“十二五”普通高等教育规划教材

数值计算方法

（第2版）

令 锋 傅守忠 陈树敏 曲良辉 编

Numerical Computation Methods

国防工业出版社
National Defense Industry Press

"十二五"普通高等教育规划教材

数值计算方法

(第2版)

令锋 傅守忠 陈树敏 曲良辉 编

国防工业出版社
·北京·

内容简介

本书阐述数值计算的基本理论和常用方法，包括误差分析与算法设计，非线性方程的数值解法，线性方程组的直接法与迭代法，插值法与最小二乘拟合法，数值积分与数值微分，常微分方程的数值解法，矩阵特征值与特征向量的计算等；并在附录中介绍了数值实验报告的基本格式、Matlab 软件的基本使用方法和 C 语言文件操作方法. 本书建议学时为 54 学时，其中含数值实验 8 学时. 书中含有较丰富的例题、习题和数值实验题，给出了典型算法的伪代码描述及 Matlab 软件提供的相应函数，并编写出版了与本书配套的复习与实验指导教材.

本书以实际应用为目的，选材恰当，体系完整，强调数值算法的设计方法和编程实现技能，可作为普通本科院校信息与计算科学、数学与应用数学、统计学、软件工程、计算机科学与技术等专业本科生学习数值分析或计算方法课程的教材，也可作为其他理工科专业本科生和研究生的参考教材.

图书在版编目(CIP)数据

数值计算方法/令锋等编.—2 版.—北京：国防工业出版社，2015.3

“十二五”普通高等教育规划教材

ISBN 978-7-118-09954-6

Ⅰ.①数...　Ⅱ.①令...　Ⅲ.①数值计算—计算方法—高等学校—教材　Ⅳ.①O241

中国版本图书馆 CIP 数据核字(2015)第 046622 号

※

国防工业出版社出版发行

(北京市海淀区紫竹院南路 23 号　邮政编码 100048)

北京奥鑫印刷厂印刷

新华书店经售

*

开本 787×1092　1/16　**印张** 10½　**字数** 234 千字

2015 年 3 月第 2 版第 1 次印刷　**印数** 1—3000 册　**定价** 25.00 元

国防书店：(010)88540777　　发行邮购：(010)88540776

发行传真：(010)88540755　　发行业务：(010)88540717

第2版前言

编写本书旨在为以培养应用型和技能型人才为目标的普通本科院校理工科学生学习计算方法或数值分析课程提供一本简明、实用、易学和易教的教材. 在教材出版的三年时间里,编者陆续收到了来自不同高校的一些读者提出的宝贵意见和建议,值此再版之际,我们谨向帮助和支持我们提高教材质量的师生表示深深的谢意.

此次修订保持了第1版的内容和特点,强调数值算法的设计方法和编程实现算法的技能,力求易学和易教. 修订的主要内容如下:

(1) 纠正了发现的编写错漏、印刷错误和叙述不严谨之处.

(2) 对个别章节的内容作了适当删改,力求更加通俗和简明.

(3) 为进一步提高数值实验教学的效果,对各章的实验题目作了适当修改和补充,并给出了验证性实验的源程序.

(4) 鉴于文件操作方法在编程中的重要作用和应用的广泛性,为引导学生在程序设计中采用文件读写技术,在附录中增加了C/C++语言文件操作简介.

尽管我们在修订中作了很大努力,但限于编者的水平,书中可能仍有谬误之处,敬请使用本书的师生和广大读者朋友批评指正.

编 者

2014年12月

第1版前言

科学计算研究以计算机为工具求解各种科学与工程计算问题的数值方法及其软件实现，并对求得的解的精度进行分析评估. 自从Galielo（伽利略）系统地引进实验方法、Newton（牛顿）奠定物理学的理论基础以来，近代科学研究的方法一直分为实验测定和理论分析两个同等重要的方面. 随着电子计算机软硬件技术的发展以及伴随的计算方法的不断改进，科学计算已经上升为一种与传统的实验测定及理论分析方法鼎立的第三种科学研究方法，三者各有其适宜的应用范围，相互促进，相互补充，相得益彰.

实验测定直接观测现实世界各种物理对象或模型的约束关系，进而总结出支配各种物理现象的物理定律. 一个物理过程的最基本、最可靠的资料无疑应由实验方法获得. 实验数据真实可信，是理论分析和科学计算的基础，但实验往往受到测量工具、环境和模型尺寸等的限制，有些条件下甚至不可能实现. 此外，实验一般耗资巨大，周期较长. 理论分析以根据物理过程和实验结果建立的数学模型为基础，研究物理现象或过程的内在规律，即求出模型的解析解. 理论分析的优点在于所得结果具有普遍性，各种影响因素清晰可见，可为进行实验研究提供理论指导，为构造新的数值算法提供理论基础，并为检验算法的精确度提供比较依据，但理论分析的对象在物理性质上需要简化，在几何形式上一般要求规则，只有高度简化或理想化的问题才能获得解析解. 对非线性问题，大部分都无法求得解析解. 有些问题虽然可求解，但解析解中含有无穷级数和特殊函数等，不能满足实际需要. 科学计算以计算机为工具，以计算数学理论为基础，通过对描述自然现象和物理过程的数学模型进行数值求解和图像显示来研究科学与工程问题的规律. 对一个确定的问题，在建立的数学模型合理且采用的求解方法有效的前提下，科学计算具有其他方法无可比拟的优势：可以模拟真实条件甚至理想条件下的状态；成本低；速度快；获得的资料完整. 在进行实验研究和理论分析时，通过科学计算可以补充资料. 在特定参数下进行一次数值计算就相当于一次实验，由此可以解决实验研究与理论分析未解决的问题，并发现一些新的物理现象. 在基本规律比较成熟的应用领域，无需实验，科学计算将独自承担起应用问题的求解，在那些理论模型复杂、实验费用昂贵甚至不可能进行实验的高新科技领域，科学计算是解决问题的主要甚至唯一的手段. 在独创性工作的先行性研究过程中，科学计算更是探索新的物理现象和规律的主要手段，历史上曾有首先通过科学计算发现新现象而后由实验证实的事件，海王星以及奇异吸引子的发现都是例证.

在科学的计算机化进程中，科学计算作为工具性、方法性、边缘交叉性的学科得到了前所未有的发展，在物理学、力学、生物学、生态学、计算机应用、航空航天、土木

工程、机械工程、地质勘探、水利水文、环境工程、医疗卫生、风险投资和经济管理等领域都有广泛的应用.科学计算与具体学科的交叉形成了许多计算性科学分支,诸如计算物理、计算力学、计算传热学、计算化学、计算生物学等,这些科学分支以数值计算方法作为基础和联系纽带,使得计算数学这一古老的数学学科成为现代科学与工程领域中一个生机勃勃的分支,为理论与实际的联系架起了桥梁.

作为介绍科学计算的基础理论与基本方法的课程,数值分析或计算方法已成为理工科许多专业本科生和研究生的专业基础课,国内外也已出版了较多的教材,然而,这些教材一般都偏重于理论证明和简单的手工跟踪算法实践,对算法的编程实现及如何进行数值实验则强调相对较少.这类教材适合研究型院校采用,对以培养应用型人才为目标的普通本科院校学生,则不完全适合.本书主要针对以培养应用型人才为目标的普通本科院校的教学需要编写,更加强调数值算法的设计和编程实现技能,力图做到简单、易教和易学,注重培养学生运用所学知识分析问题解决问题的能力.本书有如下特色.

(1) 精讲最重要的理论和常用方法,为求解较复杂的问题奠定基础.

(2) 给出典型算法的伪代码描述,以便读者选用 Excel、C 语言或 Matlab 等自己熟悉的程序设计语言实现算法.

(3) 提供基本术语和名词的英语表达.

(4) 编写出版了与本书配套的复习与实验指导书,在配套教材中补充了一定数量的典型例题,给出了本书中习题的答案,其中对部分习题做了详细解答,介绍了典型算法的实现方法,并给出了五套模拟试题及参考答案,帮助学生复习和巩固所学知识.

(5) 介绍了数值实验报告的基本格式和 Matlab 软件使用的基本方法.

学习本教材要求学生具有微积分和线性代数知识,还要求初步掌握常微分方程等基础知识,并至少掌握一种程序设计语言.书中选配的打星号(*)的习题主要供数学类专业的学生选用.

在本书编写过程中编者广泛参阅了国内外的相关教材和资料,在此谨向作者致以诚挚的感谢.本书的出版得到了肇庆学院优秀课程教学团队建设基金资助,编者在此表示衷心感谢.

限于编者的学识水平,书中难免有疏漏与欠妥之处,恳请使用本书的师生和广大读者批评指正.

编者

2011 年 10 月

目　　录

第 1 章　数值计算方法概论

本章简要介绍数值计算方法的研究对象、基本内容和特点，讨论误差的基本理论、数值算法的稳定性以及数值算法设计的原则.

1.1　数值计算方法的基本内容与特点

数值计算方法也称为数值分析或计算方法，是研究使用计算机求解各种数学问题的方法、理论及其软件实现，并对求得的结果进行分析的一个数学分支. 这里所说的数学问题即给出一组数值型数据（已知条件），根据问题所满足的性质去求另一组数值型数据（求解结果），如函数值的计算、方程组的求解等. 使用计算机进行求解时，通常需要通过离散化、逼近、插值、迭代等方法才能求得结果. 通过有限次的四则运算在计算机上求解数学问题的方法通常称为**数值方法**（Numerical Method），也称为**科学计算**（Scientific Computation），求得的**近似解**（Approximate Solution）通常称为**数值解**（Numerical Solution）.

数值计算方法以各类数学问题的数值方法为研究对象，构造求解科学与工程领域的各种数学问题的数值算法，研究算法的数学机理，对求得或将要求得的解的精度进行分析评估，通过编程和上机实现算法求得结果，分析数值结果的误差，并与相应的理论结果和可能的实验数据对比验证.

应用计算机解决实际问题的基本过程如下：

实际问题→数学模型→数值算法→程序设计→计算结果→分析结果

由实际问题应用有关科学知识和数学理论建立数学模型这一过程通常被认为是应用数学的任务，而根据数学模型提出数值算法、进行程序设计、上机计算结果和分析结果这些工作则属于数值计算方法的研究内容 .

数值计算方法是一门与计算机使用密切结合的实用性很强的课程，它既有纯数学的高度抽象性与严密科学性的特点，又有实际实验的高度技术性和应用广泛性的特点. 由于计算机的认识能力有限，人们设计的算法必须使得计算机能够接受. 例如，如果考虑用 C 语言计算一个定积分，则需要预先将求定积分的问题转化为初等运算和初等函数构成的计算问题（本书第 6 章中将介绍相应的方法），因为 C 语言中并不能直接接受“定积分”这个数学概念；由于计算机的计算能力有限，人们设计的算法应该具有较高的效率. 例如，线性代数中已经介绍了线性方程组解的存在唯一性理论和 Cramer 法则等精确解法，但用这些理论和方法还不能直接在计算机上求解线性方程组，因为，用 Cramer 法则求解一个 n 阶线性方程组要计算 $n+1$ 个行列式的值，总共需要 $(n-1)(n+1)n!$ 次乘法，当 n 充分大时，计算量将相当惊人，如对于一个仅为 20 阶的线性方程组，就要做大约 10^{21} 次乘法，即使使用每秒达百亿次的计算机去做，这项计算也需要连续工作几千年才能完成，这

自然是没有意义和价值的计算，而采用消元法（本书第 3 章中将介绍相应的方法）求解一个 n 阶线性方程组，则大约需要 $n^3/3+n^2$ 次乘法，对于 20 阶的线性方程组，使用普通微型计算机编程，可在几秒内求出答案.

在求解科学与工程计算问题时，不同的算法适合不同的问题，需要根据问题的性质设计有针对性的算法. 此外，还应该根据问题的特点，研究适合在计算机上使用、满足精度要求和节省计算时间的有效算法，在实现算法时还应该根据计算机容量、字长、速度等指标，研究具体求解步骤和程序设计技巧. 有些方法在理论上可能不够严谨，但通过实际计算和对比分析等手段，已被验证是行之有效的方法，也应该采用（如例 1－7）.

数值计算方法具有如下特点：

（1）提供面向计算机、理论可靠、计算复杂性好的数值算法. 数值分析的核心内容是研究应用计算机求解数学问题的各种数值计算方法，并对每个算法进行相关的理论分析. 对近似算法要保证收敛性和数值稳定性，并对误差进行分析，对逼近问题要保证达到要求的精度. 此外还必须保证提供的算法在计算机上切实可行，这包括要求算法有好的时间复杂性和空间复杂性.

（2）强调完成从理论到实践的全过程. 数值分析是一门实践性很强的数学课程，每个算法除了理论上要正确可行外，还要通过数值试验证明是行之有效的. 读者学完每个算法后都应该以解决实际问题为目的，通过编程或借助成熟的数学软件完成数值计算的训练，不仅要学会“怎样算”，而且必须做到“真会算”，即不仅要知道问题的解是存在的，还必须求出具体的结果.

（3）计算公式冗长且难以熟记. 数值计算方法处理问题主要采用如下方法：①“构造性“方法，许多问题的存在性证明都通过把问题的计算公式具体构造出来完成，不但证明了问题的存在性，同时还提供了具体的计算公式；②“离散化”方法，把求解连续变量的数学问题转化为求解离散变量的问题. 如把常微分方程离散成为差分方程等；③“递推化”方法，将一个复杂的计算过程归结为简单计算过程的多次重复，以便编写计算机程序计算；④“近似替代”方法，由于计算机必须在有限次运算后停止，所以数值方法常表现为一个无穷过程的截断，把一个无限过程的数学问题，转化为满足一定精度要求的有限步运算来近似替代. 这些基本特点使得数值分析课程中出现的计算公式多且繁杂，不易熟记.

根据上述特点，读者在学习过程中：第一，要注意复习微积分、线性代数、常微分方程和高级语言程序设计方法等课程的内容，这是学习本课程的基础；第二，要理解方法的基本原理和思想，掌握方法处理的技巧，并注意与计算机的结合；第三，要独立完成一定数量的习题，复习和巩固所学内容；第四，一定要认真做好数值实验，通过编程与调试实现算法，培养和提高自己分析问题和解决问题的能力.

1.2 误差的基本理论

数值计算通常是近似计算，实际结果与理论结果之间存在误差. 误差按照来源可分为 4 类：模型误差、观测误差、截断误差和舍入误差. 数值计算方法的任务之一就是对误差进行估计和控制.

1.2.1 误差来源

1. 模型误差

如 1.1 节所述,用计算机解决数值计算问题首先要建立数学模型,建立的模型仅仅是对实际问题进行抽象和理想化后的近似描述,因此,不可避免地产生误差,这种数学模型与实际问题之间出现的误差称为**模型误差**(Model Error).

2. 观测误差

针对实际问题建立的数学模型中通常包含若干物理参数,如温度、长度、电压、速度等,这些参数通常是通过观测和实验得来的,不可避免地带有误差,这种误差称为**观测误差**(Observation Error).

3. 截断误差

根据实际问题建立的数学模型通常比较复杂,很多情况下无法获得精确解,只能用数值方法求其近似解. 数学模型的精确解与数值方法的近似解之间的误差称为**截断误差**(Truncation Error). 由于截断误差是方法固有的,所以也称为方法误差(Method Error). 例如,指数函数 $f(x)=e^x$ 可展开为幂级数形式

$$e^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{3!}x^3 + \cdots + \frac{1}{n!}x^n + \cdots$$

使用计算机求值时只能取有限项作为 e^x 的近似值

$$S_n(x) = 1 + x + \frac{1}{2}x^2 + \frac{1}{3!}x^3 + \cdots + \frac{1}{n!}x^n$$

根据 Taylor 余项定理,部分和 $S_n(x)$ 作为 e^x 的近似值的余项为

$$R_n(x) = e^x - S_n(x) = \frac{x^{n+1}}{(n+1)!}e^{\xi}$$

其中 ξ 为 0 与 x 之间的数.

上例中截取无穷级数的有限项之和作为无穷级数的和的近似值时所产生的误差就是截断误差.

4. 舍入误差

由于计算机的字长有限,原始数据以及计算过程中的数据在计算机上都只能按照一定的舍入规则保留有限位,由此产生的误差称为**舍入误差**(Round - off Error).

数值计算方法中总是假定数学模型是准确的,因而不考虑模型误差和观测误差,主要研究截断误差和舍入误差对计算结果的影响.

1.2.2 绝对误差与相对误差

1. 绝对误差

定义 1-1 设 x 为准确值,x^* 为 x 的一个近似值,称

$$E_a(x) = x^* - x$$

为近似值 x^* 的**绝对误差**(Absolute Error),简称为**误差**(Error).

由定义 1-1 可以看出,误差 $E_a(x)$ 可正可负.

通常无法得到准确值 x,因而不能算出 x^* 的绝对误差 $E_a(x)$ 的准确值,只能根据测量工具或计算情况估计出误差的绝对值的一个上界,即可求出一个正数 ε_a,使得

$$|E_a(x)| = |x^* - x| \leqslant \varepsilon_a$$

正数 ε_a 称为近似值 x^* 的**绝对误差限**(Absolute Error Bound).

有了绝对误差限,就可知道准确值 x 的范围

$$x^* - \varepsilon_a \leqslant x \leqslant x^* + \varepsilon_a$$

或表示为

$$x = x^* \pm \varepsilon_a$$

绝对误差的大小在许多情况下还不能完全刻画一个近似值的精确度. 例如,测量100m 跑道的长度和测量一个人的身高,若它们的绝对误差都是 1cm,显然前者的测量精确度较后者好得多. 这表明描述一个量的近似值的精确度时,不仅要考虑绝对误差的大小,还要考虑该量本身的大小. 为此,需要引入相对误差的概念.

定义 1-2 设 x 为准确值,x^* 为 x 的一个近似值,称

$$E_r(x) = \frac{E_a(x)}{x} = \frac{x^* - x}{x} \quad (x \neq 0)$$

为近似值 x^* 的**相对误差**(Relative Error).

在实际计算中,由于准确值 x 一般是未知的,通常用 x^* 代替相对误差 $E_r(x)$ 中的分母 x,由此得近似值 x^* 的相对误差的近似表达

$$E_r(x) \approx \frac{E_a(x)}{x^*} = \frac{x^* - x}{x^*}$$

相对误差 $E_r(x)$ 可正可负,它的绝对值的上界称为相对误差限,即若存在正数 ε_r,使得

$$|E_r(x)| = \left|\frac{E_a(x)}{x}\right| \approx \left|\frac{x^* - x}{x^*}\right| \leqslant \varepsilon_r$$

成立,则称正数 ε_r 为近似值 x^* 的**相对误差限**(Relative Error Bound).

例 1-1 设有两个量 $x = 10 \pm 1$,$y = 1000 \pm 3$,求 x 与 y 的相对误差限.

解 $E_r(x) \approx \left|\dfrac{x^* - x}{x^*}\right| \leqslant \dfrac{1}{10} = 10\%$

$$E_r(y) \approx \left|\frac{y^* - y}{y^*}\right| \leqslant \frac{3}{1000} = 0.3\%$$

根据定义,x 与 y 的相对误差限分别为 10% 和 0.3%.

上例表明,y^* 近似 y 的程度要比 x^* 近似 x 的程度好得多. 相对误差更能刻画近似值的精确度.

此外,由定义可知,绝对误差和绝对误差限是有量纲的量,而相对误差与相对误差限是无量纲的量.

例 1-2 设 $x^* = 4.32$ 是由准确值 x 经过四舍五入得到的,求 x^* 的绝对误差限和相对误差限.

解 由已知得 $4.315 \leqslant x < 4.325$,于是

$$-0.005 < x^* - x \leqslant 0.005$$

所以，x^* 的绝对误差限为 $\varepsilon_a = 0.005$，相对误差限为 $\varepsilon_r \approx 0.005/4.32 \approx 0.12\%$.

2. 有效数字

当一个准确数 x 有很多位时，通常按照四舍五入原则得到 x 的近似值 x^*. 例如，无理数 $\pi = 3.1415926535897\cdots$，若按照四舍五入原则分别取 2 位和 4 位小数时，可得

$$\pi \approx 3.14, \qquad \pi \approx 3.1416$$

不管取几位小数，得到的近似数其误差的绝对值都不超过末尾数数位的半个单位，即

$$|3.14 - \pi| \leqslant \frac{1}{2} \times 10^{-2}, \ |3.1416 - \pi| \leqslant \frac{1}{2} \times 10^{-4}$$

结合例 1-2 可见，四舍五入得到的近似值，其误差的绝对值都不超过其末尾数字的半个单位.

定义 1-3 设 x 为准确值，x^* 为 x 的一个近似值，如果 x^* 的误差的绝对值不超过它的某一数位的半个单位，并且从 x^* 左起第一个非零数字到该数位共有 n 位，则称这 n 个数字为 $\boldsymbol{x}^*$ **的有效数字**（Significant Figures），也称用 x^* 近似 x 时具有 $\boldsymbol{n}$ **位有效数字.**

例 1-3 若下列近似数的绝对误差限都是 0.0005，它们各具有几位有效数字？

(1) $a = 251.234$；(2) $b = -0.208$；(3) $c = 0.002$；(4) $d = 0.00013$

解 因为 $0.0005 = \frac{1}{2} \times 10^{-3}$ 是小数点后第三位的半个单位，所以 a 有 6 位有效数字 2,5,1,2,3,4；b 有 3 位有效数字 2,0,8；c 有 1 位有效数字 2；d 没有有效数字.

有效数字还有另外一种定义方法：

定义 1-4 设准确值 x 的一个近似值 x^* 可以写成如下标准形式

$$x^* = \pm 0.a_1 a_2 \cdots a_n \times 10^m$$

其中 m 为整数，$a_i (i = 1, 2, \cdots, n)$ 是 0 到 9 中的某一数字，且 $a_1 \neq 0$. 如果 x^* 的绝对误差满足

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

则称近似值 x^* **有 $\boldsymbol{k}$ 位有效数字**，分别为 $a_1, a_2, \cdots, a_k$.

例 1-4 若分别取 3.1416 和 3.1415 作为无理数 π 的近似值，试确定它们的有效数字位数.

解 $3.1416 = 0.31416 \times 10^1$，这里 $m = 1, n = 5$. 因为

$$|3.1416 - \pi| = 0.0000073465\cdots < \frac{1}{2} \times 10^{-4}$$

所以，$1 - k = -4, k = 5$. 因此，3.1416 作为 π 的近似值具有 5 位有效数字.

$3.1415 = 0.31415 \times 10^1$，这里 $m = 1, n = 5$. 因为

$$|3.1415 - \pi| = 0.0000926\cdots < \frac{1}{2} \times 10^{-3}$$

所以，$1 - k = -3, k = 4$. 因此，3.1415 作为 π 的近似值具有 4 位有效数字.

上例表明准确值 x 的近似值 x^* 的每一位数字不一定都是有效数字，如 3.1415 作为 π 的近似值只有 4 位有效数字 3,1,4,1.

根据定义,若 x^* 是经“四舍五入”得到的近似值,那么它从第一个非零数字开始的所有数字都是有效数字.

1.3 数值算法设计的原则

定义 1－5 由基本运算和运算顺序的规则所构成的完整的解题步骤称为**算法**(Algorithm).

数学本身是严谨的,但计算机所能表示的数的位数是有限的,因而误差不可避免. 用数学上通过恒等变形获得的完全等价的两个式子在计算机上分别进行运算时,结果可能会有很大差异. 为了减少舍入误差的影响,设计数值算法时应遵循如下原则.

1. 通过简化计算步骤减少运算次数

同样一个计算问题,如果能减少运算次数,不但可节省计算时间,提高计算速度,而且能减少舍入误差的积累,这是数值计算必须遵循的原则,也是数值计算方法要研究的重要内容.

例如,计算 x^{255} 的值,如果将 x 的值逐个相乘,要用 254 次乘法,但如果写成

$$x^{255} = x \cdot x^2 \cdot x^4 \cdot x^8 \cdot x^{16} \cdot x^{32} \cdot x^{64} \cdot x^{128}$$

则只要做 14 次乘法运算即可.

又如,计算多项式

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

的值,若直接计算 $a_k x^k (k=0,1,\cdots,n)$,再逐项相加,一共需做

$$n + (n-1) + \cdots + 2 + 1 = \frac{n(n+1)}{2}$$

次乘法和 n 次加法. 若采用秦九韶算法

$$P(x) = (((a_n x + a_{n-1})x + a_{n-2})x + \cdots + a_1)x + a_0$$

则只需要 n 次乘法和 n 次加法即可.

2. 避免两个相近的数相减

如果 x^* 和 y^* 分别是准确值 x 和 y 的近似值,则 $z^* = x^* - y^*$ 是 $z = x - y$ 的近似值,此时 z^* 的相对误差满足

$$|E_r(z)| \approx \left|\frac{z^* - z}{z^*}\right| \leqslant \left|\frac{x^*}{x^* - y^*}\right| \cdot |E_r(x)| + \left|\frac{y^*}{x^* - y^*}\right| \cdot |E_r(y)|$$

所以,当 x^* 和 y^* 很接近时,z^* 的相对误差可能很大.

例如,当 $x = 1000$ 时,计算 $\sqrt{x+1} - \sqrt{x}$ 的值.

在 4 位浮点十进制数(仿机器实际计算)下直接计算得

$$\sqrt{x+1} - \sqrt{x} = \sqrt{1001} - \sqrt{1000} \approx 31.64 - 31.62 = 0.02$$

这个结果只有 1 位有效数字.

如果改用公式

$$\sqrt{x+1} - \sqrt{x} = \frac{1}{\sqrt{x+1} + \sqrt{x}} = \frac{1}{\sqrt{1001} + \sqrt{1000}} \approx \frac{1}{31.64 + 31.62} \approx 0.01581$$

则结果有 4 位有效数字.

上例表明,利用恒等式或等价关系对计算公式变形可以避免或减少有效数字的损失. 以下为几个常用的公式变换的例子:

如果 x_1 与 x_2 很接近,公式变换为

$$\lg x_1 - \lg x_2 = \lg \frac{x_1}{x_2},\ \ln x_1 - \ln x_2 = \ln \frac{x_1}{x_2}$$

如果 x 接近于 0,公式变换为

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

如果 x 充分大,公式变换为

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

$$\arctan(x+1) - \arctan x = \arctan \frac{1}{1+x(x+1)}$$

3. 避免除数绝对值远远小于被除数绝对值的除法

如果 x^* 和 y^* 分别是准确值 x 和 y 的近似值,则 $z^* = \dfrac{x^*}{y^*}$ 是 $z = \dfrac{x}{y}$ 的近似值,此时 z^* 的绝对误差满足

$$|E_a(z)| = |z^* - z| = \left|\frac{(x^*-x)y + x(y-y^*)}{yy^*}\right| \leqslant \frac{|x^*-x||y| + |x||y^*-y|}{|yy^*|}$$

$$\approx \frac{|x^*|\cdot|E_a(y)| + |y^*|\cdot|E_a(x)|}{(y^*)^2}$$

所以,若除数太小,则可能导致商的绝对误差很大.

例 1-5 在 4 位浮点十进制数(仿机器实际计算)下求解线性方程组

$$\begin{cases} 0.00001x_1 + x_2 = 1 \\ 2x_1 + x_2 = 2 \end{cases}$$

解 用消元法求解. 方程组可化为

$$\begin{cases} 0.1000 \times 10^{-4} \cdot x_1 + 0.1000 \times 10^{1} \cdot x_2 = 0.1000 \times 10^{1} \\ 0.2000 \times 10^{1} \cdot x_1 + 0.1000 \times 10^{1} \cdot x_2 = 0.2000 \times 10^{1} \end{cases}$$

若用 $\frac{1}{2}(0.1000 \times 10^{-4})$ 除第一个方程再减去第二个方程,得

$$\begin{cases} 0.1000 \times 10^{-4} \cdot x_1 + 0.1000 \times 10^{1} \cdot x_2 = 0.1000 \times 10^{1} \\ 0.2000 \times 10^{6} \cdot x_2 = 0.2000 \times 10^{6} \end{cases}$$

由此解得

$$\begin{cases} x_1 = 0 \\ x_2 = 1 \end{cases}$$