“十二五”重点图书

西安电子科技大学研究生教材建设资助项目

研究生系列教材

数值分析

Numerical Analysis

冯象初　任春丽　尚晓清　王卫卫　编著

西安电子科技大学出版社
http://www.xduph.com

“十二五”重点图书　　★研究生系列教材

西安电子科技大学研究生教材建设资助项目

数值分析

冯象初　任春丽
尚晓清　王卫卫　编著

西安电子科技大学出版社

内 容 简 介

本书系统地介绍了数值分析的理论和算法。全书共7章，内容包括三部分：第一部分是泛函分析基础，主要介绍距离空间、Banach空间、Hilbert空间的基本概念和理论；第二部分是数值逼近，包括函数的插值、逼近问题，数据处理问题，数值积分和数值微分；第三部分是数值代数，包括线性方程组、非线性方程(组)的数值解法，矩阵的特征问题。

本书内容丰富，论述翔实严谨，可作为数学系高年级本科生及电子、通信、计算机等理、工科专业研究生的教材，也可供从事科学和工程计算的科技工作者参考。

图书在版编目(CIP)数据

数值分析/冯象初等编著. —西安：西安电子科技大学出版社，2013.2
研究生系列教材
ISBN 978-7-5606-3006-9

Ⅰ. ①数… Ⅱ. ①冯… Ⅲ. ①数值分析—研究生—教材 Ⅳ. ①O241

中国版本图书馆CIP数据核字(2012)第089974号

策　　划　李惠萍
责任编辑　李惠萍　段　雷
出版发行　西安电子科技大学出版社(西安市太白南路2号)
电　　话　(029)88242885　88201467　　邮　　编　710071
网　　址　www.xduph.com　　电子邮箱　xdupfxb001@163.com
经　　销　新华书店
印刷单位　陕西天意印务有限责任公司
版　　次　2013年2月第1版　2013年2月第1次印刷
开　　本　787毫米×960毫米　1/16　　印　　张　14.5
字　　数　294千字
印　　数　1～3000册
定　　价　25.00元
ISBN 978-7-5606-3006-9/O

XDUP 3298001-1

“十二五”重点图书　　研究生系列教材

编审委员会名单

前　言

“数值分析”课程是我校数学科学系应用数学专业、信息与计算科学专业本科生的专业基础课，也是我校电子类各专业硕士研究生的学位课，每年有千余人学习此课程。学习此课程的硕士研究生来自通信、电子工程、电子机械、微电子、技术物理、计算机科学、计算机应用等多个学科和专业。

本教材编写组成员长期从事本科生和研究生的数值分析课程教学工作以及与数值分析密切相关的科研工作，在教学和科研过程中不断探索，特别是结合选课学生的专业特点和专业需要，在教学内容的设计方面进行了深入的改革。本书是在我校宋国乡教授主编的《数值分析》教材的基础上重新编写而成的。和传统的数值分析教材相比，本书增加了泛函分析基础知识、二元插值、二维梯度和方向导数的计算(图像方向场)、矩阵的广义逆、奇异值分解等内容；对大部分数值方法都给出了实例，并配套提供了适当难度和数量的习题。

本教材的主要特点是：

(1) 重视理论分析，有利于进一步巩固学生的理论基础，加强培养学生的理论分析能力；

(2) 重视理论与应用相结合，有利于加强培养学生灵活运用理论解决实际问题的能力；

(3) 教材内容与时俱进，结合专业需要，有很强的实用性。

(4) 设计有工程背景的数值试验，有利于加强培养学生的数值仿真能力。

全书共分 7 章，其中第 0 章和后续的部分章节(1.5、2.5、3.7、4.5、5.6、6.6 等节)由冯象初教授编写，第 1、2 章由任春丽副教授编写，第 3、4 章由尚晓清副教授编写，第 5、6 章由王卫卫教授编写。本书的编著者署名也按此顺序排序。

本书在编写过程中得到了西安电子科技大学理学院、研究生院和教务处的大力支持，在此深表谢意，同时，对于本书编写时所参考的相关文献的作者也一并表示感谢。

由于编者的水平有限，错漏和不妥之处在所难免，欢迎广大读者批评指正。

编 著 者

2012 年夏于西安电子科技大学

目　录

第0章 引 言

0.1 绪 论

0.1.1 数值分析

“数值分析”(Numerical Analysis)研究的是在计算机上解决数学问题的理论和数值方法，它包括数值算法的构造(计算公式和算法步骤)、算法的理论分析(误差分析、收敛性、稳定性)等。数值分析把数学理论与计算机应用紧密结合起来，既有纯数学的高度抽象性与严密科学性，又有应用的广泛性与实际实验的高度技术性。

用计算机解决科学计算问题时往往要经历以下几个过程：实际问题→建立数学模型→提出数值计算方法→程序设计→编程上机计算→分析结果并对实际问题进行解释说明。从中可以看出数值计算方法起着承上启下的作用，是连接数学模型到计算结果的重要环节。数值计算方法的根本任务是：针对具体的数学模型，研究通过计算机所能执行的基本运算(加、减、乘、除)来求得各类问题的数值解的方法，通过程序设计、运算获得计算结果，对算法和结果进行相应的理论分析，如收敛性、误差分析等，从而保证计算结果满足实际要求。

数值计算方法是在离散化的基础上进行的，其解决问题的最终结果不是解析解而是数值解。连续性问题(如：微分方程、积分方程)的求解，首先要通过特定的手段将连续问题离散化(如用差分代替微分)，然后转化为代数问题。对于给定的数学模型，采用不同的离散手段可以导致不同的数值方法，应该通过理论分析、数值试验等加以研究。例如，用计算机运算得到的结果是否收敛到实际问题的解以及收敛速度的快慢等。

例 0.1.1 求非线性方程 $f(x)=0$ 的根。

解 若已知方程的粗略近似解为 x_0，由 Taylor(泰勒)级数得

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0)$$

令其线性部分为 0，则

$$f(x_0)+f'(x_0)(x-x_0)=0$$

其解为

$$x_1=x_0-\frac{f(x_0)}{f'(x_0)}$$

得迭代公式

$$x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)}$$

这个算法称为牛顿迭代法。我们要研究在什么条件下迭代是收敛的，是否收敛到方程的根，以及收敛的速度等问题。

由于计算机的内存大小、运算速度等约束，对选定的算法要进行计算量分析。算法的计算量主要由以下两个因素决定：一是使用中央处理器(CPU)的时间，这主要由四则运算的次数决定；二是占用内存储器的空间，这主要由使用的数据量来决定。这两个因素称为时间与空间的复杂度，简称计算复杂度。

例 0.1.2 解线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$。

解 若 $\det A\neq 0$，则可用 Cramer 法则来解。设 $\boldsymbol{A}$ 为 20 阶矩阵，需计算 21 个 20 阶的行列式。计算一个 20 阶行列式需要的乘法运算量为 $19\times 20!$ 次，总的乘法运算量为

$$21\times 19\times 20!\approx 9.71\times 10^{20}\text{次}$$

若用 10 亿次/ 秒的计算机来运算，则一年可完成的乘法运算量为

$$10^9\times 365\times 24\times 3600\approx 3.15\times 10^{16}\text{次 / 年}$$

解 20 阶的方程组所需乘法运算的时间为

$$9.71\times 10^{20}\div(3.15\times 10^{16})\approx 3.08\times 10^4\text{年}$$

显然这个运算时间在实际中是不可接受的。而在实际问题中，例如大型水利工程、天气预报等，需要解的大型方程组的阶数一般都远远大于 20。这个例子说明解线性方程组的 Cramer法则在理论上虽然可行，但在实际应用中却不可行。若用第 4 章介绍的高斯消去法，则只需要 $O(n^3)$次运算。

由于计算机字长的限制，计算机只能近似地表示实数。不论计算机中的数是定点表示，还是浮点表示，它所表示的数的位数都是有限的，这说明用计算机运算得到的结果都是近似的，因此需要对算法进行误差分析，进一步要考虑算法的数值稳定性问题。

例 0.1.3 当 $b^2-4ac>0$ 时，方程 $ax^2+bx+c=0$ 有两个相异的实根 $x_{1,2}=\frac{-b\pm\sqrt{b^2-4ac}}{2a}$。若直接按此编写程序，则当 $b^2\gg 4ac$ 时，$\sqrt{b^2-4ac}\approx|b|$，所以分子中会出现两个相近数相减，这会使有效数字损失，影响计算精度，在 0.2 节中我们会给出有效数字的具体分析。

上面的几个例子告诉我们，算法设计、计算效率分析和误差分析是数值分析的核心问

题。总之，对于给定的数学模型所提出的可行、有效的算法应该是符合计算机的要求，在理论上收敛、稳定，在实际计算中精确度高，计算复杂性小，能通过试验验证的数值方法。

0.1.2 泛函分析

泛函分析(Functional Analysis)把客观世界中的研究对象抽象为元素，在集合的基础上，定义距离、范数、内积等运算。于是泛函分析将表面上彼此不相关的学科统一在它的普遍规律和共同框架之下。整个泛函分析都是用“空间”、“元素”来表示的，这使很多经典的理论具有简单明了的几何直观，因而泛函分析具有高度的抽象性、系统性和普遍性。它的观点、方法和规律可以广泛地应用于各学科。

电子计算机的出现和泛函分析在数值分析领域中的应用，使数值分析发生了革命性的变化。计算机是数值分析的计算工具，而泛函分析是进行数值方法研究的理论工具。对数值分析而言，运用泛函分析的观点与语言可使数值分析中很多定理与方法的推导变得简洁、直观，并使得结论具有普遍性。本课程将介绍泛函分析中与数值分析有关系的基本概念和理论。

0.1.3 本课程的内容及要求

课程内容：在集合的基础上给出距离空间、赋范线性空间、希尔伯特(Hilbert)空间的定义与性质；讨论距离、范数、内积及相关问题，如投影、不动点定理等；介绍工程和科学实验中最基本、最常用的数值算法——插值法、最佳逼近、数值积分与数值微分、线性与非线性方程组的数值求解、矩阵的特征值与特征向量的计算等。

课程要求：熟悉泛函分析的基本概念和理论；掌握数值分析中相关的理论分析技巧和数值求解方法；熟悉所学方法的计算过程，并在实践中能够合理选择和使用数值计算方法；培养科学计算的能力。

0.1.4 算法的实现

掌握数值计算方法的根本目的是解决所遇到的各种数学问题，为此必须在计算机上实现算法，这包含使用软件工具和自编程序两方面的含义。

本课程涉及的大多数是数值分析的基本问题，有很多现成的通用或专用数学软件包含有实现这些算法的子程序，如 Mathematica、Matlab 等，可以直接调用这些子程序或库函数求出数值解。尽管如此，自编程序仍是不可缺少的。一则，子程序和库函数是孤立的功能块，实际问题往往需要综合使用多种数学方法才能解决；二则，一个数学问题可能有多种解法，各有优缺点，需要选择合适的方法，这就需要对算法的性质有深刻的了解，只有通过理论和实践两方面才能得到解决；三则，数学软件和函数库并非包罗万象，有的问题

需要自己构造算法并编制软件。

0.2 误差的来源、基本概念及分析方法与原则

0.2.1 误差的来源

误差的来源主要有以下几方面：

(1) 模型误差。一般来说，生产和科研中遇到的实际问题是比较复杂的，要用数学模型来描述，需要进行必要的简化，忽略一些次要的因素，这样建立起来的数学模型与实际问题之间一定有误差。数学模型与实际问题之间的误差称为模型误差。

(2) 观测误差。一般数学问题包含若干参数，它们的值往往通过观测得到。实验或观测得到的数据与实际数据之间的误差称为观测误差或数据误差。

(3) 方法误差(截断误差)。一般数学问题难以求解，往往要通过近似替代，简化为较易求解的问题。数学模型的精确解与数值方法得到的数值解之间的误差称为方法误差或截断误差。例如，由 Taylor 公式得

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

用 $p_n(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}$ 近似代替 e^x，这时的截断误差为

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

ξ 介于 0 与 x 之间。

(4) 舍入误差。对数据进行四舍五入后产生的误差称为舍入误差。

由于数值分析研究的是数学问题的数值解法，所以本课程中我们只讨论方法误差(截断误差)和舍入误差。

0.2.2 误差的基本概念

1. 绝对误差和绝对误差限、相对误差和相对误差限

定义 0.2.1 设 x^* 为准确值，x 是 x^* 的近似值，称

$$e = x^* - x \tag{0.2.1}$$

为近似值 x 的绝对误差，简称误差。

误差 e 既可为正，也可为负。一般来说，准确值 x^* 是不知道的，因此误差 e 的准确值无法求出。在实际工作中，可根据相关领域的知识、经验及测量工具的精度，事先估计出

误差绝对值不超过某个正数 ε，即

$$|e| = |x^* - x| \leqslant \varepsilon \tag{0.2.2}$$

则称 ε 为近似值 x 的绝对误差限，简称为误差限或精度。

由式(0.2.2)得

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

有时将准确值 x^* 写成 $x^* = x \pm \varepsilon$。例如用卡尺测量一个圆杆的直径，其近似值为 $x = 350$ mm，由卡尺的精度知道这个近似值的误差不会超过 0.5 mm，则有

$$|x^* - x| = |x^* - 350| \leqslant 0.5 \text{ mm}$$

于是该圆杆的直径为

$$x^* = 350 \pm 0.5 \text{ mm}$$

用 $x^* = x \pm \varepsilon$ 表示准确值可以反映它的准确程度，但不能说明近似值的优劣。例如，测量一根 10 cm 长的圆钢时发生了 0.5 cm 的误差，和测量一根 10 m 长的圆钢时发生了 0.5 cm 的误差，其绝对误差都是 0.5 cm，但是，后者的测量结果显然比前者要准确得多。这说明决定一个量的近似值的优劣，除了要考虑绝对误差的大小，还要考虑准确值本身的大小，这就需要引入相对误差的概念。

定义 0.2.2　设 x^* 为准确值，x 是 x^* 的近似值，称

$$e_r = \frac{e}{x^*} = \frac{x^* - x}{x^*} \tag{0.2.3}$$

为近似值 x 的相对误差。

在实际计算中，由于准确值 x^* 总是未知的，因此也把

$$e_r = \frac{e}{x} = \frac{x^* - x}{x} \tag{0.2.4}$$

称为近似值 x 的相对误差。

在上面测量圆钢长度的例子中，前者的相对误差是 $0.5/10 = 0.05$，而后者的相对误差是 $0.5/1000 = 0.0005$。一般来说，相对误差越小，表明近似程度越好。与绝对误差一样，相对误差的准确值也无法求出。仿绝对误差限，称相对误差绝对值的上界 ε_r 为相对误差限，即

$$|e_r| = \left|\frac{x^* - x}{x}\right| \leqslant \varepsilon_r \tag{0.2.5}$$

注　绝对误差和绝对误差限有量纲(或单位)，而相对误差和相对误差限没有量纲(或单位)，通常用百分数来表示它们。

2. 有效数字及其与相对误差限的关系

当在实际运算中遇到的数位数很多时，如 π，e 等，常常采用四舍五入的原则得到近似值，为此引进有效数字的概念。

定义 0.2.3 设近似数 $x=\pm 0.a_1a_2\cdots a_n\times 10^m$，其中 a_1，a_2，…，a_n是 0 到 9 之间的自然数，$a_1\neq 0$，m 为整数。如果

$$|x^*-x|\leqslant \frac{1}{2}\times 10^{m-n} \tag{0.2.6}$$

则称近似值 x 具有 n 位有效数字。其中 a_1，a_2，…，a_n都是 x 的有效数字，也称 x 是有 n 位有效数字的近似值。

例 0.2.1 设 $x^*=3.200\ 169$，确定它的近似值 $x_1=3.2001$ 和 $x_2=3.2002$ 分别具有几位有效数字？

解 因为 $x_1=0.320\ 01\times 10^1$，$m=1$，$|x*-x_1|=0.069\times 10^{-3}<0.5\times 10^{-3}$，所以 $m-n=-3$，得 $n=4$，故 $x_1=3.2001$ 具有 4 位有效数字，而最后一位数字 1 不是有效数字。

因为 $x_2=0.320\ 02\times 10^1$，$m=1$，$|x-x_2^*|=0.31\times 10^{-4}<0.5\times 10^{-4}$，所以 $m-n=-4$，得 $n=5$，故 $x_2=3.2002$ 具有 5 位有效数字。

特别要指出的是，$x^*=3.200$ 有 4 位有效数字，而 $x=3.2$ 只有两位有效数字。

从上面的讨论可以看出，有效数字位数越多，绝对误差限就越小。

下面阐述有效数字与相对误差限的联系。

定理 0.2.1 设近似值 $x=\pm 0.a_1a_2\cdots a_n\times 10^m$ 有 n 位有效数字，则其相对误差限为 $\varepsilon_r=\dfrac{1}{2a_1}\times 10^{-n+1}$。

证明 因为 x 具有 n 位有效数字，所以由定义 0.2.3 知

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

又 $|x|\geqslant a_1\times 10^{m-1}$，所以

$$\frac{|x^*-x|}{|x|}\leqslant \frac{\frac{1}{2}\times 10^{m-n}}{a_1\times 10^{m-1}}=\frac{1}{2a_1}\times 10^{-n+1}$$

证毕。

定理 0.2.2 设近似值 $x=\pm 0.a_1a_2\cdots a_n\times 10^m$，相对误差限为 $\varepsilon_r=\dfrac{1}{2(a_1+1)}\times 10^{-n+1}$，则它至少具有 n 位有效数字。

证明 因为 $|x|\leqslant (a_1+1)\times 10^{m-1}$，所以

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

故由定义 0.2.3 知 x 至少具有 n 位有效数字。证毕。

上面的定理表明：有效数字位数越多，相对误差限也就越小。

例 0.2.2 设$\sqrt{5}$的近似值 x 的相对误差不超过 0.1%，问 x 至少具有几位有效数字？

解　设 x 至少具有 n 位有效数字，因为$\sqrt{5}=2.23\cdots$的第一个非零数字是 2，即 x 的第一位有效数字 $a_1=2$，根据题意及定理知，

$$\frac{|\sqrt{5}-x^*|}{|x|}\leqslant\frac{1}{2a_1}\times10^{-n+1}=\frac{1}{2\times2}\times10^{-n+1}\leqslant10^{-3}$$

得 $n\geqslant3.398$，故取 $n=4$，即 x 至少具有 4 位有效数字，其相对误差不超过 0.1%。

0.2.3　减少误差的若干原则

在用计算机实现算法时，我们输入计算机的数据一般是有误差的(如观测误差等)，计算机运算过程的每一步又会产生舍入误差，由十进制转化为机器数也会产生舍入误差，这些误差在迭代过程中还会逐步传播和积累，因此我们必须研究这些误差对计算结果的影响。但一个实际问题往往需要亿万次以上的计算，且每一步都可能产生误差，因此我们不可能对每一步误差进行分析和研究，只能根据具体问题的特点进行研究，提出相应的误差估计。特别地，如果我们在构造算法的过程中注意了以下几项原则，那么将有效地减少和避免误差的危害，控制误差的传播和积累。

(1) 避免两个相近的数相减。

在数值计算中两个相近的数相减会造成有效数字的严重损失，从而导致误差增大，影响计算结果的精度。

例 0.2.3　当 $x=10003$ 时，计算 $\sqrt{x+1}-\sqrt{x}$的近似值。

解　若使用 6 位十进制浮点运算，运算时取 6 位有效数字，结果

$$\sqrt{x+1}-\sqrt{x}=100.020-100.015=0.005$$

只有一位有效数字，损失了 5 位有效数字，使得绝对误差和相对误差都变得很大，影响计算结果的精度。若改用

$$\sqrt{x+1}-\sqrt{x}=\frac{1}{\sqrt{x+1}+\sqrt{x}}=\frac{1}{100.020+100.015}=0.00499913$$

则其结果有 6 位有效数字，与精确值 0.00499912523117984…较为接近。

(2) 防止重要的小数被大数“吃掉”。

在数值计算中，参加运算的数的数量级有时相差很大，而计算机的字长又是有限的，因此，如果不注意运算次序，那么就可能出现小数被大数“吃掉”的现象。这种现象在有些情况下是允许的，但在有些情况下，这些小数很重要，若它们被“吃掉”，就会造成计算结果的失真，影响计算结果的可靠性。

例 0.2.4　求二次方程 $x^2-(10^9+1)x+10^9=0$ 的根。

解　用因式分解易得方程的两个根为 $x_1=10^9$，$x_2=1$，但用求根公式 $x_{1,2}=\frac{-b\pm\sqrt{b^2-4ac}}{2a}$编制程序，如果在只能将数表示到小数后 8 位的计算机上运算，那么首先

要对$-b$的阶级即小数位数进行分析：

$$-b = 10^9 + 1 = 0.1000000 \times 10^{10} + 0.0000000001 \times 10^{10}$$

而计算机上只能达到8位，故在计算机上$0.0000000001\times10^{10}$不起作用，即视为0，于是

$$-b = 0.1000000 \times 10^{10} = 10^9$$

类似地，有$\sqrt{b^2-4ac}=|b|=10^9$，故所得两个根为$x_1=10^9$，$x_2=0$，x_2严重失真的原因是大数吃掉小数的结果。

如果把x_2的计算公式写成$x_2=\dfrac{-b-\sqrt{b^2-4ac}}{2a}=\dfrac{2c}{-b+\sqrt{b^2-4ac}}$，则

$$x_2 = \frac{2\times10^9}{10^9+10^9} = 1$$

再如，已知$x=3\times10^{12}$，$y=7$，$z=-3\times10^{12}$，求$x+y+z$。如果按$x+y+z$的次序来编程序，x“吃掉”y，而x与z互相抵消，其结果为零。若按$(x+z)+y$的次序来编程序，其结果为7。由此可见，如果事先大致估计一下计算方案中各数的数量级，编制程序时加以合理的安排，那么重要的小数就可以避免被“吃掉”。此例还说明，用计算机做加减运算时，交换律和结合律往往不成立，不同的运算次序会得到不同的运算结果。

(3) 避免出现除数的绝对值远远小于被除数绝对值的情况。

在用计算机实现算法的过程中，如果用绝对值很小的数作除数，往往会使舍入误差增大。即在计算y/x时，若$|x|\ll1$，则可能产生较大的舍入误差，对计算结果带来严重影响，应尽量避免。

例 0.2.5 在4位浮点十进制数下，用消去法解线性方程组

$$\begin{cases} 0.00003x_1 - 3x_2 = 0.6 \\ x_1 + 2x_2 = 1 \end{cases}$$

解 仿计算机实际计算，将上述方程组写成

$$\begin{cases} 0.3000\times10^{-4}x_1 - 0.3000\times10^{1}x_2 = 0.6000\times10^{0} & (1) \\ 0.1000\times10^{1}x_1 + 0.2000\times10^{1}x_2 = 0.1000\times10^{1} & (2) \end{cases}$$

(1)÷(0.3000×10^{-4})−(2)(注意：在第一步运算中出现了用很小的数作除数的情形，相应地在第二步运算中出现了大数“吃掉”小数的情形)，得

$$\begin{cases} 0.3000\times10^{-4}x_1 - 0.3000\times10^{1}x_2 = 0.6000\times10^{0} \\ -0.1000\times10^{6}x_2 = 0.2000\times10^{5} \end{cases}$$

解得

$$x_1=0, \quad x_2=-0.2$$

而原方程组的准确解为$x_1=1.399\ 972\cdots$，$x_2=-0.199\ 986\cdots$。显然上述结果严重失真。

如果反过来用第二个方程消去第一个方程中含x_1的项，那么就可以避免很小的数作除数的情形。即(2)×(0.3000×10^{-4})−(1)，得