21 世纪高等院校教材

数 值 分 析 原 理

封建湖　车刚明　聂玉峰　编著

科 学 出 版 社

北　京

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 世纪高等院校教材)
ISBN 978-7-03-009730-9

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

责任编辑:胡华强 杨 波/责任校对:陈玉凤
责任印制:张克忠/封面设计:王 浩

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

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:B5(720×1000)
2007 年 6 月第八次印刷 印张:21 1/4
印数:17 001—19 500 字数:378 000

定价:25.00 元
(如有印装质量问题,我社负责调换〈路通〉)

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未来研究工作的需要,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.对于其他同类书籍中没有给予足够重视,但根据作者多年的计算经验证明是非常有效的、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 月于西安

目　　录

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.

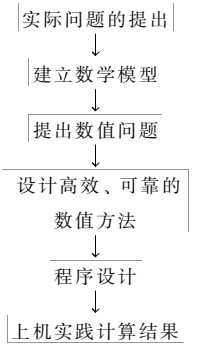

图 1.1

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

$$\begin{cases} \dfrac{\mathrm{d}y}{\mathrm{d}x} = x + y^2 & x \in [0,1], \\ y(0) = y_0, \end{cases}$$

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)方法求得其解.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，因而掌握数值方法的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^*.如果存在正数 $\varepsilon^*=\varepsilon(x^*)$,使得绝对误差 $|e^*|=|x^*-x|\leqslant\varepsilon^*$,则称 ε^* 为 x^* 近似 x 的一个**绝对误差限**,简称误差限.

此时有 $x\in[x^*-\varepsilon^*,x^*+\varepsilon^*]$,工程上习惯用 $x=x^*\pm\varepsilon^*$ 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 $x=100\pm1$ 和 $y=10000\pm5$,尽管对 x 和 y 的测量绝对误差限满足 $\varepsilon(x^*)=1<5=\varepsilon(y^*)$,但绝对误差限在真值中所占的比例却有不等关系:

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

因此,我们并不认为测量 $x=100\pm1$ 比 $y=10000\pm5$ 更精确.

定义 1.2　设 x^* 是准确值 $x(\neq 0)$ 的一个近似，称 $e_r(x^*)=\dfrac{x^*-x}{x}$ 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^*.

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

即 e_r^* 与 $\dfrac{e^*}{x^*}=\dfrac{x^*-x}{x^*}$ 相差一个较 e_r^* 高一阶的无穷小量，故有时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 $\varepsilon_r^*=\varepsilon^*/x^*$ 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 $\varepsilon_r(y^*)<\varepsilon_r(x^*)$ 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：

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

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

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

四舍情形下的误差限

$$|x^*-x|=10^m\times\underline{0.\overset{n\text{个}0}{00\cdots0}x_{n+1}\cdots}\leqslant 10^m\times 0.\overset{n\text{个}0}{00\cdots0}5=\frac{1}{2}\times 10^{m-n}.$$

五入情形下的误差限

$$|x^*-x|=10^m\times\left|\overset{n-1\text{个}0}{0.00\cdots01}-\underline{0.\overset{n\text{个}0}{00\cdots0}x_{n+1}\cdots}\right|$$

$$=10^{m-n}\times|1-\underline{0.x_{n+1}}\cdots|\leqslant\frac{1}{2}\times 10^{m-n}.$$

综合以上两点,“四舍五入”法的误差限是$\frac{1}{2}\times10^{m-n}$. #

定义 1.3 设 x 的近似值 x^* 有如下标准形式

$$x^*=\pm10^m\times0.\underline{x_1\,x_2\cdots x_n\cdots x_p}, \tag{1.3}$$

其中 m 为整数,$\{x_i\}\subset\{0,1,2,\cdots,9\}$且 $x_1\neq0$, $p\geqslant n$. 如果有

$$|e^*|=|x^*-x|\leqslant\frac{1}{2}\times10^{m-n},$$

则称 x^* 为 x 的具有 ***n* 位有效数字**的近似数,或称 x^* 精确到小数点后第 n 位,其中数字 $x_1, x_2,\cdots, x_n$ 分别被称为 x^* 的第一、第二、…、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,它们是有效数吗?

解 真值 $x=12.49=10^2\times0.1249$, $m=2$.

$$|e_1^*|=|x_1^*-x|=0.01\leqslant\frac{1}{2}\times10^{-1}=\frac{1}{2}\times10^{2-3},$$

$$|e_2^*|=|x_2^*-x|=0.09\leqslant\frac{1}{2}\times10^{0}=\frac{1}{2}\times10^{2-2},$$

$$|e_3^*|=|x_3^*-x|=0.01\leqslant\frac{1}{2}\times10^{-1}=\frac{1}{2}\times10^{2-3},$$

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 $|e_r^*|\leqslant\frac{1}{2x_1}\times10^{1-n}$；

2° 若相对误差 $|e_r^*|\leqslant\frac{1}{2(x_1+1)}\times10^{1-n}$，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 $|e^*|\leqslant\frac{1}{2}\times10^{m-n}$. 而相对误差

$$|e_r^*|=\left|\frac{e^*}{x^*}\right|\leqslant\frac{1}{2|x^*|}\times10^{m-n}$$

$$\leqslant\frac{1}{2\times10^m\times\underline{0.x_1}}\times10^{m-n}=\frac{1}{2x_1}\times10^{1-n}.$$

2° 绝对误差

$$|e^*|=|e_r^*|\cdot|x^*|\leqslant\frac{1}{2(x_1+1)}\times10^{1-n}\times10^m\times\underline{0.x_1x_2\cdots x_n}$$

$$\leqslant\frac{1}{2(x_1+1)}\times10^{m+1-n}\times\underline{0.(x_1+1)}=\frac{1}{2}\times10^{m-n},$$

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 $\frac{1}{2}\times10^{-3}$，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

$$|e_r^*|\leqslant\frac{1}{2x_1}\times10^{1-n}=\frac{1}{8}\times10^{1-n}.$$

令 $\frac{1}{8}\times10^{1-n}\leqslant\frac{1}{2}\times10^{-3}$，解得 $n\geqslant3.4$. 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

截断误差;初始数据在计算机内的浮点表示一般有舍入误差;每次运算一般又会产生新的舍入误差,并传播以前各步已经引入的误差;考虑到误差有正有负,误差积累的过程一般包含有误差增长和误差相消的过程,并非简单的单调增长;运算次数非常之多,不可能人为地跟踪每一步运算.这些因素注定了对误差进行准确估计是很困难的.

本小节中,在每一步都是准确计算的假设下,即不考虑截断误差和由运算进一步引入的舍入误差,介绍分析初值误差传播规律的泰勒(Taylor)方法和区间分析法,然后引入坏函数值点的概念.

1. 用泰勒公式分析初值的误差传播规律

设可微函数 $y=f(x_1,x_2,\cdots,x_n)$中的自变量 x_1、x_2、…、x_n 是相互独立的.用它们的近似值进行计算,得到函数值 y 的近似值 $y^*=f(x_1^*,x_2^*,\cdots,x_n^*)$.

当 x_1^*、x_2^*、…、x_n^* 很好地近似了相应的真值时,利用多元函数的一阶泰勒公式可求得 y^* 的绝对误差和相对误差分别为①

$$\begin{aligned} e(y^*) &= y^*-y \approx \sum_{i=1}^{n} f'_i(x_1^*,\cdots,x_n^*)(x_i^*-x_i) \\ &= \sum_{i=1}^{n} f'_i(x_1^*,\cdots,x_n^*)e(x_i^*), \end{aligned} \tag{1.4}$$

$$\begin{aligned} e_r(y^*) &= \frac{e(y^*)}{y^*} \approx \sum_{i=1}^{n} \frac{x_i^*}{y^*} f'_i(x_1^*,\cdots,x_n^*)\frac{e(x_i^*)}{x_i^*} \\ &= \sum_{i=1}^{n} \frac{x_i^*}{y^*} f'_i(x_1^*,\cdots,x_n^*)e_r(x_i^*). \end{aligned} \tag{1.5}$$

进而得到如下绝对误差限和相对误差限的传播关系:

$$\varepsilon(y^*) \approx \sum_{i=1}^{n} \left| f'_i(x_1^*,\cdots,x_n^*) \right| \varepsilon(x_i^*), \tag{1.6}$$

$$\varepsilon_r(y^*) \approx \sum_{i=1}^{n} \left| \frac{x_i^*}{y^*} f'_i(x_1^*,\cdots,x_n^*) \right| \varepsilon_r(x_i^*). \tag{1.7}$$

① 当 $f'_i(x_1^*,x_2^*,\cdots,x_n^*)(1\leqslant i\leqslant n)$的绝对值均很小时(如驻点),需要使用二阶的泰勒公式分析误差传播状况.

由式(1.6)可得到二元函数算术运算(+ − × ÷)的误差限传播不等式：

$$\varepsilon(x_1^* \pm x_2^*) \approx \varepsilon(x_1^*) + \varepsilon(x_2^*), \tag{1.8}$$

$$\varepsilon(x_1^* x_2^*) \approx |x_2^*|\varepsilon(x_1^*) + |x_1^*|\varepsilon(x_2^*), \tag{1.9}$$

$$\varepsilon\left(\frac{x_1^*}{x_2^*}\right) \approx \frac{|x_2^*|\varepsilon(x_1^*) + |x_1^*|\varepsilon(x_2^*)}{|x_2^*|^2} \quad (x_2^* \neq 0). \tag{1.10}$$

由式(1.7)、(1.9)和(1.10)可得到二元函数算术运算的相对误差限传播关系式：

$$\varepsilon_r(x_1^* + x_2^*) \approx \max\{\varepsilon_r(x_1^*), \varepsilon_r(x_2^*)\} \quad (x_1^* x_2^* > 0), \tag{1.11}$$

$$\varepsilon_r(x_1^* x_2^*) \approx \varepsilon_r(x_1^*) + \varepsilon_r(x_2^*) \quad (x_1^* x_2^* \neq 0), \tag{1.12}$$

$$\varepsilon_r\left(\frac{x_1^*}{x_2^*}\right) \approx \varepsilon_r(x_1^*) + \varepsilon_r(x_2^*) \quad (x_1^* x_2^* \neq 0). \tag{1.13}$$

这里需要说明的是，对于具体的一组数据，上面给出的误差限传播公式是实际误差的一个粗糙偏大的估计. 如对于加法运算的估计(1.8)式，它包括了误差源 $e(x_1^*)$和 $e(x_2^*)$同号且同时均达到了误差限这一最坏的情况，实际情况往往并非这么坏.

下面通过算例考察相近数相减情形下的误差传播.

例 1.4　已知 $x_1 = 100.002$ 和 $x_2 = 100$ 的近似值 $x_1^* = 100.003$ 和 $x_2^* = 99.999$. 试分析函数 $y = x_1 - x_2$ 的近似值 $y^* = x_1^* - x_2^*$ 的绝对误差、相对误差和有效数字.

解　$y = x_1 - x_2 = 0.2 \times 10^{-2}$，　$y^* = x_1^* - x_2^* = 0.004$，

$$e(y^*) = y^* - y = 0.002, \quad e_r(y^*) = \frac{y^* - y}{y} = 100\%.$$

由 $|e(y^*)| = 0.002 \leqslant 0.005 = \frac{1}{2} \times 10^{-2} = \frac{1}{2} \times 10^{-2-0}$ 知 y^* 有零位有效数字. #

在这一算例中，虽然相近数 x_1^* 和 x_2^* 均有 5 位有效数字，但它们的差却连一位有效数字也没有. 可见，在数值计算中应该尽量避免相近数相减. 如当正数 x 充分大时，可按如下方法变换算式以提高数值计算的精度：

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

$$\sqrt{x+1}-\sqrt{x}=\frac{1}{\sqrt{x+1}+\sqrt{x}},$$

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

$$\ln(x-\sqrt{x^2-1})=-\ln(x+\sqrt{x^2-1}).$$

当$|x|$充分小时,可使用如下变换

$$1-\sqrt{1-x^2}=\frac{x^2}{1+\sqrt{1-x^2}},$$

$$\arctan x-x=-\frac{x^3}{3}+\frac{x^5}{5}-\cdots,$$

$$\sin x-x=-\frac{x^3}{3!}+\frac{x^5}{5!}-\cdots.$$

提高相关算式的计算精度.

*2.区间分析法

在实际问题中,常常知道初始数据的误差范围,我们可以用区间分析法分析初值误差的传播规律.

设 $x\in[a_1,b_1]$, $y\in[a_2,b_2]$,则有如下结论成立:

1° $x+y\in[a_1+a_2,b_1+b_2]$;

2° $-x\in[-b_1,-a_1]$;

3° $1/x\in[1/b_1,1/a_1]$, $a_1b_1>0$;

4° $xy\in[\min\Xi,\max\Xi]$, $\Xi=\{a_1a_2,a_1b_2,b_1a_2,b_1b_2\}$.

类比于如上结论,我们规定如下的区间运算规则:

1° $[a_1,b_1]+[a_2,b_2]=[a_1+a_2,b_1+b_2]$;

2° $-[a_1,b_1]=[-b_1,-a_1]$;

3° $1/[a_1,b_1]=[1/b_1,1/a_1]$, $a_1b_1>0$;

4° $[a_1,b_1]\times[a_2,b_2]=[\min\Xi,\max\Xi]$.

又考虑到 $x-y=x+(-y)$, $x\div y=x\times\frac{1}{y}$,定义

5° $[a_1,b_1]-[a_2,b_2]=[a_1,b_1]+(-[a_2,b_2])$

$$=[a_1-b_2,b_1-a_2];$$

6°　$[a_1,b_1]/[a_2,b_2]=[a_1,b_1]\times(1/[a_2,b_2])$

$$=[a_1,b_1]\times\left[\frac{1}{b_2},\frac{1}{a_2}\right],a_2b_2>0.$$

对于精确数 p 和近似数 x 之间的算术运算$\circ$(代表加、减、乘、除中任一种运算)有

7°　$p\circ x\in[p,p]\circ[a_1,b_1]$.

所谓区间分析法就是应用如上一些区间运算法则进行误差分析,具体操作过程可参阅如下例子:

例 1.5　已知方程组$\begin{cases}3x+ay=10\\5x+by=20\end{cases}$的求解算法 $y=\dfrac{10}{3b-5a}$, $x=\dfrac{10-ay}{3}$. $a=2.100\pm\dfrac{1}{2}\times10^{-3}$, $b=3.300\pm\dfrac{1}{2}\times10^{-3}$,试估计计算 x,y 的误差范围.

解　记 $\varepsilon=\dfrac{1}{2}\times10^{-3}$,于是有

$$a\in[2.100-\varepsilon,2.100+\varepsilon],\qquad b\in[3.300-\varepsilon,3.300+\varepsilon].$$

由

$$3b-5a\in[9.9-3\varepsilon,9.9+3\varepsilon]-[10.5-5\varepsilon,10.5+5\varepsilon]$$

$$=[-0.6-8\varepsilon,-0.6+8\varepsilon],$$

得

$$y=\frac{10}{3b-5a}\in\frac{10}{[-0.6-8\varepsilon,-0.6+8\varepsilon]}$$

$$=\left[\frac{10}{-0.6+8\varepsilon},\frac{10}{-0.6-8\varepsilon}\right]\approx[-16.7785,-16.5563],$$

故 $y=-16.6674\pm0.1111$.

由

$$-\frac{ay}{3}\in-\left[\frac{2.1-\varepsilon}{3},\frac{2.1+\varepsilon}{3}\right]\times[-16.7785,-16.5563]$$

$$\approx[0.6998,0.7002]\times[16.5563,16.7785]$$

$$\approx[11.5861,11.7483],$$

得

$$x=\frac{10-ay}{3}\in[14.9194,15.0816],$$

故 $x=15.0005\pm0.0811$. #

3. 计算函数值的条件数

设 x^* 是 x 的较好近似,由微分中值定理知,可微函数 $f(x)$在这两点的函数值之差满足

$$\begin{aligned} f(x^*)-f(x)&=f'(x+\theta(x^*-x))(x^*-x),\quad 0<\theta<1\\ &\approx f'(x)(x^*-x), \end{aligned} \tag{1.14}$$

即有

$$e(f(x^*))\approx f'(x)e(x^*). \tag{1.15}$$

上式反应了函数值绝对误差与自变量绝对误差之间的关系,并且有如下结论:当 $|f'(x)|<1$ 时,函数值的扰动比自变量的微小变化还要小;而当 $|f'(x)|$很大时,自变量的微小变化,将引起函数值较大的扰动,此时,称 x 是函数 f 在绝对误差意义下的**坏函数值点**.

从(1.15)式可以推出函数值相对误差与自变量相对误差之间的如下联系

$$e_r(f(x^*))=\frac{e(f(x^*))}{f(x)}\approx x\frac{f'(x)}{f(x)}e_r(x^*). \tag{1.16}$$

这一近似等式表明:当 $\left|x\dfrac{f'(x)}{f(x)}\right|$ 很大时,自变量的微小变化,将引起函数值较大的扰动,此时,称 x 是函数 f 在相对误差意义下的**坏函数值点**.

基于如上分析,我们称 $|f'(x)|=\mathrm{cond}_a(f)$和 $\left|x\dfrac{f'(x)}{f(x)}\right|=\mathrm{cond}_r(f)$分别为在绝对误差意义下和相对误差意义下在 x 点**计算函数值的条件数**.它是函数自身在点 x 处固有的特征. 对于相同的自变量扰动,条件数越大,计算出的函数值误差越大. 要提高函数值的计算精度,通常只有通过提高初值精度来实现.

§1.3 舍入误差分析及数值稳定性

1.3.1 浮点数系及其运算的舍入误差

计算机中通常配置有两种类型的算术运算:定点数运算和浮点数运算. 所谓点是指小数点,用浮点数进行计算是指用常数个**数字**进行工作;而用定点数进行计算是指用常数个**小数**位数进行工作. 这里我们仅介绍较多使用的浮点数系及其运算的舍入误差.

1. 浮点数系以及舍入误差的产生

一个浮点数的表示由正负号、小数形式的尾数、以及确定小数点位置的阶三部分组成. 单精度实数用 32 位的二进制数据表示浮点数的这三个信息，其中数值符号占 1 位，尾数占 23 位，阶数占 8 位.

对于规范化的浮点数(除零外)，23 位的二进制尾数形式是：

$$(0.\underline{1\,\alpha_2\,\alpha_3\cdots\alpha_{23}})_2 = 2^{-1} + \sum_{i=2}^{23} \alpha_i 2^{-i}, \quad \alpha_i \in \{0,1\},$$

式中 2^{-i} 表示尾数中小数点后第 i 位的权. 当尾数的首位小于 5 时，可通过不断乘以 2 使之首位大于或等于 5，相应的二进制阶数需要减去乘以 2 的次数.

在 8 位的阶数中，有 1 位表示阶数的符号，7 位表示二进制的阶数数值，于是阶数数值的范围是 $0 \sim 2^7 - 1$.

综合上面关于阶数和尾数的讨论，单精度实数集合为

$$R_S = \{0\} \cup \left\{ a \,\middle|\, a = \pm 2^p \left(2^{-1} + \sum_{i=2}^{23} \alpha_i 2^{-i} \right), p \in Z \text{ 且 } |p| \leqslant 2^7 - 1 \right\}.$$

其中元素是能够准确表示的数，称之为机器数.

设 $a \in R_S$，与之相邻的能够准确表示的机器数是 $b = a + 2^{p-23}$ 和 $c = a - 2^{p-23}$. 这样，在区间 (c,a) 和 (a,b) 上的实数无法准确表示. 通常计算机系统规定：不能精确表示的实数用与之最近的机器数表示①. 我们将实数 x 在机器中的浮点(float)表示记为 $fl(x)$. 将由此表示产生的误差 $fl(x) - x$ 称之为**舍入误差**. 如当 $x \in \left[\frac{c+a}{2}, \frac{a+b}{2} \right) = [a - 2^{p-1-23}, a + 2^{p-1-23})$ 时用 a 表示，即有 $fl(x) = a$. 相对误差 e_r^* 满足：

$$|e_r^*| = \left| \frac{fl(x) - x}{fl(x)} \right| \leqslant \frac{2^{p-1-23}}{2^{p-1}} = 2^{-23} \approx 10^{-6.923}.$$

上式表明单精度实数 $fl(x)$ 能够有 6～7 位有效数字.

二进制阶数数值上限 $2^7 - 1$ 相应于十进制的阶数数值上限是 $38((2^7-1)\cdot \lg 2 \approx 38.23)$. 结合阶数的符号，除零外，单精度实数的量级不大于 10^{38} 不小于 10^{-38}. 当输入、输出或中间计算过程中出现量级大于 10^{38} 的数据时，因单精度实数无法正确表示该数据，将导致程序的非正常停止，称此现象为上溢

① 当某一实数距离两个机器数同样近时，为了表示的惟一性，还需要附加一条其他规定，并且随机器系统的不同而有所差异.

(overflow). 而当出现量级小于 10^{-38}的非零数据时，一般计算机将该数置为零，精度损失，称此现象为下溢(underflow). 当数据有可能出现上溢或下溢时，可通过乘积因子变换数据，使之正常表示.

一般地，设在某一浮点系统中，尾数占 t 位二进制数(未计算尾数的符号位)，阶数占 s 位二进制数(未计算阶数的符号位)，实数 x 的浮点表示 $fl(x)$共需要 $t+s+2$ 位的二进制数位. 当不出现溢出时，绝对误差 e^* 和相对误差 e_r^* 分别满足如下估计：

$$|e^*|=|fl(x)-x|\leqslant 2^{p-1}\cdot 2^{-t}, \tag{1.17}$$

$$|e_r^*|=\left|\frac{fl(x)-x}{fl(x)}\right|\leqslant\frac{2^{p-1-t}}{2^{p-1}}=2^{-t}, \tag{1.18}$$

其中 p 由 $2^{p-1}\leqslant|x|<2^p$ 确定.上溢界和下溢界分别是 $2^{2^{s}-1}=10^{(2^{s}-1)\lg 2}$和 $2^{-2^{s}}=10^{-2^{s}\lg 2}$. 对于单精度实数有 $t=23, s=7$.

2. 浮点运算舍入误差分析

定理 1.2 设实数 x 满足 $2^{p-1}\leqslant|x|<2^p$ 且 $|p|\leqslant 2^s-1$，则 x 的浮点表示 $fl(x)$满足

$$fl(x)=x(1+\delta),\quad |\delta|\leqslant 2^{-t},$$

其中 s, t 分别为浮点系统中给二进制阶数数值以及尾数数值的表示所分配的二进制数位.

证明 设 $fl(x)=x(1+\delta)$，则有 $\delta=\dfrac{fl(x)-x}{x}$，进而由式(1.17)得

$$|\delta|=\left|\frac{fl(x)-x}{x}\right|\leqslant\frac{2^{p-1-t}}{2^{p-1}}=2^{-t}. \qquad \#$$

我们用符号$\circ$表示加减乘除四种算术运算之一，并将浮点数 $fl(x)$与 $fl(y)$的算术运算理想地简化为：首先计算出 $fl(x)\circ fl(y)$的精确值[①]，然后用浮点数表示 $fl(fl(x)\circ fl(y))$.这样便由定理 1.2 得到如下推论.

推论 1.1 $fl(fl(x)\circ fl(y))=(fl(x)\circ fl(y))(1+\delta), |\delta|\leqslant 2^{-t}$.

利用该推论可以推导出算术表达式求值的误差界.

例 1.6 对三同号数的算术运算 $a+b+c$ 作舍入误差分析.

解 这里对运算$(a+b)+c$作误差分析.

① 这一理想简化的依据是：CPU 中的运算器能够精确到较浮点数系更多的数位.

$$fl(fl(a)+fl(b))=(fl(a)+fl(b))(1+\delta_3)$$
$$=(a(1+\delta_1)+b(1+\delta_2))(1+\delta_3),$$
$$fl(fl(fl(a)+fl(b))+fl(c))$$
$$=\{fl(fl(a)+fl(b))+fl(c)\}(1+\delta_5)$$
$$=\{[a(1+\delta_1)+b(1+\delta_2)](1+\delta_3)+c(1+\delta_4)\}(1+\delta_5)$$
$$=a+b+c+a(\delta_1+\delta_3+\delta_1\delta_3+\delta_5+\delta_1\delta_5+\delta_3\delta_5+\delta_1\delta_3\delta_5)+b(\delta_2+\delta_3+\delta_2\delta_3+\delta_5+\delta_2\delta_5+\delta_3\delta_5+\delta_2\delta_3\delta_5)+c(\delta_4+\delta_5+\delta_4\delta_5).$$

设 $|\delta_i|\leqslant\varepsilon\leqslant 2^{-t}$, $i=1,2,3,4,5$. 于是得到

$$|fl(fl(fl(a)+fl(b))+fl(c))-(a+b+c)|$$
$$\leqslant(|a|+|b|)(3\varepsilon+3\varepsilon^2+\varepsilon^3)+|c|(2\varepsilon+\varepsilon^2).$$　　#

从上述例题的结果可以看出:浮点机器数的加法并不一定满足结合律,先加绝对值较小的两数,然后再和另外一数相加,将会有较小的舍入误差.这一事实更为深刻的意义在于:**数学上等价的算法在数值上并不总是等效的**.

例 1.6 分析误差的思路是:首先论证近似计算 $f^*(a_1^*,a_2^*,\cdots,a_m^*)=f(a_1+\delta a_1,a_2+\delta a_2,\cdots,a_m+\delta a_m)$,然后估计出摄动量 $\delta a_i(i=1,2,\cdots,m)$ 的大小,进而得到 $|f^*(a_1^*,a_2^*,\cdots,a_m^*)-f(a_1,a_2,\cdots,a_m)|$ 的估计. 这种将误差估计转化为原始数据摄动的方法,称之为**向后误差分析法**.

我们通常的思路是:对每一步找出舍入误差界,随着计算过程逐步向前分析,直到估计出最后结果的误差界,这一方法称之为**向前误差分析法**.

1.3.2　算法的数值稳定性

上一例题定量地分析了舍入误差的积累效应,从其过程可以看到,舍入误差分析是非常繁杂困难的,而舍入误差不可避免,运算量又相当大,为此,人们提出了"数值稳定性"这一概念对舍入误差是否影响产生可靠的结果进行定性的分析.

一个算法,如果在运算过程中舍入误差在一定条件下能够得到控制,或者舍入误差的增长不影响产生可靠的结果,则称该算法是**数值稳定**的,否则称其为数值不稳定.

下面讨论两种算法计算积分 $I_n=\int_0^1\frac{x^n}{x+5}\mathrm{d}x$ 的数值稳定性.

由 $I_n=\int_0^1\frac{x+5-5}{x+5}x^{n-1}\mathrm{d}x=\int_0^1 x^{n-1}\mathrm{d}x-5I_{n-1}=\frac{1}{n}-5I_{n-1}$ 得到递推

公式 $I_n = \frac{1}{n} - 5I_{n-1}(n = 1,2,\cdots)$，而 $I_0 = \int_0^1 \frac{1}{x+5}\mathrm{d}x = \ln\frac{6}{5} \approx 0.1823$. 小数点后保留 4 位小数，用该公式计算出前十个数据的近似结果以及绝对误差参见表 1.1

由上述递推公式可以得到变形公式：$I_{n-1} = \frac{1}{5n} - \frac{I_n}{5}$，而

$$\frac{1}{6(n+1)} = \int_0^1 \frac{x^n}{6}\mathrm{d}x \leqslant I_n \leqslant \int_0^1 \frac{x^n}{5}\mathrm{d}x = \frac{1}{5(n+1)},$$

结合如上估计以及变形公式得到计算积分的第二种方法. 取 $I_{10}^* = \frac{1}{2}\left(\frac{1}{55} + \frac{1}{66}\right) \approx 0.0167$，相关计算结果也参见表 1.1.

表 1.1　两种计算积分方法的计算结果比较

n	$I_n^* = 1/n - 5I_{n-1}^*$		$I_{n-1}^* = (1/n - I_n^*)/5$	
	I_n^*	$\lvert I_n^* - I_n\rvert$	I_n^*	$\lvert I_n^* - I_n\rvert$
0	0.1823	0.00002	0.1823	0.2156×10^{-6}
1	0.0885	0.0001	0.0884	0.7784×10^{-7}
2	0.0575	0.0005	0.0580	0.3892×10^{-6}
3	0.0458	0.0027	0.0431	0.3873×10^{-6}
4	0.0208	0.0135	0.0343	0.6330×10^{-7}
5	0.0958	0.0673	0.0285	0.3165×10^{-6}
6	−0.3125	0.3368	0.0243	0.2491×10^{-6}
7	1.7054	1.6842	0.0212	0.3262×10^{-6}
8	−8.4018	8.4206	0.0189	0.6308×10^{-6}
9	42.1200	42.1031	0.0167	0.2265×10^{-5}
10	−210.5002	210.5156	0.0167	0.1332×10^{-4}

第一种方法的计算结果以及计算公式均表明，舍入误差的传播依 5 的幂次进行增长，因而是一种不稳定的方法. 第二种方法的舍入误差在一定范围内依$\frac{1}{5}$的幂次进行传播，随着计算的深入误差不仅没有增长而且越来越小，因而是一种稳定的方法. 值得注意的是，后者方法的绝对误差没有随着计算的进一步进行而趋于零，其原因是计算递推过程中仅仅保留了 4 位小数.

总之，除了算法的正确性之外，在算法设计中至少还应注意如下几个方

面的问题：

1. 尽量避免两个相近的数相减；
2. 合理安排量级相差很大的数之间的运算次序，防止大数“吃掉”小数；
3. 尽可能避免绝对值很小的数做分母；
4. 防止出现溢出；
5. 简化计算步骤以减少运算次数；
6. 选用数值稳定性好的算法.

习 题 1

1. 请指出如下有效数的绝对误差限、相对误差限和有效数字位数：
$$49\times10^{-2},\quad 0.0490,\quad 490.00.$$
2. 将 22/7 作为 π 的近似值，它有几位有效数字，绝对误差限和相对误差限各为多少？
3. 要使 $\sqrt{101}$ 的相对误差不超过 $\frac{1}{2}\times10^{-4}$，至少需要保留多少位有效数字.
4. 设 x^* 为 x 的近似数，证明 $\sqrt[n]{x^*}$ 的相对误差大约为 x^* 相对误差的 $\frac{1}{n}$ 倍.
5. 某矩形的长和宽大约为 100cm 和 50cm，应该选用最小刻度为多少 cm 的测量工具，才能保证计算出的面积误差不超过 0.15cm^2.
6. 设 $x=5\pm0.1$，$y=5\pm0.1$，试估计出 $a=y/(x+1)$ 的取值范围.
7. 论证当 x^* 是 x 的较好近似时，函数值的相对误差、自变量的相对误差、相对误差意义下的条件数之间满足如下近似公式
$$\varepsilon_r(f^*)\approx \mathrm{cond}_r(f(x^*))\varepsilon_r(x^*)$$
8. 计算函数 $y=\sin(n^3x)$ 在 $x=0.0001$ 附近的函数值，当 $n=100$ 时，试估计满足函数值相对误差不超过 0.1%时的自变量相对误差限和绝对误差限.
9. 对于 32 位单精度实数系统，使用迭代格式算法
$$x_0=4,\ x_{n+1}=x_n^2,\ n=1,2,3,\cdots$$
迭代多少次将产生上溢.
10. 请设计出求解方程 $x^2+2px+q=0$ 根的一个有效算法，要求它也能够适用于 $p^2\gg|q|$ 时的情形. 用所设计算法以及求根公式计算 $p=240.05$，$q=1.00$ 时方程根的近似值(计算过程保留 2 位小数)，并给出两个根近似值的相对误差界.
11. 设有 64 位浮点系统：尾数符号占 1 位，尾数数值占 52 位，阶码符号占 1 位，阶码数值占 10 位. 请推算在此系统下实数的浮点表示能够有多少位有效数字，并计算该浮点系统的上溢界和下溢界.

第二章　函数插值

在科学与工程计算中，常会碰到函数表达式过于复杂不便于计算，而又需要计算众多点处的函数值；或者无表达式仅仅有一些采样点处的函数值，而又需要计算非采样点处的数据这类问题，此时希望建立复杂函数或者未知函数的一个便于计算的近似表达式．在数值积分、数值微分、常微分方程数值解等方面还会直接或间接地遇到函数的近似表达问题．本章研究的函数插值法则是建立函数近似表达式的一种基本方法．

§2.1　插值问题

已知定义于区间$[a,b]$上的实值函数 $f(x)$在 $n+1$ 个互异节点$\{x_i\}_{i=0}^n \subset [a,b]$处的函数值$\{f(x_i)\}_{i=0}^n$．若函数集合 Φ 中的函数 $\varphi(x)$满足

$$\varphi(x_i)=f(x_i),\qquad i=0,1,\cdots,n, \tag{2.1}$$

则称 $\varphi(x)$为 $f(x)$在函数集合 Φ 中关于节点$\{x_i\}_{i=0}^n$的一个插值函数[①]，并称 $f(x)$为被插值函数，$[a,b]$为插值区间，$\{x_i\}_{i=0}^n$为插值节点，(2.1)式为插值条件．

引入符号 $\mathcal{M}=\max\{x_i\}_{i=0}^n$，$\widetilde{m}=\min\{x_i\}_{i=0}^n$，当用插值函数 $\varphi(x)$计算被插值函数 $f(x)$在点 $x\in(\widetilde{m},\mathcal{M})$处近似值的方法称之为**内插法**，若用来计算点 $x\in[a,b]$但 $x\notin[\widetilde{m},\mathcal{M}]$处近似值的方法称之为**外插法**．

当函数集合 Φ 表示一些三角函数的多项式集合时的插值方法称之为三角插值；当函数集合 Φ 为一些有理分式的集合或者是多项式的集合时的插值方法分别称之为有理插值和代数插值．

下面针对代数插值讨论插值函数的存在性、惟一性、构造方法以及误差估计．

鉴于插值条件(2.1)式共含有 $n+1$ 个约束方程，而 n 次多项式恰有 $n+1$个待定系数，于是取函数集合 Φ 为不超过 n 次的多项式集合 $\boldsymbol{P}_n=\mathrm{span}\{1,x,x^2,\cdots,x^n\}$，即有

---

① 关于带导数的插值在后面小节叙述．

$$\boldsymbol{P}_n=\{\varphi(x)\mid \varphi(x)=a_0+a_1x+a_2x^2+\cdots+a_nx^n,\ a_i\in R,0\leqslant i\leqslant n\},$$

进而插值问题等价于确定系数$\{a_i\}_{i=0}^n$使得插值条件(2.1)成立，即

$$\begin{bmatrix}1 & x_0 & x_0^2 & \cdots & x_0^n\\ 1 & x_1 & x_1^2 & \cdots & x_1^n\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & x_n & x_n^2 & \cdots & x_n^n\end{bmatrix}\begin{bmatrix}a_0\\ a_1\\ a_2\\ \vdots\\ a_n\end{bmatrix}=\begin{bmatrix}f(x_0)\\ f(x_1)\\ f(x_2)\\ \vdots\\ f(x_n)\end{bmatrix},\tag{2.2}$$

线性方程组(2.2)的系数矩阵是范德蒙德(Vandermonde)矩阵，又节点$\{x_i\}_{i=0}^n$互异，故系数矩阵非奇异，线性方程组(2.2)存在惟一解. 这样得到如下定理

定理 2.1（存在惟一性）　满足插值条件(2.1)的不超过 n 次的插值多项式是存在惟一的.

该定理的几何解释是，平面上有且仅有一条不超过 n 次的代数曲线恰好通过给定的 $n+1$ 点$\{(x_i,f(x_i))\}_{i=0}^n$. 上面的分析过程也指出通过求解线性方程组(2.2)可以求得该代数曲线的 $n+1$ 个系数值.

称被插值函数 $f(x)$与插值函数 $\varphi(x)$之间的误差 $R_n(x)=f(x)-\varphi(x)$为插值余项. 它满足如下定理.

定理 2.2（误差估计）　设 $f^{(n)}(x)$在区间$[a,b]$上连续，$f^{(n+1)}(x)$在区间(a,b)内存在. $\varphi(x)$是满足插值条件(2.1)的不超过 n 次的插值多项式. 则对任意 $x\in[a,b]$，存在 $\zeta=\zeta(x)\in(a,b)$，使得

$$R_n(x)=f(x)-\varphi(x)=\frac{f^{(n+1)}(\zeta)}{(n+1)!}\omega_{n+1}(x)\tag{2.3}$$

成立，式中 $\omega_{n+1}(x)=\prod_{i=0}^{n}(x-x_i)$. 进而当$|f^{(n+1)}(x)|$在区间$(a,b)$有上界 M_{n+1}时，有

$$|R_n(x)|\leqslant\frac{M_{n+1}}{(n+1)!}|\omega_{n+1}(x)|.\tag{2.4}$$

证明　由插值条件(2.1)知

$$R_n(x_i)=f(x_i)-\varphi(x_i)=0,\qquad i=0,1,2,\cdots,n,$$

因此可设插值余项

$$R_n(x)=k(x)\omega_{n+1}(x). \tag{2.5}$$

当 x 为某一插值节点时，对任意的 $\zeta\in(a,b)$结论(2.3)成立. 当点 x 与插值节点 x_0、x_1、…、x_n 互不相同时，它们均是以 t 为自变量的辅助函数

$$g(t)=f(t)-\varphi(t)-k(x)\omega_{n+1}(t)$$

在区间$[a,b]$上的 $n+2$ 个互异零点. 由函数 $f(x)$和多项式函数的光滑性知，$g^{(n)}(t)$在区间$[a,b]$上连续，$g^{(n+1)}(t)$在区间(a,b)内存在.

对函数 $g(t)$在区间$[a,b]$上的 $n+2$ 个互异零点形成的 $n+1$ 个子区间上使用罗尔(Rolle)定理，函数 $g'(t)$在区间(a,b)上至少有 $n+1$ 个互异零点. 这 $n+1$ 零点又形成 n 个子区间，对 $g'(t)$在这些子区间上使用罗尔定理，函数 $g''(t)$在区间(a,b)上至少有 n 个互异零点. 以此类推，函数 $g^{(n+1)}(t)$在区间(a,b)上至少有 1 个零点 $\zeta=\zeta(x;x_0,x_1,\cdots,x_n)$.

由函数 $g(t)$的形式知 $g^{(n+1)}(t)=f^{(n+1)}(t)-(n+1)!\,k(x)$，将 $g^{(n+1)}(t)$的零点 ζ带入得到

$$k(x)=\frac{f^{(n+1)}(\zeta)}{(n+1)!},$$

将上式带入(2.5)知结论(2.3)成立. 结论(2.4)可由(2.3)直接得到. #

从此定理结论可以看到，插值误差与节点$\{x_i\}_{i=0}^n$和点 x 之间的距离有关，节点距离 x 越近，一般地插值误差越小. 特别地，当被插值函数 $f(x)$自身就是不超过 n 次的多项式，则有 $f(x)\equiv\varphi(x)$.

§2.2 插值多项式的构造方法

建立插值多项式的方法简称为插值法. 上节提到的插值法需要求解线性方程组，这里介绍更为简便实用的方法：拉格朗日(Lagrange)插值法和牛顿(Newton)插值法，并将用之构造的满足插值条件(2.1)的插值多项式分别记为 $L_n(x)$和 $N_n(x)$. 由插值多项式的存在惟一性定理 2.1 知，这两种方法构造出的插值多项式是恒等的，即有 $L_n(x)\equiv N_n(x)$.

2.2.1 拉格朗日插值法

针对 $n+1$ 个互异的插值节点$\{x_i\}_{i=0}^n$，我们引入如下辅助问题：

构造不超过 n 次的插值多项式 $l_i(x)$，使之满足插值条件

$$l_i(x_j)=\delta_{ij}=\begin{cases}1, & j=i,\\ 0, & j\neq i,\end{cases}\quad j=0,1,2,\cdots,n. \tag{2.6}$$

插值条件(2.6)要求不超过 n 次的插值多项式 $l_i(x)$在除节点 x_i 外的其余节点处的函数值为零，故 $l_i(x)$必然可表示为如下形式

$$l_i(x)=c(x-x_0)(x-x_1)\cdots(x-x_{i-1})(x-x_{i+1})\cdots(x-x_n),$$

由插值条件 $l_i(x_i)=1$ 可求得常数

$$c=\frac{1}{(x_i-x_0)(x_i-x_1)\cdots(x_i-x_{i-1})(x_i-x_{i+1})\cdots(x_i-x_n)},$$

这样得到插值函数

$$\begin{aligned}l_i(x)&=\frac{(x-x_0)(x-x_1)\cdots(x-x_{i-1})(x-x_{i+1})\cdots(x-x_n)}{(x_i-x_0)(x_i-x_1)\cdots(x_i-x_{i-1})(x_i-x_{i+1})\cdots(x_i-x_n)}\\&=\frac{\omega_{n+1}(x)}{(x-x_i)\omega'_{n+1}(x_i)}.\end{aligned}$$

当 $i=0,1,2,\cdots,n$ 时，便得到了 $n+1$ 个 n 次插值多项式 $l_0(x)$、$l_1(x)$、$\cdots$、$l_n(x)$，称它们为关于节点$\{x_i\}_{i=0}^n$的**拉格朗日插值基函数**. 这些基函数仅依赖于插值节点$\{x_i\}_{i=0}^n$，并满足

$$l_i(x_j)=\delta_{ij},\quad i,j=0,1,2,\cdots,n. \tag{2.7}$$

利用式(2.7)可以验证不超过 n 次的多项式

$$L_n(x)=\sum_{i=0}^{n}f(x_i)l_i(x)$$

满足式(2.1)列出的所有插值条件. 结合插值多项式的存在惟一性定理 2.1 知 $L_n(x)$正是所需要建立的插值多项式，称之为**拉格朗日插值多项式**. 该插值多项式具有结构清晰紧凑的特点，常用于理论分析. 当被插值函数 $f(x)$ 满足插值误差估计定理 2.2 的条件时有误差估计

$$R_n(x)=f(x)-L_n(x)=\frac{f^{(n+1)}(\zeta)}{(n+1)!}\omega_{n+1}(x) \tag{2.8}$$

和

$$|R_n(x)|=|f(x)-L_n(x)|\leqslant\frac{M_{n+1}}{(n+1)!}\left|\omega_{n+1}(x)\right|. \tag{2.9}$$