭代法	75
3.4.2	关于一个收敛充分条件的证明	78
3.5	联立方程组的牛顿解法	79
	习题三	81
4	插值与拟合	83
4.1	插值与拟合问题	83
4.1.1	插值问题	83
4.1.2	拟合问题	84
4.2	拉格朗日插值	85

4.2.1 拉格朗日插值多项式	85
4.2.2 插值余项与误差估计	89
4.3 均差与牛顿插值	91
4.3.1 均差及其性质	91
4.3.2 牛顿插值	92
4.4 差分与等距节点插值	94
4.4.1 差分及其性质	94
4.4.2 等距节点插值	96
4.5 其他插值方法	99
4.5.1 埃尔米特插值	99
4.5.2 分段低次插值	102
4.5.3 三次样条插值	105
4.6 曲线拟合	112
4.6.1 曲线拟合与函数逼近	112
4.6.2 曲线拟合的最小二乘法	125
习题四	131
5 数值积分与数值微分	134
5.1 数值积分问题	134
5.1.1 数值求积的基本思想	134
5.1.2 代数精度	135
5.1.3 插值型求积公式	136
5.1.4 数值积分收敛性与稳定性	137
5.2 牛顿-柯特斯数值积分	138
5.2.1 柯特斯系数	138
5.2.2 偶阶求积公式的代数精度	139
5.2.3 牛顿-柯特斯求积公式的余项	140
5.3 复化求积公式	141
5.3.1 复化梯形公式	141
5.3.2 复化辛普森求积公式	142
5.4 龙贝格求积公式	144
5.4.1 梯形法的递推化	144
5.4.2 龙贝格算法	145
5.4.3 理查森外推加速法	147
5.5 高斯求积公式	150
5.5.1 勒让德多项式	150
5.5.2 高斯-勒让德求积公式	153
5.5.3 高斯-切比雪夫求积公式	155

5.6 蒙特卡罗数值积分	157
5.6.1 蒙特卡罗方法的基本思想及特点	157
5.6.2 蒙特卡罗方法伪随机数的产生和检验	158
5.6.3 蒙特卡罗随机抽样	161
5.6.4 随机变量分布函数的构造	163
5.7 数值微分	164
5.7.1 中点方法与误差分析	164
5.7.2 插值型的求导公式	165
5.7.3 利用数值积分求导	168
5.7.4 数值微分的外推算法	170
习题五	171
6 微分方程数值解法	174
6.1 引言	174
6.2 简单的数值方法	175
6.2.1 欧拉(Euler)法	175
6.2.2 后退欧拉法	176
6.2.3 梯形方法与改进的 Euler 公式	177
6.2.4 单步法的有关概念	180
6.3 显式龙格-库塔方法	183
6.3.1 显式龙格-库塔方法的一般形式	183
6.3.2 2 阶显式 R-K 方法	184
6.3.3 3 阶与 4 阶显式 R-K 方法	185
6.4 线性多步法	187
6.4.1 线性多步法的一般公式	187
6.4.2 基于数值积分的方法	188
6.4.3 基于 Taylor 展开的方法	189
6.4.4 预测-校正方法	193
6.5 一阶方程组和高阶方程	193
6.5.1 一阶方程组	193
6.5.2 化高阶方程为一阶方程组	195
6.6 边值问题的数值解法	195
习题六	198
参考文献	200

1 绪论

1.1 现代科学技术研究的一般过程

高科技的发展,促使计算机科学的发展突飞猛进.计算机的高速、大容量和多功能,又为现代科学技术的发展提供了最优、最快的新途径.一般而言,现代科学技术研究可按如下四个阶段进行.

1.1.1 工程问题数学化(数学建模)

采用恰当的数学语言,描述自然科学、社会科学及管理 and 决策科学各领域中关键而核心的问题,常称为数学建模.要建立一个好的数学模型,对于单方面的专家都是很困难的,必须由各相关领域的专家和数学工作者,特别是从事计算数学、应用数学研究工作的学者,紧密结合,相互取长补短才有可能.这是因为评价一个模型的优劣主要有两点:其一,用什么样的数学语言,才能真正反映工程实际;其二,所用的数学语言,可否在计算机中实现,这二者缺一不可.因此,要求参与建模工作的工程专家必须精通专业,并且不但要有一定的数学和计算数学的基础知识,还要掌握宽广的数学知识,清楚工程问题数学化面临的主要问题和采用哪种数学语言描述此问题更为恰当.工程中的数学问题,一般可分为三类:其一,连续型(确定性),即能用数学解析式刻画的工程问题;其二,离散型(统计型),找不到确定的数学解析式来描述的工程问题;其三,不确定型(随机型).本书重点讨论连续型.

1.1.2 数学问题数值化(算法与分析)

从工程实际中抽象出的数学问题,绝大多数不能直接用计算机来识别,因此,先进的计算工具——计算机,不能直接求解相应的数学问题,自然不能用于解决相应的工程问题.把数学问题数值化,就是如何根据不同的数学问题,寻求相应的方法.此方法(常称为数值计算方法)可以用四则运算和一些逻辑运算或者直接用

计算机语言来描述相应的数学问题,便于用计算机求解的数学问题,此方法的优劣,直接关系能否把计算机用于解决高科技问题.由此可见,数值计算在当代科学技术的地位与作用,它直接关系到能否用现代的数学方法、最先进的计算工具去解决现代科学技术中的管理问题、规划和决策问题以及各领域中高科技中的关键问题.因此对数值计算的算法构造及优劣的分析,是每一个科技工作者、决策者不可缺少的基础知识.对于即将走上工作岗位的硕士研究生而言则是必须掌握的.

1.1.3 数值问题机器化(程序设计)

要求程序设计者,用最简练的机器语言,最快的速度,最少的存储量来设计软件,并获得准确的计算结果.要达到这些结果,程序设计者必须掌握数值计算方法的构思途径,算法的关键与难点,熟悉计算机软硬件的基础知识,能灵活应用某种机器语言,准确无误的描述每一个算法,并能以最快的速度发现和解决计算过程中出现的各种异常问题,这是检验程序设计者水平高低的客观标准,也是衡量一个决策者、工程师或管理工作水平高低的重要标志.程序设计,是每一个科技工作者必须具备的技能,掌握这种技能入门快,见效也快,但要真正成为一个高级程序设计者,却是十分困难的,必须具备丰富的想象力以及总结归纳和设计的能力,具有总设计师、总工程师的能力与水平.

1.1.4 科学实验

前面三个阶段,只是为现代科学技术研究提供了一种途径,也可以说是捷径,但这种途径是否能真正解决科技中的问题,被实际生产部门直接采用,还必须将第三阶段获得的计算结果,在科学实验室里进行检验,看看是否与工程实际相符,能否推广应用,若不相符,分析不符的根源何在,确定返回前三阶段的某一阶段从新开始,重复上述的工作直到满意为止.这一步是必不可少的,且很成熟的,只要有相应的实验设备和原材料都可进行.前三阶段的实质是把一个工程问题置入计算机中,在计算机中和做大量的模拟实验,当基本上与实际问题相符时,再经过此步,这样可以减少实验次数,节省大量的原材料,缩短设计周期,还能使性能达到最佳,是现代科学的必经途径.

任何一个有竞争力的新产品,任何一项能产生社会和经济效益的科研课题,必须经过上述的四个过程.第一阶段是根本,模型是否反映工程实际问题的需要,取决于当代高级科技人员的理论水平;第二阶段是桥梁,所构造的算法是否能取代原数学模型,是否具有有效性和可实现性,取决于从事计算数学的研究人员的理论水平和创新能力;第三阶段是检验前两阶段的有效性和可行性的方法和手

段;第四阶段则是检验该项研究成果是否有实际应用价值,同时也检验了前三阶段研究工作的可靠性,也是能否见到经济和社会效益的关键.

1.2 数值计算研究的主要问题

数学来源于实践,从工程实际中抽象出来的数学问题,大体上分为如下三类:(1)线性与非线性方程组的求解;(2)逼近函数的构造方法(数值逼近);(3)微分方程数值解.除此之外,还有离散型和不确定(随机)型.在高科技中,有许多的工程问题很难用现有的数学语言来刻画,必须从工程实际出发,探索新的数学语言,构造相应的数值求解方法,不断产生新的数学分支,这需要众多的工程专家、数学工作者紧密结合,深入到科研的前沿,去发现、解决新问题,从而促进科学技术高速发展.在本书中,只对上述三类问题探索数值求解方法及其与算法有关的理论进行分析,为了使读者有一个宏观的印象,先对上述三类问题作如下简介.

1.2.1 线性和非线性方程组的数值解法

此类问题是后两类问题的基础,一般表示法如下:

$$f_j(x_1, x_2, \dots, x_n) = b_j, j = 1, 2, \dots, m.$$

在一般情况下: $n \geq m$.

记 $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$, $\mathbf{b} = (b_1, b_2, \dots, b_m)^T$, $\mathbf{F}(\mathbf{x}) = [f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_m(\mathbf{x})]^T$. 上述方程组可简记为 $\mathbf{F}(\mathbf{x}) = \mathbf{b}$.

当 $f_j(\mathbf{x})$ ($j = 1, 2, \dots, m$) 均为线性函数时,称 $\mathbf{F}(\mathbf{x}) = \mathbf{b}$ 为线性方程组: $\mathbf{Ax} = \mathbf{b}$, 其中 $\mathbf{A} = (a_{ij})_{m \times n}$, 常用的方法是直接法与迭代解法,对于大型稀疏矩阵 $\mathbf{A}$ 常采用不完全共轭梯度法,详见本书第2章.

当 $f_j(\mathbf{x})$ ($j = 1, 2, \dots, m$) 为非线性函数时,称 $\mathbf{F}(\mathbf{x}) = \mathbf{b}$ 为非线性方程组,惯用的方法是迭代法,详见本书第3章.

1.2.2 数值逼近

在科学实验、新产品性能分析和大规模的科学计算中,可获得大量的离散数据,如何通过这些数据获得原物体的各种性能指标,惯用的方法是将离散数据解析化,即构造一个或多个函数,取代离散数据,此函数常称为逼近函数.构造逼近函数的常规方法的难点在于如何根据离散数据的分布特点,选择相应的逼近函数所在的函数类;当逼近函数所在的函数类确定后,如何评价逼近程度、逼近的优劣,从而获得相应的逼近方法.根据不同的目的与需要又将逼近方法分为插值问题与拟合问题,本书主要介绍了拉格朗日插值、牛顿插值、埃尔米特插值、分段插

值、样条插值和曲线拟合等,数值积分与数值微分,详见本书第3章、第4章和第5章.

1.2.3 微分方程数值解

在最优控制、机械工程及流体工程中将许多问题抽象出的数学问题归结为求常微分方程(组)的初值问题:

$$\begin{cases} f[x, y(x), y'(x), \dots, y^{(m)}(x)] = 0, \\ y^{(k)}(x_0) = y_k \quad (k = 0, 1, \dots, m-1). \end{cases}$$

初值条件也可换成两点边值问题,如:

$$\begin{aligned} y^{(k)}(a) &= y_k \quad (k = 0, 1, \dots, l-1), \\ y^{(m-k)}(b) &= y_k \quad [k = m, m-1, \dots, m-(l-1)]. \end{aligned}$$

上述微分方程要求精确的理论解,一般情况下是十分困难的. 这些问题解的存在性、唯一性都很难确定,并且近代科学技术中,抽象出的问题比上述问题更为复杂,因此只能求解近似解,目前常用的方法是差分法,详见本书第6章.

1.3 误差与数值计算的误差估计

1.3.1 误差的来源与分类

在数值计算过程中,估计计算结果的精确度是十分重要的工作,而影响精确度的因素是各种各样的误差,它们可分为两大类:一类称为“过失误差”,它一般是由人为造成并可以避免的,故在“数值计算”中我们不讨论它;而另一类称为“非过失误差”,这在“数值计算”中往往是无法避免的,也是我们将要研究的对象. 按照它们的来源,误差可分为以下4种:

(1) 模型误差

用数值计算方法解决实际问题时,首先必须建立数学模型. 由于实际问题的复杂性,在对实际问题进行抽象与简化时,往往为了抓住主要因素而忽略了一些次要因素,这样就会使得建立起来的数学模型只是复杂客观现象的一种近似描述,它与实际问题之间总会存在一定的误差. 我们把数学模型与实际问题之间出现的这种误差称为模型误差.

(2) 观测误差

数学模型中往往包含一些由观测或实验带来的物理量,由于测量工具精度和测量手段的限制,它们与实际大小之间必然存在误差,这种误差称为观测误差.

(3) 截断误差

由实际问题建立起来的数学模型,在很多情况下要得到准确解是困难的,通

常用数值方法求出它的近似解. 例如常用有限过程逼近无限过程, 用能计算的问题代替不能计算的问题. 这种数学模型的精确解与由数值方法求出的近似解之间的误差称为截断误差, 由于截断误差是数值计算方法固有的, 故又称为方法误差. 例如用函数 $f(x)$ 的泰勒 (Taylor) 展开式的部分和 $S_n(x)$ 去近似代替 $f(x)$, 其余项 R_n 就是真值 $f(x)$ 的截断误差.

(4) 舍入误差

用计算机进行数值计算时, 由于计算机的位数有限, 计算时只能对超过位数的数字进行四舍五入, 由此产生的误差称为舍入误差. 例如用 2.718 28 作为无理数 e 的近似值产生的误差就是舍入误差. 应请读者注意的是, 虽然少量的舍入误差是微不足道的, 但在计算机上完成了千百万次运算之后, 舍入误差的积累却可能是十分惊人的.

综上所述, 数值计算中除了可以完全避免的过失误差外, 还存在难以回避的模型误差、观测误差、截断误差和舍入误差. 在这四种误差来源的分析中, 前两种误差是客观存在的, 后两种误差是由计算方法所引起的. 因此, 后两种误差将是数值计算方法的主要研究对象, 讨论它们在计算过程中的传播和对计算结果的影响, 并找出误差的界, 对研究误差的渐近特性和改进算法的近似程度具有重大的实际意义.

1.3.2 误差与有效数字

(1) 绝对误差与绝对误差限

设某一量的精确值为 x , 其近似值为 x^* , $E(x) = x - x^*$ 为近似值 x^* 的绝对误差, 简称误差.

当 $E(x) > 0$ 时, 称 x^* 为弱近似值或亏近似值; 当 $E(x) < 0$ 时, 称 x^* 为强近似值或盈近似值.

一般而言, 某一量的精确值 x 是不知道的, 因而 $E(x)$ 也无法求出, 但往往可以估计出 $E(x)$ 的上界, 即存在 $\varepsilon > 0$, 使得 $|E(x)| = |x - x^*| \leq \varepsilon$, 此时, 称 ε 为近似值 x^* 的绝对误差限, 简称误差限或精度. ε 越小, 表示近似值 x^* 的精度越高. 显然有 $x^* - \varepsilon \leq x \leq x^* + \varepsilon$. 有时也用 $x = x^* \pm \varepsilon$ 来表示近似值 x^* 的精度或精确值 x 的所在范围. 绝对误差是有量纲的. 例如, 用毫米刻度的卷尺去测量一长度为 x 的物体, 测得其近似值为 $x^* = 84 \text{ mm}$, 由于卷尺以毫米为刻度, 所以其误差不超过 0.5 mm , 即 $|x - 84| \leq 0.5 \text{ mm}$. 这样, 虽然不能得出准确值 x 的长度是多少, 但从这个不等式可以知道 x 的范围是 $83.5 \text{ mm} \leq x \leq 84.5 \text{ mm}$. x 必在 $[83.5 \text{ mm}, 84.5 \text{ mm}]$ 内.

(2) 相对误差与相对误差限

用绝对误差来刻画某个近似值的精确程度是有局限性的,在很多场合中它无法显示出近似值的准确程度.如测量 10 cm 和 10 m 的两个长度,若它们的绝对误差都是 1 cm,显然前者的测量结果比后者的准确.由此可见,决定一个量的近似值的精确度,除了要看绝对误差的大小外,还必须考虑该量本身的大小,为此引入相对误差的概念.

绝对误差与精确值之比,即称 $E_r(x) = \frac{E(x)}{x} = \frac{x - x^*}{x}$ 为近似值 x 的相对误差.在实际中,由于精确值 x 一般无法知道,因此往往取 $E_r^*(x) = \frac{x - x^*}{x^*}$ 作为近似值 x^* 的相对误差.

类似于绝对误差的情况,若存在 $\delta > 0$,使得 $|E_r^*(x)| = |\frac{x - x^*}{x^*}| \leq \delta$,则称 δ 为近似值 x^* 的相对误差限.相对误差是无量纲的数,通常用百分比表示,称为百分误差.

根据上述定义可知,当 $|x - x^*| < 1 \text{ cm}$ 时,测量 100 m 物体时的相对误差 $|E_r(x)| \leq \frac{1}{10\,000} = 0.01\%$,测量 10 m 物体时的相对误差 $|E_r(x)| \leq \frac{1}{1\,000} = 0.1\%$.可见前者的测量结果要比后者精确.所以,在分析误差时,相对误差更能刻画误差的特性.

(3)有效数字

为了能给出一种数的表示法,使之既能表示其大小,又能表示其精确程度,于是需要引进有效数字的概念.在实际计算中,当准确值 x 有很多位数时,我们常按四舍五入的原则得到 x 的近似值 x^* .例如无理数 $e = 2.718\,281\,828\cdots$,按四舍五入原则分别取三位和六位小数时,则得 $e \approx 2.72$ 和 $e \approx 2.718\,28$.

不管取几位小数得到的近似值,其绝对误差都不超过末位数的半个单位,即

$$|e - 2.72| \leq \frac{1}{2} \times 10^{-2}, |e - 2.718\,28| \leq \frac{1}{2} \times 10^{-5}.$$

下面我们将四舍五入抽象成数学语言,进而引进“有效数字”的概念.

若近似值 x^* 的绝对误差限是某一位的半个单位,就称其“准确”到这一位,且若从该位直到 x^* 的第一位非零数字共有 n 位,则称近似值 x^* 有 n 位有效数字.

引入有效数字概念后,我们规定所写出的数字都应该是有效数字,且在同一计算问题中,参加运算的数,都应该有相同位数的有效数字.

例如,下列各数 358.467, 0.004 275 11, 8.000 034, $8.000\,034 \times 10^3$ 的具有 5 位有效数字的近似值分别是 358.47, 0.004 275 1, 8.000 0, 8 000.0.

注意:8.000 034 的 5 位有效数字的近似值是 8.000 0,而不是 8,因为 8 只有 1 位有效数字.前者精确到 0.000 1,而后者仅精确到 1,两者相差是很大的.显然,前者远较后者精确,由此可见,有效数字尾部的零不能随意省去,以免损失精度.

一般而言,任何一个实数 x 经四舍五入后得到的近似值 x^* 都可写成如下标准形式

$$x^* = \pm (a_1 \times 10^{-1} + a_2 \times 10^{-2} + \cdots + a_n \times 10^{-n}) \times 10^m, \quad (1.1)$$

所以,当其绝对误差限满足

$$|x - x^*| \leq \frac{1}{2} \times 10^{m-n} \quad (1.2)$$

时,则称近似值 x^* 具有 n 位有效数字,其中 m 为整数, a_1 是 1~9 中的一个数字, $a_2, a_3, \cdots, a_n$ 是从 0~9 中的数字.

根据上述有效数字的定义,不难验证 e 的近似值 2.718 28 具有 6 位有效数字.事实上,

$$2.718\ 28 = (2 \times 10^{-1} + 7 \times 10^{-2} + 1 \times 10^{-3} + 8 \times 10^{-4} + 2 \times 10^{-5} + 8 \times 10^{-6}) \times 10.$$

这里 $m=1, n=6$. 因为 $|e - 2.718\ 28| = 0.000\ 001\ 828 < \frac{1}{2} \times 10^{-5}$, 所以它具有 6 位有效数字.

有效数字不但给出了近似值的大小,而且还给出了它的绝对误差限.例如有效数字 3567.82, 0.423×10^{-2} , 0.4230×10^{-3} 的绝对误差限分别为 $\frac{1}{2} \times 10^{-2}$, $\frac{1}{2} \times 10^{-5}$, $\frac{1}{2} \times 10^{-6}$. 必须注意,在有效数字的记法中, 0.423×10^{-2} 与 $0.423\ 0 \times 10^{-3}$ 是有区别的,前者只有 3 位有效数字,而后者则具有 4 位有效数字.

还需指出的是,一个准确数值的有效数字的位数,应当说有无穷多位,例如 $1/8 = 0.125$ 不能说只有 3 位有效数字.

有效数字与绝对误差、相对误差有如下关系:

(1) 若某数 x 的近似值 x^* 有 n 位有效数字,则此近似值 x^* 的绝对误差限为 $|x - x^*| \leq \frac{1}{2} \times 10^{m-n}$. 由此可见,当 m 一定时,有效数字位数 n 越多,其绝对误差限越小.

(2) 若用 (1.1) 式表示的近似值 x^* 具有 n 位有效数字,则其相对误差限为

$$|E_r^*(x)| \leq \frac{1}{2a_1} \times 10^{-(n-1)}. \quad (1.3)$$

反之,若 x^* 的相对误差限满足:

$$|E_r^*(x)| \leq \frac{1}{2(a_1 + 1)} \times 10^{-(n-1)}, \quad (1.4)$$

则 x^* 至少具有 n 位有效数字.

事实上, 由 (1.1) 式可知 $a_1 \times 10^{m-1} \leq |x^*| \leq (a_1 + 1) \times 10^{m-1}$, 所以

$$|E_r^*(x)| = \left| \frac{x - x^*}{x^*} \right| \leq \frac{0.5 \times 10^{m-n}}{a_1 \times 10^{m-1}} = \frac{1}{2a_1} \times 10^{-(n-1)}.$$

反之, 由 $|x - x^*| = |x^*| |E_r^*(x)|$

$$\leq (a_1 + 1) \times 10^{m-1} \times \frac{1}{2(a_1 + 1)} \times 10^{-(n-1)} = \frac{1}{2} \times 10^{m-n}$$

知, x^* 至少具有 n 位有效数字. 可见有效数字的位数越多, 相对误差限就越小, 即近似值的有效位数越多, 用这个近似值去代替准确值的精度就越高.

例 1.1 为使 $\sqrt{20}$ 的近似值的相对误差小于 1%, 问至少应取几位有效数字?

解 $\sqrt{20}$ 的近似值的首位非零数字是 $a_1 = 4$, 由 (1.3) 式有

$$|E_r^*(x)| = \frac{1}{2 \times 4} \times 10^{-(n-1)} < 1\%,$$

解之得 $n > 2$, 故取 $n = 3$ 即可满足要求. 也就是说只要 $\sqrt{20}$ 的近似值具有 3 位有效数字, 就能保证 $\sqrt{20} \approx 4.47$ 的相对误差小于 1%.

1.3.3 数值计算的误差估计

数值计算中误差产生与传播的情况非常复杂, 参与运算的数据往往都是些带有误差的近似数, 这些数据的误差在多次运算中会进行传播, 使计算结果又产生一定的误差, 这就是误差的传播问题. 以下介绍利用函数的 Taylor 公式来估计误差的一种常用方法.

设可微函数 $y = f(x_1, x_2, \dots, x_n)$ 中的自变量 $x_1, x_2, \dots, x_n$ 相互独立, 且 $x_1^*, x_2^*, \dots, x_n^*$ 依次是 $x_1, x_2, \dots, x_n$ 的近似值, 则 y 的近似值 $y^* = f(x_1^*, x_2^*, \dots, x_n^*)$.

$$\begin{aligned} E(y^*) &= f(x_1, x_2, \dots, x_n) - f(x_1^*, x_2^*, \dots, x_n^*) \\ &\approx \sum_{i=1}^n \frac{\partial f(x_1^*, x_2^*, \dots, x_n^*)}{\partial x_i} (x_i - x_i^*) \\ &= \sum_{i=1}^n \frac{\partial f(x_1^*, x_2^*, \dots, x_n^*)}{\partial x_i} E(x_i^*), \end{aligned} \quad (1.5)$$

$$\begin{aligned} E_r(y^*) &= \frac{E(y^*)}{y^*} \approx \sum_{i=1}^n \frac{\partial f(x_1^*, x_2^*, \dots, x_n^*)}{\partial x_i} \frac{x_i^*}{y^*} \frac{E(x_i^*)}{x_i^*} \\ &= \sum_{i=1}^n \frac{x_i^*}{y^*} \frac{\partial f(x_1^*, x_2^*, \dots, x_n^*)}{\partial x_i} E_r(x_i^*). \end{aligned} \quad (1.6)$$