普通高等教育“十二五”规划教材
工 科 数 学 精 品 丛 书
海 军 院 校 重 点 教 材

数值计算方法

何汉林 李 薇 王 炜 主编

科 学 出 版 社

普通高等教育“十二五”规划教材
工科数学精品丛书
海军院校重点教材

数值计算方法

何汉林　李　薇　王　炜　主编

科　学　出　版　社
北　京

内 容 简 介

本书着重介绍了与现代计算有关的数值计算方法的基本概念、基本理论和基本方法，以方便读者今后的继续学习和科学研究、工程计算等打下坚实的数值分析与科学计算基础. 本书理论叙述严谨、精练，概念明确，系统性较强. 全书内容主要包括：线性代数方程组求解、非线性方程求根、插值方法、数值积分与微分、微分方程数值解法、最佳平方逼近理论、矩阵特征值与特征向量计算、数值计算方法的 MATLAB 实现.

本书可作为高等学校工科各专业和部分理科专业高年级本科生及工科研究生的数值分析课程教材，也可作为教学和科研工作者的参考书.

图书在版编目(CIP)数据

数值计算方法/何汉林，李薇，王炜主编. —北京：科学出版社，2016.2

(工科数学精品丛书)

普通高等教育"十二五"规划教材　海军院校重点教材

ISBN 978-7-03-047198-7

Ⅰ.①数…　Ⅱ.①何…　②李…　③王…　Ⅲ.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(2016)第 016837 号

责任编辑：王雨舸/责任校对：董艳辉　王　晶

责任印制：彭　超/封面设计：苏　波

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

武汉市首壹印务有限公司印刷

科学出版社发行　各地新华书店经销

*

开本：787×1092　1/16

2016 年 2 月第　一　版　印张：19 3/4

2016 年 2 月第一次印刷　字数：457 000

定价：43.90 元

(如有印装质量问题，我社负责调换)

前　言

科学技术、工业生产等各领域的问题通过建立数学模型与数学产生了紧密的联系，但实际应用中所导出的数学模型往往不能方便地求出精确解，只能简化模型或利用其他方法求出近似解，数值计算方法就是求模型近似解的重要方法. 本书包含科学计算中一些最基本的数值方法及其分析，这些方法可以直接应用于某些科学与工程计算，也是其他一些更复杂数值计算方法的基础. 本书可作为高等学校计算机、物理等各专业高年级本科生和工科各专业研究生的数值分析或计算方法课程教材，也可作为教学和科研工作者的参考书.

本书编者长期从事数学教学与科研工作，多次承担本科数值分析课程以及研究生相关课程的教学任务，积极跟踪教改、课改动态，积累了比较丰富的教学实践经验. 本书力求反映数值分析课程的最新发展动态，注意汲取国内外同类教材的优点，在教材内容取材与具体编排等方面反复研究斟酌，在保持课程知识体系相对完整和严谨的同时，精选优化教学内容，注重理论联系实际，内容由浅入深、循序渐进，突出重点，加强应用，启发思维，努力使编写的教材成为一部特色鲜明、针对性和适用性强的教学用书.

本书共 9 章：第 1 章绪论，主要介绍误差与误差分析的有关概念，MATLAB 软件的基本知识；第 2 章解线性方程组的直接法，主要介绍解线性方程组的两种重要的消去法和各种三角分解法，并给出了方程组的性态、误差分析和病态方程组的解法；第 3 章解线性方程组的迭代法，主要介绍解线性方程组的雅可比迭代法、赛德尔迭代法和超松弛迭代法，并讨论了其收敛性；第 4 章非线性方程求根，主要介绍迭代法的构造原理，迭代法的收敛性判别，牛顿迭代法及其变形；第 5 章插值法，主要介绍拉格朗日插值公式、牛顿插值公式、埃尔米特插值公式以及三次样条插值并讨论其余项；第 6 章数值积分与数值微分，主要介绍常用的数值求积和数值微分公式及其误差估计、求积公式的代数精确度、收敛性和稳定性等；第 7 章常微分方程的数值解法，主要介绍单步法，单步法的收敛性、稳定性，几种常用的多步法以及常微分方程组的单步法和二阶方程边值问题的差分方法；第 8 章最佳平方逼近，主要介绍最佳平方逼近的基本概念、基本方法，常用的正交多项式以及曲线拟合的最小二乘法；第 9 章矩阵特征值和特征向量的计算，主要介绍求矩阵最大及最小特征值和相应特征向量的幂法与反幂法、正定矩阵全部特征值和相应特征向量的雅可比方法以及指定特征值的吉文斯-豪斯霍尔德方法和 QR 方法. 每章最后一节介绍算法的 MATLAB 实现. 全书例题和习题比较丰富，在每章后配置习题，书后附有 3 套模拟试卷，大部分习题难易程度适中，侧重于对所学知识、方法的理解与运用，部分习题具有扩充视野的作用，是对正文内容的某种补充. 书后附有习题答案，便于读者自学和检测评估学习效果.

本书第 1、2、3 章由李薇编写，第 4、5、6 章由王炜编写，第 7、8、9 章及模拟试卷由何汉林编写，全书由何汉林负责统稿、定稿工作.

本书被海司院校部列为海军级重点教材，它的编写出版得到了科学出版社领导和编辑的大力支持和帮助，海军工程大学训练部领导、理学院领导以及应用数学系的同事们给予我们关心和支持.在此谨向他们表示诚挚的谢意.本书编写参考了大量资料，对于书末所列参考书目的作者们也一并表示由衷的敬意和真诚的感谢.

由于编者水平所限，不足之处在所难免，敬请专家和读者批评指正.

编　者

2015 年 10 月

目　　录

第1章 绪 论

本章简要介绍数值分析的研究对象、内容和特点，介绍误差的基本概念以及有效数字与误差的关系，并讨论了数值计算的误差估计问题. 其中利用函数的泰勒(Taylor) 展开式估计误差是误差定量估计的一种基本方法. 除此之外，本章还着重讨论了误差的定性分析以及在数值计算中应当普遍遵循的若干原则. 考虑到基于 MATLAB 软件的数值计算日益广泛，本章最后简要介绍了 MATLAB 软件的基础知识.

1.1 数值分析的研究对象与特点

数值分析也称为计算方法，是数学科学的一个重要分支，它研究用计算机求解数学问题的数值计算方法及其软件实现. 随着电子技术的发展和科学技术的进步，电子计算机的使用日益广泛. 计算机作为科学计算的主要工具越来越不可缺少，因而需要研究适合计算机使用的数值计算方法. 为了更具体地说明数值分析的研究对象，我们考察用计算机解决科学计算问题的一般过程，可以概括为

实际问题 → 数学模型 → 计算方法 → 程序设计 → 上机计算

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，通常作为应用数学的任务. 而根据数学模型提出求解的计算方法及编出程序并上机算出结果，进而对计算结果进行分析，这一过程则是计算数学的任务，也是数值分析的研究对象. 因此，数值分析就是研究用计算机解决数学问题的数值方法及其理论. 它的内容包括：误差理论、线性与非线性方程(组) 的数值解、矩阵的特征值与特征向量计算、曲线拟合与函数逼近、插值方法、数值积分与数值微分、常微分方程与偏微分方程数值解等. 本书只介绍科学与工程计算中最常用的基本数值方法，包括线性方程组与非线性方程求根、插值与最佳平方逼近、数值积分与数值微分、常微分方程数值解法、矩阵的特征值与特征向量计算. 这些都是数值分析中最基础的内容.

数值分析是一门与计算机密切结合的实用性很强的数学课程，它既具有纯数学的高度抽象性与严密科学性的特点，又具有应用广泛性与实际试验的高度技术性的特点. 例如，考虑线性方程组的解，线性代数中只介绍解的存在唯一性及有关理论和精确解法，用这些理论和方法还不能直接在计算机上求解. 我们知道，用克拉默(Cramer) 法则求解一个 n 阶线性方程组，要算 $n+1$ 个 n 阶行列式，总共需要 $(n-1)(n+1)n!$ 次乘法，当 n 较大时，计算量是相当惊人的. 而如果用消元法，求解一个 n 阶线性方程组大约需要 $\frac{1}{3}n^3+n^2$ 次乘法. 这一简单的例子告诉我们，能否正确地制定算法，是科学计算成败的关键. 另外，求解这类问题还应根据方程特点，研究适合计算机使用的、满足精度要求的、计算时间节省的有效算法及其相关理论. 在实现这些算法时往往还要根据计算机容量、字长、速度等指标，研究具体求解步骤和程序设计技巧. 有的方法在理论上虽不够严密，但通过实际计

算、对比分析等手段,证明是行之有效的方法,也应该采用.这些都是数值分析应有的特点.

概括地说,数值分析是研究适合于在计算机上使用的实际可行、理论可靠、计算复杂性好的数值计算方法.具体说就是:

(1) 面向计算机.要根据计算机特点提供实际可行的算法,即算法只能由计算机可执行的加减乘除四则运算和各种逻辑运算组成.

(2) 要有可靠的理论分析.数值分析中的算法理论主要是连续系统的离散化及离散型方程数值求解.有关基本概念包括误差、稳定性、收敛性、计算量、存储量等,这些概念用于刻画计算方法的可靠性、准确性、效率以及使用的方便性.

(3) 要有良好的计算复杂性及数值试验.计算复杂性是算法好坏的标志,它包括时间复杂性(指计算时间多少)和空间复杂性(指占用存储单元多少).对很多数值问题使用不同算法,其计算复杂性将会大不一样,例如对 20 阶的线性方程组若用克拉默法则作为算法求解,其乘除法运算次数需要 9.7×10^{20},若用每秒运算 1 亿次的计算机计算也要 30 万年,这是无法实现的,而用数值分析中介绍的高斯(Gauss)消去法求解,其乘除法运算次数只需 3060 次,这说明选择算法的重要性.当然有很多数值方法不可能事先知道其计算量,故对所有数值方法除理论分析外,还必须通过数值试验检验其计算复杂性.

1.2　数值计算的误差

数值计算的结果与实际问题的真值之间往往存在差异,这种差异称为误差.除了极个别的情况外,数值计算总是存在误差.数值分析的任务之一是将误差控制在一定的允许范围内或者至少对误差有所估计.

1.2.1　误差的来源与分类

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与实际问题之间存在一定的差异,我们把数学模型与实际问题之间出现的这种差异称为模型误差.只有实际问题提法正确,建立数学模型时又抽象、简化得合理,才能得到好的结果.由于这种误差难于用数量表示,通常都假定数学模型是合理的,这种误差可忽略不计,在数值分析中不予讨论.在数学模型中往往包含一些由观测或实验得到的物理量,如温度、重量、长度等,由于测量工具精度和测量手段的限制,它们与实际量大小之间必然存在差异,这种差异称为观测误差.观测误差在数值分析中也不予讨论.数值分析中所涉及的误差,主要指以下两类:

第一类是截断误差或方法误差,它是指将数学问题转化为数值计算问题时产生的误差,通常是用有限过程近似无限过程时产生的误差.例如,计算 $f(x)=\mathrm{e}^x$ 的值,用泰勒公式展开前 4 项为

$$\mathrm{e}^x \approx 1+x+\frac{x^2}{2}+\frac{x^3}{6}=P_3(x)$$

当 $|x|<1$ 时其截断误差为

$$R_3(x)=\mathrm{e}^x-P_3(x)=\frac{1}{4!}\mathrm{e}^{\xi}x^4 \quad (\xi\text{在 }0\text{ 与 }x\text{ 之间})$$

第二类是舍入误差,数值计算时由于计算是有限位的,所以原始数据、中间结果和最后结果都要舍入,这就产生舍入误差. 在十进制运算中一般采用四舍五入,例如 $\frac{1}{3}$ 写成 0.3333,$\pi\approx 3.1416$ 等,都有舍入误差.

数值分析需要讨论这两类误差在计算过程中的传播和对计算结果的影响,研究控制它们的方法以保证最终结果有足够的精度.

1.2.2 误差与有效数字

定义 1.1 设准确值 x 的近似值为 x^*,则 $e(x^*)=x^*-x$ 称为近似值 x^* 的绝对误差,简称误差,$e_r(x^*)=\dfrac{x^*-x}{x}$ 称为近似值 x^* 的相对误差.

绝对误差可正可负,一般来说 $e(x^*)$ 的准确值很难求出,往往只能求 $|e(x^*)|$ 的一个上界 $\delta(x^*)$,即

$$|e(x^*)|=|x^*-x|\leqslant\delta(x^*)$$

$\delta(x^*)$ 称为 x^* 的误差限. 相对误差 $e_r(x^*)$ 当 $x=0$ 时没有意义,且准确值 x 往往未知,故常用 $\dfrac{x^*-x}{x^*}$ 作为相对误差,而称 $\delta_r(x^*)=\dfrac{\delta(x^*)}{|x^*|}$ 为相对误差限.

例 1.1 已知 $\pi=3.141\,592\,6\cdots$,若取近似数为 $x^*=3.14$,则

$$e(x^*)=x^*-\pi=-0.001\,592\,6\cdots,\qquad |e(x^*)|\leqslant 0.002=\delta(x^*)$$

为 x^* 的误差限,而相对误差限

$$\delta_r(x^*)=\frac{\delta(x^*)}{3.14}<0.0007$$

通常在 x 有多位数字时,若取前有限位数的数字作近似值,都采用四舍五入原则. 例如,$x=\pi$,

取 3 位:$x^*=3.14$,$|e(x^*)|\leqslant 0.002$,

取 5 位:$x^*=3.1416$,$|e(x^*)|\leqslant 0.000\,01$,

它们的误差限都不超过近似数 x^* 末位数的半个单位,即

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

定义 1.2 设 x^* 是 x 的一个近似数,表示为

$$x^*=\pm 0.a_1a_2\cdots a_n\times 10^k \tag{1-1}$$

每个 $a_i(i=1,2,\cdots,n)$ 均为 $0,1,\cdots,9$ 中的一个数字,且 $a_1\neq 0$,如果

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

则称 x^* 近似 x 有 n 位有效数字.

例如,用 3.14 近似 π 有 3 位有效数字,用 3.1416 近似 π 有 5 位有效数字. 显然,近似数的有效位数越多,相对误差限就越小,反之也正确.

定理 1.1 设 x 的近似数 x^* 表示为式(1-1),如果 x^* 具有 n 位有效数字,则其相对误差限为

$$\frac{|x^*-x|}{|x^*|}\leqslant\delta_r(x^*)=\frac{1}{2a_1}\times 10^{-(n-1)} \tag{1-2}$$

反之,若

$$\frac{|x^*-x|}{|x^*|}\leqslant\delta_r(x^*)=\frac{1}{2(a_1+1)}\times 10^{-(n-1)} \tag{1-3}$$

则 x^* 至少具有 n 位有效数字.

证 由式(1-1),得

$$a_1\times 10^{k-1}\leqslant|x^*|\leqslant(a_1+1)\times 10^{k-1}$$

所以当 x^* 有 n 位有效数字时,

$$\frac{|x^*-x|}{|x^*|}\leqslant\frac{\frac{1}{2}\times 10^{k-n}}{a_1\times 10^{k-1}}=\frac{1}{2a_1}\times 10^{-(n-1)}$$

反之,由式(1-3),有

$$|x^*-x|\leqslant|x^*||\delta_r(x^*)|\leqslant(a_1+1)\times 10^{k-1}\times\frac{1}{2(a_1+1)}\times 10^{-(n-1)}=\frac{1}{2}\times 10^{k-n}$$

故 x^* 有 n 位有效数字.

例 1.2 下列近似数有几位有效数字?其相对误差限是多少?

(1) $x=\mathrm{e}\approx 2.71828=x^*$　　(2) $x=0.030021\approx 0.0300=x^*$.

解 (1) 由 $|2.71828-\mathrm{e}|<\frac{1}{2}\times 10^{-5}$,因 $k=1$,故 $n=6$,即 x^* 有 6 位有效数字.因 $a_1=2$,相对误差限 $\delta_r(x^*)=\frac{1}{4}\times 10^{-5}$.

(2) $|0.0300-x|<\frac{1}{2}\times 10^{-4}$,因 $k=-1$,故 $n=3$,即 x^* 有 3 位有效数字.由 $a_1=3$ 知,相对误差限 $\delta_r(x^*)=\frac{1}{6}\times 10^{-2}$.

1.2.3 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近似数,它们都有误差.而这些数据的误差在多次运算中又会进行传播,使计算结果产生一定的误差,这就是误差的传播问题.以下介绍利用函数的泰勒公式来估计误差的一种常用方法.

设一元函数 $f(x)$ 具有二阶导数,自变量 x 的一个近似值是 x^*,$f(x)$ 的近似值为 $f(x^*)$,用 $f(x)$ 在 x^* 点的泰勒展开估计误差,可得

$$|f(x)-f(x^*)|\leqslant|f'(x^*)(x-x^*)|+\frac{1}{2}|f''(\xi)(x-x^*)^2|$$

其中 ξ 在 x 与 x^* 之间,如果 $f'(x^*)\neq 0$,且 $|f''(\xi)|$ 与 $|f'(x^*)|$ 有相同数量级,而 $|x-x^*|$ 很小,则 $f(x^*)$ 的近似误差限与近似相对误差限分别为

$$\delta(f(x^*))\approx|f'(x^*)|\delta(x^*),\quad \delta_r(f(x^*))\approx\left|\frac{f'(x^*)}{f(x^*)}\right|\delta(x^*) \tag{1-4}$$

如果 f 为多元函数,自变量为 $x_1,x_2,\cdots,x_n$,其近似值为 $x_1^*,x_2^*,\cdots,x_n^*$,类似于一元函数,可用多元函数 $f(x_1,x_2,\cdots,x_n)$ 的泰勒展开,取一阶近似得误差限为

$$\delta(f(x_1^*,x_2^*,\cdots,x_n^*))\approx\sum_{i=1}^{n}\left|\frac{\partial f(x_1^*,x_2^*,\cdots,x_n^*)}{\partial x_i}\right|\delta(x_i^*) \tag{1-5}$$

相对误差限为

$$\begin{aligned}&\delta_r(f(x_1^*,x_2^*,\cdots,x_n^*))\\ \approx&\sum_{i=1}^{n}\left|\frac{\partial f(x_1^*,x_2^*,\cdots,x_n^*)}{\partial x_i}\right|\frac{\delta(x_i^*)}{|f(x_1^*,x_2^*,\cdots,x_n^*)|}\\ \approx&\sum_{i=1}^{n}\left|\frac{\partial f(x_1^*,x_2^*,\cdots,x_n^*)}{\partial x_i}\right|\frac{|x_i^*|}{|f(x_1^*,x_2^*,\cdots,x_n^*)|}\delta_r(x_i^*)\end{aligned} \tag{1-6}$$

以上两式中的各项

$$\left|\frac{\partial f(x_1^*,x_2^*,\cdots,x_n^*)}{\partial x_i}\right| \quad 和 \quad \left|\frac{x_i^*}{f(x_1^*,x_2^*,\cdots,x_n^*)}\right|\left|\frac{\partial f(x_1^*,x_2^*,\cdots,x_n^*)}{\partial x_i}\right|$$

分别为各个 $x_i^*\ (i=1,2,\cdots,n)$ 对 $f(x_1^*,x_2^*,\cdots,x_n^*)$ 的绝对误差限和相对误差限的增长因子,它们分别表示绝对误差限 $\delta(x_i^*)$ 和相对误差限 $\delta_r(x_i^*)$ 经过传播后增大或缩小的倍数.

若把式(1-5)用到两个或多个数的算术运算中,则可得到近似数 x_1^* 及 x_2^* 的四则运算误差估计

$$\delta(x_1^*\pm x_2^*)=\delta(x_1^*)+\delta(x_2^*) \tag{1-7}$$

$$\delta(x_1^*x_2^*)\approx|x_1^*|\delta(x_2^*)+|x_2^*|\delta(x_1^*) \tag{1-8}$$

$$\delta\left(\frac{x_1^*}{x_2^*}\right)\approx\frac{|x_1^*|\delta(x_2^*)+|x_2^*|\delta(x_1^*)}{|x_2^*|^2}\quad(x_2^*\neq 0) \tag{1-9}$$

例 1.3 设测得某场地长 l 的值为 $l^*=110\text{ m}$,宽 d 的值为 $d^*=80\text{ m}$,已知

$$|l-l^*|\leqslant 0.2\text{ m},\quad |d-d^*|\leqslant 0.1\text{ m}$$

试求面积 $S=ld$ 的绝对误差限与相对误差限.

解 因 $S=ld,\dfrac{\partial S}{\partial l}=d,\dfrac{\partial S}{\partial d}=l$,由式(1-5)知

$$\delta(S^*)\approx\left|\left(\frac{\partial S}{\partial l}\right)^*\right|\times\delta(l^*)+\left|\left(\frac{\partial S}{\partial d}\right)^*\right|\times\delta(d^*)$$

其中$\left(\dfrac{\partial S}{\partial l}\right)^*=d^*=80\text{ m},\left(\dfrac{\partial S}{\partial d}\right)^*=l^*=110\text{ m},\delta(d^*)=0.1\text{ m},\delta(l^*)=0.2\text{ m}$,从而绝对误差限为

$$\delta(S^*)\approx 80\times 0.2+110\times 0.1=27(\text{m}^2)$$

相对误差限为

$$\delta_r(S^*)\approx\frac{\delta(S^*)}{|S^*|}=\frac{27}{80\times 110}=0.31\%$$

1.3 误差定性分析与避免误差危害

上述误差估计方法只对运算量很少的情形适用,对大规模数值计算的舍入误差估计,目前尚无有效的方法做出定量分析. 为了确保数值计算结果的正确性,应对数值计算问

题进行定性分析，以保证其舍入误差不会影响计算的精度，这就是本节要讨论的问题.

1.3.1 病态问题与条件数

对一个数值问题，往往由于问题本身而使计算结果相对误差很大，这种问题就是病态问题. 例如计算函数值 $f(x)$，若 x 的近似值为 x^*，其相对误差为 $\frac{x^*-x}{x}$，函数值 $f(x^*)$ 的相对误差为 $\frac{f(x^*)-f(x)}{f(x)}$，它们相对误差之比的绝对值为

$$\left|\frac{[f(x)-f(x^*)]/f(x)}{(x-x^*)/x}\right| \approx \left|\frac{xf'(x)}{f(x)}\right| = C_p \tag{1-10}$$

C_p 称为计算函数值 $f(x)$ 的条件数. 如果 C_p 很大，将引起函数值 $f(x^*)$ 的相对误差很大，出现这种情况时，就认为问题是病态的. 例如 $f(x)=x^n$，$f'(x)=nx^{n-1}$，则 $C_p=n$，它表示相对误差可能放大 n 倍. 如 $n=10$，有 $f(1)=1$，$f(1.02)\approx 1.24$. 若 $x=1$，$x^*=1.02$，则自变量相对误差为 2%，而函数值 $f(1.02)$ 的相对误差为 24%，这时就认为问题是病态的. 一般情况下若条件数 $C_p\geqslant 10$，则认为是病态问题，C_p 越大病态越严重.

其他计算问题也要分析是否病态，例如解线性方程组，如果输入数据有微小误差，引起解的相对误差很大，就认为此方程组是病态方程组.

例 1.4　分析方程组 $\begin{cases}x+ay=1,\\ ax+y=0\end{cases}$ 是否病态.

解　当 $a=1$，系数矩阵奇异，方程组无解；当 $a\neq 1$，解为 $x=\frac{1}{1-a^2}$，$y=\frac{-a}{1-a^2}$；当 $a\approx 1$，若输入数据 a 有误差，则解的误差很大. 例如，当 $a=0.99$，解 $x\approx 50.25$，当 a 有误差，$a^*=0.991$，则解 $x^*\approx 55.81$，误差 $|x-x^*|\approx 5.56$ 很大，问题病态. 对此例中 $x=(1-a^2)^{-1}$ 应用式(1-10) 求 C_p，得

$$C_p=\left|\frac{ax'(a)}{x(a)}\right|=\left|\frac{2a^2}{1-a^2}\right|$$

当 $a=0.99$ 时 $C_p\approx 100$，故问题病态. 只有当 $C_p\ll 1$ 时问题才为良态.

1.3.2 算法的数值稳定性

一个数值方法进行计算时，由于原始数据有误差，在计算中这些误差会传播，有时误差增长很快使计算结果误差很大，影响了结果的可靠性.

定义 1.3　一个算法如果原始数据有扰动(即误差)，而计算过程舍入误差不增长，则称此算法是数值稳定的. 否则，若舍入误差增长，则称算法不稳定.

例 1.5　计算积分 $I_n=\int_0^1 x^n e^{x-1}\,dx\ (n=0,1,\cdots)$. 当 $n=0$ 时，$I_0=1-e^{-1}$，对 I_n 用分部积分法，得

$$I_n=1-nI_{n-1}\quad (n=1,2,\cdots) \tag{1-11}$$

若计算 I_0 时取 $e^{-1}\approx 0.3679$，由式(1-11) 依次计算 $I_1,I_2,\cdots,I_9$，考察其计算结果是否正确

(取4位有效数字).

解 由 $I_0 = 1 - e^{-1} \approx 0.6321 = I_0^*$,故有误差 $e_0 = I_0 - I_0^*$. 由式(1-11),计算

$$I_n^* = 1 - nI_{n-1}^* \quad (n = 1,2,\cdots,9)$$

得

$$I_1^* = 0.3679, \quad I_2^* = 0.2642, \quad I_3^* = 0.2074,$$
$$I_4^* = 0.1704, \quad I_5^* = 0.1480, \quad I_6^* = 0.1120,$$
$$I_7^* = 0.2160, \quad I_8^* = -0.7280, \quad I_9^* = 7.552$$

显然,结果不正确,因为 $I_n > 0$,而 $I_8^* < 0$. 实际上,

$$e_n = I_n - I_n^* = -n(I_{n-1} - I_{n-1}^*) = \cdots = (-1)^n n! e_0$$

当 n 增大时 $|e_n|$ 是递增的,且 I_9^* 的误差达到 $-9!e_0$,是严重失真的. 它表明式(1-11)给出的算法是不稳定的. 如果在式(1-11)中将算法改为

$$I_{n-1} = \frac{1}{n}(1 - I_n) \quad (n = 9,8,\cdots,2,1) \tag{1-12}$$

由于 $\frac{1}{n+1}e^{-1} \leqslant I_n \leqslant \frac{1}{n+1}$,取 $I_n \approx \frac{1}{2}\frac{1}{n+1}(e^{-1}+1)$. 当 $n = 9$,

$$I_9 = \frac{1}{20}(1 + e^{-1}) \approx 0.0684 = \tilde{I}_9$$

再由式(1-12),可求出

$$\tilde{I}_8 = 0.01035, \cdots, \tilde{I}_2 = 0.2643, \tilde{I}_1 = 0.2673, \tilde{I}_0 = 0.6321$$

此时,

$$\tilde{\varepsilon}_{n-1} = I_{n-1} - \tilde{I}_{n-1} = -\frac{1}{n}(I_n - \tilde{I}_n), \quad |\tilde{\varepsilon}_0| = \frac{1}{n!}|\tilde{\varepsilon}_n|$$

计算是稳定的.

注:数值不稳定的算法是不能使用的.

1.3.3 避免误差危害的若干原则

数值计算中除了要分清问题是否病态和算法的数值稳定性外,还应尽量避免误差危害. 通常运算中应注意以下若干原则:

(1) 避免用绝对值很小的数作除数.

例如,$\delta\left(\frac{x}{y}\right) \approx \frac{|x|\delta(y) + |y|\delta(x)}{|y|^2}$,当 $|x| \gg |y|$ 时,舍入误差会扩大.

例 1.6 x, y 的舍入误差均为 0.5×10^{-3},而 $y = x\times10^{-7}$,则 $\frac{x}{y}$ 的舍入误差为

$$\frac{(|x| + |x|\times10^{-7})\times0.5\times10^{-3}}{|x|^2\times10^{-14}} \approx \frac{1}{|x|}\times0.5\times10^{11}$$

绝对值很小的数作除数有时还会造成计算机的溢出而停机.

(2) 避免两个相近数相减,以免有效数字损失.

因两数之差 $x - y$ 的相对误差为 $e_r(x-y) = \frac{e(x-y)}{x-y}$,当 x 与 y 很接近时,两数之差

$x-y$ 的相对误差会很大,有效数位将严重丢失.

避免办法是进行变换.

例 1.7 $\sqrt{170}-13=0.038\,404\,8\cdots$,如用四位有效数字计算

$$\sqrt{170}-13=13.04-13=0.04$$

结果只有一位有效数字. 如改为

$$\sqrt{170}-13=\frac{1}{\sqrt{170}+13}=\frac{1}{13.04+13}=0.038\,40$$

结果有四位有效数字.

新算法避免了两个相近数的相减.

例 1.8 求 $x^2-16x+1=0$ 的小正根.

解 方程的两根为 $x_1=8-\sqrt{63}$ 及 $x_2=8+\sqrt{63}$,若取 $\sqrt{63}\approx 7.94$,小正根

$$x_1=8-\sqrt{63}\approx 8-7.94=0.06$$

只有一位有效数字. 为避免两相近数相减可改用

$$x_1=8-\sqrt{63}=\frac{1}{8+\sqrt{63}}\approx\frac{1}{15.94}\approx 0.0627$$

仍有三位有效数字.

(3) 注意运算次序,防止大数"吃掉"小数,如多个数相加减,应按绝对值由小到大的次序运算.

计算机在进行运算时,首先要把参加运算的数对阶,即把两数都写成绝对值小于 1 而阶码相同的数. 如若在十进制计算机(阶码为 9) 中表示 $a=10^8+1$,则必须将它改写成

$$a=0.1\times 10^9+0.000\,000\,001\times 10^9$$

如果计算机只能表示 8 位小数,则 0.000 000 001 在该计算机中表示为机器数 0,于是算出 $a=0.1\times 10^9$,大数"吃"了小数,这种情况有时允许,有时不允许.

例 1.9 $a=10^{10}, b=10, c=-a$. 则

$$(a+b)+c=(10^{10}+10)-10^{10}\approx 10^{10}-10^{10}=0$$

b 被大数吃掉了. 如按 $(a+c)+b=0+b=b$,b 就没有被吃掉. 这也是构造算法时要注意的问题.

例 1.10 一元二次方程 $x^2-(10^9+1)x+10^9=0$,其精确解为 $x_1=10^9, x_2=1$. 如用求根公式:$x_{1,2}=\dfrac{-b\pm\sqrt{b^2-4ac}}{2a}$ 和字长为 8 位的计算机求解,有

$$\sqrt{b^2-4ac}=\sqrt{10^{18}-2\times 10^9+1}\approx\sqrt{10^{18}}=10^9 \quad 及 \quad 10^9+1\approx 10^9$$

则

$$x_1\approx\frac{-(-10^9)+10^9}{2}=10^9,\quad x_2\approx\frac{-(-10^9)-10^9}{2}=0$$

x_2 的值与精确解差别很大. 若用

$$x_2=\frac{-b-\sqrt{b^2-4ac}}{2a}=\frac{2c}{-b+\sqrt{b^2-4ac}}\approx\frac{2\times 10^9}{-(-10^9)+10^9}=1$$

因此,算法的选用很重要.

(4) 注意简化计算步骤,尽量减少运算次数.

例 1.11 计算 x^{255} 的值.

如果逐个相乘要用 254 次乘法. 若将其改写为

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

只需 14 次乘法.

例 1.12 计算多项式的值:$P_n(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$.

计算 a_kx^k 须 k 次乘法运算,若保留 x^k 的运算结果,计算 $P_n(x)$ 共需 $2n-1$ 次乘法和 n 次加法运算. 如写成

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

用递推算法:$u_0=a_n,\ u_k=u_{k-1}x+a_{n-k}\quad(k=1,2,\cdots,n)$. 最终 $P_n(x)=u_n$,共需 n 次乘法和 n 次加法运算. 此算法称为秦九韶算法,也称霍纳(Horner)算法(秦九韶于1247年提出此算法,比霍纳 1819 年提出此算法早 500 多年).

例 1.13 计算 ln2 的近似值,要求误差小于10^{-5}.

解法 1 用级数

$$\ln 2=1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\cdots+\frac{(-1)^{n-1}}{n}+\cdots$$

的前 n 项部分和来计算. 截断误差 $R_1\leqslant\dfrac{1}{n+1}$,若$\dfrac{1}{n+1}\leqslant 10^{-5}$,则需 $n\geqslant 10^5$. 即要取前十万项求和,计算量大,舍入误差积累,将使有效数字丢失.

解法 2 用级数

$$\ln\frac{1+x}{1-x}=2x\left(1+\frac{1}{3}x^2+\frac{1}{5}x^4+\cdots+\frac{x^{2n}}{2n+1}+\cdots\right)$$

来计算. 当 $x=\dfrac{1}{3}$ 时,有

$$\ln 2=\frac{2}{3}\left[1+\frac{1}{3\times 9}+\frac{1}{5\times 9^2}+\cdots+\frac{1}{(2n+1)\times 9^n}+\cdots\right]$$

取前 5 项之和作近似值,产生的截断误差为

$$\begin{aligned}R_2&=\frac{2}{3}\left(\frac{1}{11\times 9^5}+\frac{1}{13\times 9^6}+\frac{1}{15\times 9^7}+\cdots\right)<\frac{2}{3}\times\frac{1}{11\times 9^5}\left(1+\frac{1}{9}+\frac{1}{9^2}+\cdots\right)\\&=\frac{2}{3\times 11\times 9^5}\times\frac{1}{1-\frac{1}{9}}=\frac{1}{12\times 11\times 9^4}<\frac{1}{10^5}\end{aligned}$$

最后指出误差的定性分析中首先要分清问题是否病态,如果问题计算结果相对误差很大就是病态问题,对病态问题计算结果就可能不可靠. 对非病态问题(即良态问题)主要考虑算法的稳定性,对不稳定的算法计算结果也不可靠;另外,计算过程中还要根据给出的上述原则尽量避免舍入误差增长.

1.4 MATLAB 软件简介

MATLAB是MATrix LABoratory("矩阵实验室")的缩写,是由美国MathWorks公司开发的集数值计算、符号计算和图形可视化三大基本功能于一体的,功能强大、操作简单的