普通高等教育“十一五”国家级规划教材

信息与计算科学专业教材系列

数值分析

林成森 编著

科学出版社
www.sciencep.com

普通高等教育“十一五”国家级规划教材

信息与计算科学专业教材系列

数值分析

林成森 编著

科学出版社

北京

内 容 简 介

本书系统阐述了数值分析的基本概念和理论. 内容包括: 数值计算的误差, 解线性方程组的直接法和迭代法, 线性方程组的最小二乘解, 矩阵特征值问题, 插值法, 函数逼近, 曲线拟合, 数值积分, 解非线性方程和方程组的数值方法.

本书适合高等院校信息与计算、数学、应用数学、计算机应用等专业的本科生作为教材, 也可供工程技术人员参考.

图书在版编目(CIP)数据

数值分析/林成森 编著. —北京: 科学出版社, 2007

普通高等教育"十一五"国家级规划教材·信息与计算科学专业教材系列

ISBN 978-7-03-018441-2

I. 数… II. 林… III. 数值计算-高等学校-教材 IV. O241

中国版本图书馆 CIP 数据核字(2007) 第 003307 号

责任编辑: 姚莉丽 祖翠娥 / 责任校对: 张怡君

责任印制: 张克忠 / 封面设计: 陈 敬

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

双青印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2007 年 1 月第 一 版 开本: B5(720×1000)

2007 年 1 月第一次印刷 印张: 26 3/4

印数: 1—3 000 字数: 510 000

定价: 33.00 元

(如有印装质量问题,我社负责调换(环伟))

前　言

自然科学、工程技术、经济、医学和人文等领域中的许多问题都可以应用有关学科知识和数学理论用数学语言描述为数学问题或建立数学模型. 这些问题中只有很少一部分可以给出解析解, 而绝大多数则得不到准确解或求解的计算工作量很大, 需要借助于计算机进行数值求解. 解这类问题通常称为科学计算. 当代科学计算已渗透到极广泛的专业领域中, 如计算物理、计算化学、计算生物学和计算经济学等. 数值分析或称数值计算方法, 又是它们的基础. 本书系统地介绍了基本常用的数值计算方法和一些现代数值方法及有关理论分析. 主要内容有解线性方程组的直接法和迭代法、插值法、函数逼近、数据的最小二乘拟合、数值积分、解非线性方程和方程组的数值方法、常微分方程初值问题和边值问题的数值解法、求线性方程组的最小二乘解的数值方法、矩阵特征值问题.

本书除了介绍常用的数值方法外, 还比较强调数值分析的基本原理和基本理论分析, 阐述严谨、详细、深入浅出. 为了加深读者对书中内容的理解, 我们给出了较多例题, 在每章后面配有相当数量的习题. 在书末给出习题答案, 并且绝大多数证明题都给予了提示供读者参考.

学习数值分析必须紧密结合计算机. 本书介绍的数值方法的算法用伪程序 (pseudocode) 给出, 以便在计算上实现这些算法, 读者容易把它们译成 FORTRAN、Pascal、C 语言或其他程序语言. 随着计算机技术的迅速发展, 计算机语言多样化及数学软件的普及, 已有现成数学软件如集成化软件包 Matlab 等, 方便了读者使用. 因此, 我们对一些比较复杂的数值方法不给出算法.

为了面向更多读者和减少篇幅, 我们在书中将一些读者需要了解但证明较难或较繁的理论结论述而不证.

本书可作为高等院校数学、计算机等系各专业及理工类有关专业学生的教材. 如果授课学时较少, 则可适当删减书中某些章节内容和一些理论, 这并不影响本书的系统性. 本书也可供工程技术人员参考.

由于作者水平有限, 书中尚有不足之处, 敬请各位老师和读者批评指正.

作　者

2005.7

目　录

第1章 误 差

1.1 数值方法

自然科学、工程技术、经济和医学等领域中遇到的许多问题都可以应用有关学科知识和数学理论用数学语言描述为数学问题, 或者说建立其数学模型. 然而, 这些数学问题往往得不到它的准确解, 或者解这种问题的计算工作量很大, 只能借助计算机求其近似解 (称为**数值解**或**计算解**). 例如, 人造卫星接收站的天线结构问题可化为成千上万阶线性方程组, 不借助于计算机则无法求其解. 应用计算机解数学问题 (数学模型) 的步骤首先是提出能在计算机上实现的数值方法, 然后用计算机语言编写程序, 最后上机计算求其结果.

数值方法是对给定问题的输入数据和所需结果 (输出) 之间的一种明确的数学描述. 例如, 我们用 Newton 法 (将在 8.4 节中讨论) 计算 $\sqrt{2}$. 输入 $\sqrt{2}$ 的一个初始近似值 $x_0(x_0>0)$, 如取 $x_0=2$, 由迭代公式

$$x_n=\frac{1}{2}\left(x_{n-1}+\frac{2}{x_{n-1}}\right),\quad n=1,2,\cdots$$

产生一个序列 $x_0,x_1,\cdots,x_n,\cdots$. 可以证明 $\lim\limits_{n\to\infty}x_n=\sqrt{2}$. 因此, 当 n 充分大, 如 $n=m$ 时, 终止计算. 这样, 我们取 x_m 作为 $\sqrt{2}$ 的近似值, 即 $\sqrt{2}\simeq x_m$. x_m 就是我们所需的计算结果.

为使一个数值方法在计算机上得到实现, 还需给出数值方法的**算法**, 它是用算术运算 (加、减、乘、除) 和逻辑运算完整地描述数值方法, 按一定顺序执行这些运算, 经有限步把输入数据转换成输出结果数据或信息. 例如, 我们用伪程序 (pseudocode) 给出计算 $\sqrt{2}$ 的 Newton 法的一种算法:

输入 初始近似值 x_0; 最大迭代次数 m.

输出 $\sqrt{2}$ 的近似值 p 或迭代失败信息.

step 1 $p_0\leftarrow x_0$.

step 2 对 $n=1,\cdots,m$ 做 step3 $\sim$ 4.

step 3 $p\leftarrow\left(p_0+\dfrac{2}{p_0}\right)/2$.

step 4 若 $|p-p_0|<10^{-8}$, 则输出 (p), 停机, 否则 $p_0\leftarrow p$.

step 5 输出 ('Method failed');
停机.

算法中 "←" 表示赋值. 若输出 "Method failed(方法失败)", 则表明迭代 m 次得到的 x_m, x_{m-1} 仍不满足 $|x_m - x_{m-1}| < 10^{-8}$.

读者很容易把伪程序描述的算法译成 FORTRAN、Pascal、C 语言或其他程序语言, 然后上机计算.

建立一个数值方法 (算法) 的基本原则应该是①便于在计算机上实现; ②计算工作量尽量小; ③存储量尽量小; ④问题准确解与计算解的误差小.

数值分析又称计算方法或数值计算方法. 它是研究运用计算机解数学问题的数值方法及其相关理论. 它是一门有丰富内容和自身理论体系的课程, 既有纯数学的系统性和严密性特点又与纯数学不同 (正如建立数值方法的基本原则所指出的).

1.2 误　　差

使用计算机求数学问题的数值解, 由于下面一些原因会产生误差.

(1) 数据误差. 用计算机进行数值计算时, 输入数据 (初始数据) 往往是近似的. 例如, $\pi = 3.14159265\cdots$, 在计算机上只能取有限位小数, 如取 $\pi \simeq 3.14159$. 这就产生了误差. 初始数据的这种误差称为**数据误差**. 有的输入数据是由实验或观测得到的. 由于观测手段的限制、测量仪器精密程度的影响, 得到的初始数据也会有一定的误差. 这种误差又称为**观测误差**.

(2) 截断误差. 求一级数的和或无穷序列的极限时, 我们取有限项作为它们的近似值, 它与级数和或极限之间的误差称为**截断误差**. 例如, 用 e^x 的幂级数展开式

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

计算 e^x 时, 取级数的前 $n+1$ 项的部分和 s_n 作为 e^x 的近似:

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

于是用 s_n 作为 e^x 的近似时就有截断误差:

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

它是由于截去 e^x 的幂级数展开的余项而产生的. 又如前述的用 Newton 法计算 $\sqrt{2}$, 我们取 $\sqrt{2} \simeq x_m$, 其截断误差是 $\sqrt{2} - x_m$.

(3) 离散误差. 在数值计算中, 我们常常用近似公式来求数学问题的近似解. 例如, 求曲边梯形 $abBA$(图 1.2.1) 的面积:

$$S = \int_a^b f(x)\mathrm{d}x.$$

若用梯形 $abBA$ 的面积

$$T=\frac{b-a}{2}(f(a)+f(b))$$

作为 S 的近似值, 则产生误差 $S-T$. 这种误差称为**离散误差**. 它是由于把连续型问题离散化而产生的.

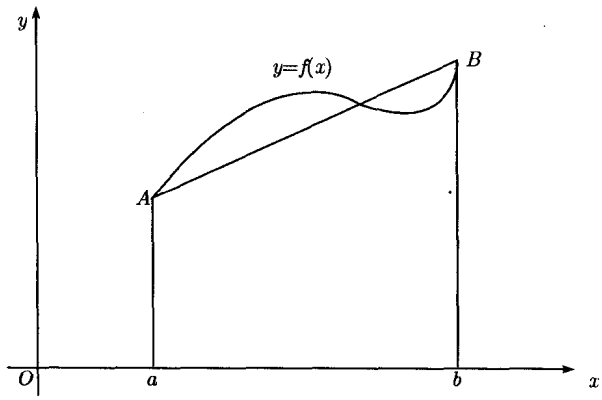

图 1.2.1

截断误差和离散误差统称为**方法误差**. 它是用数值方法求数学问题的近似解时由于使用近似公式导致数学问题的精确解与近似解之间产生的误差.

(4) 数值计算过程中的误差. 计算器或计算机只有有限位计算能力, 用数值方法解数学问题一般不能求得问题的准确解. 在进行数值计算过程中, 初始数据或计算得中间结果数据要用“四舍五入”或其他规则取近似值. 由此产生的误差称为**舍入误差**.

在一个数值方法中, 通常至少有上述一种误差出现. 在许多数值方法中, 会有上述多种误差出现.

在 1.1 节中我们提到, 用数值方法求数学问题的数值解时要求问题的数值解与精确解的误差越小越好, 即要求数值解的精确度越高越好. 因此, 我们首先要给出误差大小的度量. 有两种衡量误差大小的方法: 一是绝对误差, 二是相对误差.

假设某一量的准确值 (真值) 为 x, $\bar{x}$ 是 x 的一个近似值. 我们称 $\bar{x}$ 与 x 的差

$$e_{\bar{x}}=x-\bar{x}$$

为 $\bar{x}$ 的**绝对误差**(简称**误差**). 于是, 我们有

$$x=\bar{x}+e_{\bar{x}}.$$

通常, 我们不能算出准确值 x, 因此也不能算出绝对误差 $e_{\bar{x}}$ 的准确值, 只能估计误

差的大小范围. 若误差的绝对值不超过某一个正数 ε, 即

$$|e_{\bar{x}}| = |x - \bar{x}| \leqslant \varepsilon,$$

我们则称 ε 为近似值 $\bar{x}$ 的一个**绝对误差界**. 例如, $\pi = 3.14159265\cdots$, 取 $\bar{\pi} = 3.14$, 则

$$|\pi - \bar{\pi}| < 0.002.$$

绝对误差还不足以刻画近似数的精确程度. 例如, $x = 100$(厘米), $\bar{x} = 99$, 则 $e_{\bar{x}} = 1$; 而 $y = 10000$(厘米), $\bar{y} = 9950$, 则 $e_{\bar{y}} = 50$. 从表面上看, 后者的绝对误差是前者的 50 倍. 但是, 前者每厘米长度产生了 0.01 厘米的误差, 而后者每厘米长度只产生 0.005 厘米的误差. 因此, 要决定一个量的近似值的精确程度, 除了要看误差的大小外, 往往还要考虑该量本身的大小. 我们定义

$$r_{\bar{x}} = \frac{x - \bar{x}}{x}$$

为 $\bar{x}$ 的**相对误差**. 因为一个量的准确值往往是不知道的, 因此常常将 $\bar{x}$ 的相对误差 $r_{\bar{x}}$ 定义为

$$r_{\bar{x}} = \frac{x - \bar{x}}{\bar{x}}.$$

一般说来, 我们不能准确地计算出相对误差. 然而, 像绝对误差那样, 可以估计它的大小范围, 即指定一个正数 δ, 使得

$$|r_{\bar{x}}| \leqslant \delta.$$

我们称 δ 为 $\bar{x}$ 的一个**相对误差界**.

在数值分析中有关误差的讨论还会遇到有效数字的概念.

设实数 a 的近似数取为

$$a = \pm a_0 a_1 \cdots a_m . a_{m+1} \cdots a_n,$$

其中 $a_i (i = 0, 1, \cdots, n)$ 都是 $0, 1, \cdots, 9$ 中的一个数字, 且 $m \neq 0$ 时, $a_0 \neq 0$. 若 $\bar{a}$ 的绝对误差满足

$$|a - \bar{a}| \leqslant \frac{1}{2} \times 10^{-(n-m)}, \tag{1.2.1}$$

即不超过 $\bar{a}$ 的最末一位的半个单位, 则 $\bar{a}$ 的第一位 (自左至右) 非零数字 (设为 a_s) 到 $\bar{a}$ 的最末一位数字 a_n 中的 $n + 1 - s$ 位数字的每一个数字都称为 $\bar{a}$ 的**有效数字**, 并说近似数 $\bar{a}$ 是具有 $n + 1 - s$ 位有效数字的**有效数**.

例 1.2.1　设 π 的近似数取为 $\bar{\pi} = 3.1416$, 则 $|\pi - 3.1416| < 0.000008 < \frac{1}{2} \times 10^{-4}$. 因此 $\bar{\pi} = 3.1416$ 有 5 位有效数字, 其中 3, 1, 4, 1, 6 都是有效数字. 若取 $\bar{\pi} = 3.142$,

则 $|\pi - 3.142| < 0.0005 = \frac{1}{2} \times 10^{-3}$. 因此, $\bar{\pi} = 3.142$ 有 4 位有效数字, 其中 3, 1, 4, 2 都是有效数字.

例 1.2.2 设数 $a = 0.0330551$ 的近似值 $\bar{a} = 0.033056$, 则 $|a - \bar{a}| = 0.0000009 < \frac{1}{2} \times 10^{-5}$. 因此 $\bar{a}$ 有 5 位有效数字, 其中 3, 3, 0, 5, 6 都是有效数字.

例 1.2.3 有效数 100.00 与有效数 100 不同. 前者具有 5 位有效数字: 1, 0, 0, 0, 0, 而后者只有 3 位有效数字: 1, 0, 0.

定理 1.2.1 设数 a 的近似数

$$\bar{a} = \pm a_0 a_1 \cdots a_m . a_{m+1} \cdots a_n \tag{1.2.2}$$

具有 $n+1-s$ 位有效数字, 则其相对误差有估计式

$$\left| \frac{a - \bar{a}}{\bar{a}} \right| \leqslant \frac{1}{2a_s} \times 10^{-(n-s)}, \tag{1.2.3}$$

其中 $a_s \neq 0$ 是 $\bar{a}$ 的第一位有效数字.

证明 首先, 若 $s = 0$, 此时 $a_0 \neq 0$, 由 (1.2.2) 知

$$|\bar{a}| \geqslant a_0 \times 10^m.$$

再由 (1.2.1) 式有

$$\left| \frac{a - \bar{a}}{\bar{a}} \right| \leqslant \frac{1}{a_0 \times 10^m} \times \frac{1}{2} \times 10^{-(n-m)} = \frac{1}{2a_0} \times 10^{-n}.$$

其次, 若 $s \neq 0, a_s \neq 0$, 此时 $m = 0$, 由 (1.2.2) 式知

$$|\bar{a}| \geqslant a_s \times 10^{-s}.$$

再由 (1.2.1) 式有

$$\left| \frac{a - \bar{a}}{\bar{a}} \right| \leqslant \frac{1}{a_s \times 10^{-s}} \times \frac{1}{2} \times 10^{-n} = \frac{1}{2a_s} \times 10^{-(n-s)}.$$

这个定理表明, 一个近似数的有效数字越多, 其相对误差则越小, 因而精确度越高. 就例 1.2.3, 有效数 100.00 的精确度比 100 的精确度高. 据 (1.2.3) 式, 前者的相对误差界为 $\frac{1}{2} \times 10^{-4}$, 而后者的相对误差界为 $\frac{1}{2} \times 10^{-2}$.

1.3 浮点运算和舍入误差

1.3.1 计算机中数的表示

计算器或计算机只能执行有限位数字的数的算术运算 (加、减、乘、除运算).

假设提供给计算机一个只有有限位小数的数 x, 它可表示成

$$x = \pm 10^J \sum_{k=1}^{t} d_k 10^{-k}, \tag{1.3.1}$$

其中 J 是整数, $d_1, d_2, \cdots, d_t$ 都是 $0, 1, 2, \cdots, 9$ 中的一个数字. 例如:

$$2004.09 = 10^4(2 \times 10^{-1} + 0 \times 10^{-2} + 0 \times 10^{-3} + 4 \times 10^{-4} + 0 \times 10^{-5} + 9 \times 10^{-6}),$$

$$-0.00105 = -10^{-2}(1 \times 10^{-1} + 0 \times 10^{-2} + 5 \times 10^{-3}).$$

若记

$$a = \sum_{k=1}^{t} d_k 10^{-k} = 0.d_1 d_2 \cdots d_t, \tag{1.3.2}$$

则

$$x = \pm a \times 10^J. \tag{1.3.3}$$

例如:

$$2004.09 = 0.200409 \times 10^4, \qquad -0.00105 = -0.105 \times 10^{-2}.$$

(1.3.1) 或 (1.3.3) 式是通常的数的十进制系统计数法, 其中的 10 称为十进制系统的**基数**. 在计算机中, 还广泛采用二进制、八进制和十六进制系统表示数的方法, 它们的基数分别为 2, 8 和 16.

一般地, 一个 p 进制数 x 可以表示成

$$x = \pm p^J \sum_{k=1}^{t} d_k p^{-k}, \tag{1.3.4}$$

其中 $d_k (k = 1, 2, \cdots, t)$ 都是 $0, 1, \cdots, p-1$ 中的一个数字. (1.3.4) 式或写成

$$x = \pm a \times p^J, \tag{1.3.5}$$

其中

$$a = \sum_{k=1}^{t} d_k p^{-k} = 0.d_1 d_2 \cdots d_t. \tag{1.3.6}$$

我们称 a 为数 x 的**尾数**(其值小于 1). 自然数 t 为计算机的**字长**, 它表示数 x 的尾数的位数. J 是整数, 称为数 x 的**阶**, 它用来确定该数的小数点的位置.

在各种计算机中, 有各自规定的字长 t, 以及阶 J 的范围: $-L \leqslant J \leqslant U$($L$ 和 U 为正整数或零). L, U 的大小表明计算机中表示的数的范围大小. 我们称数 x 的表示式 (1.3.4) 或 (1.3.5) 为 x 的**浮点**表示. 再假设 $x \neq 0$ 时, $d_1 \neq 0$, 则称 x 为**规格化浮点数**.

按 (1.3.4) 或 (1.3.5) 规定的规格化浮点数的全体组成的集合记作 F. 我们称 F 为一个**规格化浮点数系**. 规格化浮点数系 F 是一个有限的离散的数集合. 在数值计算中通常取基数 $p=10$.

例 1.3.1 设 $p=10$, $t=4$, $L=U=5$, 则规格化浮点数系 F 中最大的数是 $0.9999\times10^5=99990$, 最小的数是 $-0.9999\times10^5=-99990$, 除数 0 外绝对值最小的数是 $\pm0.1000\times10^{-5}=\pm10^{-6}$.

1.3.2 浮点运算和舍入误差

规格化浮点数系 F 是一个离散的有限集合. 在使用计算机进行数值计算时, 若提供给计算机的数 x 的绝对值大于 F 中最大正数, 则在计算机上出现**上溢**; 若 $x(\neq0)$ 的绝对值小于 F 中最小正数, 则在计算机上出现**下溢**, 此时计算机将数 x 作为 0 处理, 称其为**机器 0**. 上溢和下溢统称为**溢出**. 在以后的讨论中, 我们将假定所提供的初始数据或中间结果数据不会发生溢出现象. 初始数据或中间计算结果数据 x 可能不在规格化浮点数系 F 中, 因此要用 F 中最接近 x 的数 x_R 作为 x 的近似值. 现设十进制数 $x\in F$, 则可按舍入 (四舍五入) 的规则取 x_R, 即若

$$x=\pm a\times10^J, \tag{1.3.7}$$

其中

$$a=0.d_1\cdots d_td_{t+1}\cdots d_n\cdots,\quad 0\leqslant d_i\leqslant9\quad(i=1,\cdots,n,\cdots),\quad d_1>0, \tag{1.3.8}$$

则取

$$\bar a=\begin{cases}0.d_1\cdots d_t, & 0\leqslant d_{t+1}\leqslant4,\\ 0.d_1\cdots d_t+10^{-t}, & d_{t+1}\geqslant5,\end{cases} \tag{1.3.9}$$

$$x_R=\pm\bar a\times10^J. \tag{1.3.10}$$

于是有

$$\left|\frac{x_R-x}{x}\right|=\left|\frac{\pm\bar a\times10^J-(\pm a)\times10^J}{\pm a\times10^J}\right|$$
$$=\left|\frac{\bar a-a}{a}\right|.$$

由于

$$a\geqslant10^{-1},\qquad|\bar a-a|\leqslant\frac12\times10^{-t},$$

因此

$$\left|\frac{\bar a-a}{a}\right|\leqslant\frac12\times10^{-t+1}=5\times10^{-t},$$