高等学校研究生教材

应用数值分析

YINGYONG SHUZHI FENXI

（第三版）

文世鹏　张　明　编著

石油工业出版社
PETROLEUM INDUSTRY PRESS

高等学校研究生教材

应用数值分析

（第三版）

文世鹏　张　明　编著

石油工业出版社

内 容 提 要

本书系统论述了求解数学问题的数值算法的理论原理和方法。全书共分十一章，包括数值代数和数值逼近的主要内容、非线性问题的数值方法、常微分方程数值解法和 MATLAB 应用基础五个部分。书中的算法还包括有近代较新的方法，并与科学计算软件 MATLAB 的实现相结合。

本书作为工科研究生数值分析课的教学用书，但书中内容与本科生用的计算方法课程内容也相衔接，并力求理论与实际应用并重，循序渐进且自成体系。本书既可作为高等学校数值分析课的教材，也可供科技工程人员参考和自学。

图书在版编目(CIP)数据

应用数值分析/文世鹏，张明编著.
北京：石油工业出版社，2005.7
高等学校研究生教材
ISBN 7-5021-5124-9

Ⅰ. 应…
Ⅱ. ①文… ②张…
Ⅲ. 数值计算-研究生-教材
Ⅳ. 0241

中国版本图书馆 CIP 数据核字(2005)第 071805 号

出版发行：石油工业出版社
(北京安定门外安华里 2 区 1 号 100011)
网 址：www.petropub.com.cn
总 机：(010)64262233 发行部：(010)64210392
经 销：全国新华书店
印 刷：石油工业出版社印刷厂

2005 年 7 月第 3 版 2005 年 7 月第 3 次印刷
787×1092 毫米 开本：1/16 印张：26.5
字数：674 千字 印数：3501—6500 册

定价：35.00 元
(如出现印装质量问题，我社发行部负责调换)

第三版前言

为了满足教学的需要,《应用数值分析》在第二版的基础上又出了第三版。

本书的第三版较前两版有较多的更新变动。全书内容的章节安排有了变化,有些章节重新撰写,有些章节有删减也有增补。在函数插值、最小二乘拟合以及数值积分部分都增加了二维问题的数值方法。涉及应用数学软件 MATLAB 的第十一章“MATLAB 与数值计算”和附录“MATLAB 入门基础”这两部分,根据 MATLAB 更新版本的变化也相应地进行了调整修订。全书仍然保持了自成体系以及对不同层次要求的读者提供较多可选择的自学阅读空间的特点。

在教学过程中,使用本书作为工科研究生数值分析课教材的任课教师和学生对本书的再版提出了很多好的建议,抱有很好的期望,在此,对他们表示衷心的感谢。在本书第三版的编写中,支丽霞对全书习题作了修正补充;明辉对与 MATLAB 应用有关的第十一章的附录进行审核和删补;苏宇和熊天智负责书稿的编辑,部分章节打印、抄写和绘制图形。由于他们有成效的工作才使本书能顺利及时地出版,在此表示诚挚的谢意。同时也要感谢石油工业出版社的大力支持。

第三版书稿完成之后,虽感尽了力,但仍觉有许多不尽人意的地方。随着计算机科学的迅速发展,与计算机应用紧密相关的数值计算类课程的教学内容和教学方法也要随之不断改进,这也是编写本书第三版的缘由之一。

限于编者水平,书中一定会有疏漏不妥之处,诚恳地希望使用本书的读者批评指正。

文世鹏　张　明

2005 年 3 月于中国石油大学

第二版前言

本书自出版以来，作为研究生数值分析课教材在教学中使用过多次，受到读者欢迎，并获得2000年度中国石油天然气集团公司颁发的优秀教学成果奖。根据教学需要和读者要求，决定对原书进行必要的修订补充后重新出版。

这次再版除了对原书的内容和例题有一些订正外，最大的变动是新增加了第十二章“MATLAB与数值计算”以及“MATLAB应用基础”。MATLAB作为近代优秀的功能强大而又使用方便的科学计算应用软件，已在国内外得到了广泛的应用。这两年，我们在教学中已经把数值分析课与MATLAB相结合，用MATLAB作为学习数值分析课的基本计算机语言和计算工具，取得了很好的效果。这次再版增补的内容也是我们教学实践积累的结果。

在此，要感谢我们的学生对本书再版的关心和支持鼓励，提出了很多好的建议；感谢负责再版的工作人员的努力，使得新书能顺利及时出版。

文世鹏　张　明

2001.7.30

第一版前言

随着计算机科学的发展,数值计算的方法已成为自然科学、工程技术领域中解决实际问题最有效的手段。科学计算能力是现代高层次人材综合素质中所不可缺少的。而高等学校中以现代数值方法的理论原理、计算及计算机实现为主要内容的数值分析课,在培养研究生的科学计算能力方面有着重要的、不可替代的作用。因此,大多数工程技术和应用科学方面的硕士研究生都选择数值分析课作为必修的学位课。这本《应用数值分析》就是作者在为工科硕士研究生多年讲授数值分析课的基础上形成的,并根据积累的教学经验和信息,对内容进行了认真的选择增补和重新安排。

研究生数值分析课是在大学本科计算方法课基础上的深入和提高,主要包括计算数学中数值代数和数值逼近及相关内容。在计算数学中这两部分的内容很多,而作为一本教材如何选择安排内容呢? 根据多年的教学实践及国内外同类学校相关信息,我们有以下几点认识:

1. 数值分析课不应该只是罗列现有的基本算法及重复已有的程序编码流程。应该把主要精力放在提高学生分析算法,能较快接受和应用现代高效算法及构造设计专用算法的能力上。这就要求学生具有计算数学中必要的理论基础,如矩阵论基础、泛函分析的基本思想方法以及用抽象数学空间、向量和矩阵认识和处理问题的能力。

2. 工科硕士研究生学习数值分析课的主要目的,不是为了将来从事计算数学中数值代数和数值逼近本身的研究工作。因此,在数值分析课中虽然应重视加强理论基础,但不能、也没有必要过于理论化和专门化,其目的应该是让学生较好地了解一个好的数值方法是在什么理论原理的基础上产生的,以及怎样认识、比较和选择不同算法。

3. 由于计算机科学的迅速发展,不断出现一些新的数值方法。因此,在数值分析课中不能只停留在学习成熟的传统的基本数值方法上,还需要接触和了解一些近代出现的新的数值方法。

4. 要处理好与大学本科计算方法内容的衔接与提高深入。

以上几点虽然认识到了,但不一定能在实际中做得很好,我们应尽力沿着这个方向努力。全书共分十一章,内容大致分为四个部分。

第一部分:包括第一章,第二章。第一章是对大学本科线性代数重要内容的总结复习和补充深化,这是学习本书必备的基础,熟悉这部分的读者可以跳过本章。第二章是矩阵论中矩阵分解的内容,包括重要的三角分解、正交分解和奇异值分解。本书中这部分内容与矩阵论课中不同的是,不仅讨论矩阵分解的理论原理,还讨论实现这些矩阵分解,并适合在计算机上计算的具体算法。安排第一部分这两章内容,使得本书可以基本上自成体系。

第二部分:包括第三,四,五,六,十章。这些章包括了数值代数中的三个主要内容:线性代数方程组的解法,最小二乘问题和代数特征值问题;并专门安排了第六章非线性方程组的数值方法。非线性问题是当前各个领域中非常重要、非常活跃的问题,在第六章中重点讨论了基本而有效的牛顿型算法和无约束优化算法。

第三部分:包括第七,八,九章。这一部分是数值逼近的主要内容,包括函数插值、函数最佳逼近和数值积分。在函数最佳逼近中,重点介绍了最佳平方逼近。另外,还介绍了二维区域

的函数插值和曲面拟合问题。

第四部分是第十一章,常微分方程初边值问题的解法。除了初值问题外,对常微分方程边值问题也作了较多介绍,其中打靶法也包括了作者自己做过的研究课题的结果。

本书的主要对象是非计算数学专业的工科硕士研究生,具有相当于大学本科数学基础知识的工程技术人员也可以本书作为学习进修的用书,对部分工程学科的博士生也是有用的参考书。

本书的内容比实际教学的内容要多,在安排教学计划时,可以根据不同的对象和不同的要求对内容进行选择和重新组合。根据国家教委曾提出过的硕士研究生数值分析课的基本要求和作者教学实践,除去打“*”号的章节,其他的章节作为本书的主要内容,可以在 60 学时的课堂教学中完成。为了熟悉算法和熟悉应用数值计算软件,配合课堂理论教学可安排课后有 30~40 小时计算机的实习时间。

本书由文世鹏主编,并编写了第一章至第五章和第十章内容,承担审核修订全书。杨小远、张明、陈安乐共同参加编写工作,提供了第六,七,八,九,十一章内容。在成书过程中,梁景伟、潘建华也提供过部分章节初稿,卢名高、蔡璋琏提出过许多好的建设性意见,在此表示感谢。

书中错误和不当之处,恳请读者批评指正。

编　者

1999 年 5 月

目　　录

第一章　数值计算与误差分析

数值分析的主要内容是研究适合用计算机求解各种数学问题的数值计算方法，包括构造算法及数值计算的误差分析等问题。数值计算方法是计算数学中重要的基础学科，它有很强的理论性与实用性，并且与计算机科学紧密相关。近代随着计算机科学的迅速发展，数值计算方法已广泛地应用于自然科学与工程技术的各个领域之中，它是求解各种数学问题进行科学计算必不可少的主要方法。

本章作为《应用数值分析》教材的绪论，介绍有关数值计算与误差分析的基本概念和基础知识。

第一节　数值问题与数值方法

求解一个科学技术工程实际问题都是先把实际问题抽象近似地归纳为数学问题，建立一个数学模型。建立数学模型过程中要涉及许多不同的数学学科，但是工程实际问题中常用到的主要还是线性代数、矩阵计算、微积分和微分方程中的一些基本的数学问题。直接求解数学问题很难得到原问题能用数学解析式表达的解析解（即精确解）。真正实用而有效的方法是通过离散化将数学问题改变成一个数值问题。所谓数值问题是指可以在原问题自变量的原始数据和用输出数据表示的结果之间建立的一个确定的函数关系。对这个数值问题进行数值运算就可得到数值解（也称为计算解或近似解）。当然，也有的数学模型本身就是一个可以进行数值计算的数值问题。因此，求解工程实际问题进行科学计算所面对的实际上是数值问题。

求解数值问题必须要构造适合用计算机作为计算工具的数值方法，这种数值方法是指求解数值问题的、按确定顺序可以在计算机上依次执行的计算公式，也就是通常所说的数值计算方法。求解同一个数学问题可以构造多种算法，选用不同的数值方法。数值方法的好坏与原数学问题的性态有关，与数值方法本身的稳定性、计算解的精度和可靠性，以及计算量的大小有关，也还要考虑用计算机进行计算的特点等问题。作为研究生基础数学课的数值分析课程的主要内容就是讨论介绍有效的、实用的近代数值方法。

以计算机作为计算工具用数值方法求解数学问题时要设计计算机上的算法，即数值算法。数值算法有不同的种类，我们这里讨论的数值算法是指确定性的串行算法。在选用好的数值方法（Numerical Method）基础上，数值算法（Numerical Algorithm）是专门为在计算机上可进行确定计算而设计的算法，它不同于数值方法之处是它不是一系列的计算公式，而是一个完整的按规定顺序可执行的具体计算步骤。这个算法还要具备输入和输出数据及算法步骤的确定性，算法每一步在计算机上的可执行性以及通过有限次计算步骤能得到确定的数值结果。另外，还要考虑算法的复杂性使计算速度尽量快，占用计算机存储空间尽量少。通过数值算法可以编制出数值计算软件，使得科学计算更加方便。因此，数值方法不同于数值算法，而数值算法也不等同于程序设计。设计好的数值算法进而编制实用的数值软件要涉及计算机科学与软件工程等其他学科。

在数值分析课程中介绍重要的基本的数值方法，同时也要介绍和讨论一些好的数值算法，

使读者了解在求解数值问题的数值方法基础上如何进一步设计计算机上的数值算法，进而为编制程序作准备。但是在课程内不能做深入详细的讨论，课程的重点还是介绍计算机上常用的近代数值方法。同时，也要配合课程学习进行初步的数值算法设计和编制程序上机实习的训练，以提高学生应用计算机解决各种科学技术及工程实际问题的能力。

例 1－1 给出 x_0，计算多项式的值 $P_n(x_0)=a_nx_0^n+a_{n-1}x_0^{n-1}+\cdots+a_1x_0+a_0$。

这是一个数值问题，原始数据为 x_0 和系数 $a_k(k=0,1,\cdots,n)$。若按多项式的结构直接对各项分别作乘法再相加就是一个数值方法，可以输出数值结果 $P_n(x_0)$。但是，这种算法的计算量是

$$\text{乘法次数}=1+2+\cdots+n=n(n+1)/2$$

如果 $n=100$，则要进行大约 10^4 的乘法运算，这个计算量太大，而且随着计算次数的增加，计算误差也会增加。另外，用计算机计算时还需要占用 $2n+1$ 个存储单元。这个计算方法不能作为一个好的数值方法。但是，还可以构造另外的数值方法，为此将计算的结构和顺序改变一下，求 $P_n(x_0)$ 的计算式改写为

$$P_n(x_0)=x_0(\cdots x_0(x_0(a_nx_0+a_{n-1})+a_{n-2})+\cdots+a_1)+a_0$$

按这个公式计算的乘法次数只有 n 次，在计算机上只要 $n+2$ 个存储单元，这是一个好的数值方法。这个算法是我国南宋时期的著名数学家秦九韶于公元 1247 年首先提出的称为秦九韶算法，它比西方英国人霍纳(Horner)提出同样的算法要早 500 多年。在这个数值方法的基础上可以设计一个用于计算机计算的串行数值算法：

① 读入数据 $a_0,a_1,\cdots,a_n,x_0$

② 置 $S_n=a_n$

③ 对 $k=n-1,n-2,\cdots,1,0$

$S_k=x_0S_{k+1}+a_k$

④ 输出 $P_n(x_0)=S_0=x_0S_1+a_0$

这个算法也就是现在计算多项式的值所用的数值算法。

例 1－2 求解线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2=b_1\\a_{21}x_1+a_{22}x_2=b_2\end{cases}$$

这是一个线性代数中最基本而又是最简单的数学问题，但它本身就是一个离散形式的数值问题。可以构造不同的数值方法解出 x_1 和 x_2。比如，用熟知的克莱姆(Cramer)法则就得到以下的计算公式

$$x_1=\frac{b_1a_{22}-b_2a_{12}}{a_{11}a_{22}-a_{12}a_{21}}\quad\text{和}\quad x_2=\frac{b_2a_{11}-b_1a_{21}}{a_{11}a_{22}-a_{12}a_{21}}$$

这是一个求解原问题的数值方法，给出原始数据可以由数值计算得到解 x_1 和 x_2。但是，当方程组的规模增大时，这个数值方法就不能用来求解原数学问题。因为用克莱姆法则的计算公式求解 n 阶规模的线性方程组要计算 $n+1$ 个 n 阶行列式，而每一个 n 阶行列式要作 $(n-1)(n!)$ 次乘法。当 $n=20$ 时，为求出方程组的解就要进行大约 10×10^{20} 次的乘法，这样大的计算量就是用每秒钟可进行一千万次计算的计算机来计算也需要 300 万年才能得到计算结果，这是无法接受的。更何况实际上要求解的线性方程组的规模是 n 为数百、数千甚至上万，因此，这个数值方法是不能应用的。求解线性方程组有效的近代数值方法和数值算法将在第三章中介绍。

由以上实例可以看到数学问题和数值问题之间的关系，求解数值问题的数值方法和数值算法之间的关系以及为了求解数值问题研究构造实用有效的好的数值方法是多么重要。

第二节　数值计算的误差分析

数值计算中参与计算的数都是近似的有误差的数，所求出的计算解也是近似的解。因此，误差分析是数值计算中的重要问题，也是一个困难的问题。以下主要介绍误差分析的基本知识。

一、误差的来源

用数值计算方法求解一个实际的科学技术或工程问题一般包括以下几个阶段：

① 建立数学模型并转化为数值问题；

② 构造或选用数值方法；

③ 确定原始数据；

④ 设计数值算法编制程序或调用数值软件中的程序在计算机上求出计算解。

以上每一阶段都存在误差会影响计算解的精度。可以把所有这些误差归结为四类：模型误差、原始数据误差、方法误差和计算误差。

1. *模型误差*

为了用数值方法求解实际的问题，首先要忽略一些次要的因素提出简化的假设。在这些假设的条件下就可用简明的数学语言来描述实际问题和物理现象各种变量之间的关系，建立一个数学模型，然后才能用数值方法求解这个数学问题。显然这种简化的数学模型只是原问题的一个近似，它们之间的误差就是模型误差。数值方法无法改进由模型误差产生的解的不精确性。因此，数值分析中不讨论模型误差，只能假设数学模型是合理的。

2. *原始数据误差*

原始数据误差也称观测误差。数学模型中的一些物理量和原始数据都要通过观察和测量得到，而在观察和测量中要受到方法和仪器精度的限制必然会引进误差。这些误差会影响以这些数据为基础的任何计算解的精度。在数值分析中一般也不直接讨论这种误差，而是用分析舍入误差的方法来分析这些误差在计算中的传播和影响。

3. *方法误差*

方法误差又称截断误差，它是指用数值方法求解某个数学问题的逼近误差，其中包括用截取有限项近似无穷级数所产生的误差；也包括用容易计算的简化问题代替不便计算的复杂问题所带来的误差等。

例 1－3　已知 Taylor 展开式 $e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots + \frac{1}{n!}x^n + \cdots$，求 e^{-1} 的近似值，并估计误差。

解：利用展开式的前三项，取 $n = 2e^{-1} \approx 1 + (-1) + \frac{1}{2}(-1)^2 = 0.5$

截断误差　　$|R_2| = |e^{-1} - 0.5| \leqslant \frac{1}{3!} < 1.7 \times 10^{-1}$

4. *计算误差*

计算误差又称舍入误差，由于机器字长的限制使得任一个实数在机器中要被舍入而近似地表示成机器浮点数而引起的误差。所以，原始数据表示成机器数有误差，而每一次运算又可

能产生新的误差。这种舍入误差在数值计算的每一步计算中都会产生，并且逐步传播和扩散从而影响计算解的精度。

二、截断误差分析

1. 截断误差的估计方法

截断误差是由用一种近似的数值方法来求解数学问题而产生的误差。在函数的 Taylor 展开中，用近似的 Taylor 多项式代替函数 $f(x)$，有

$$f(x) = P_n(x) + R_n(x)$$

其中 $R_n(x) = \dfrac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}$ 是近似代替的余项，也就是用 $P_n(x)$ 近似 $f(x)$ 的截断误差。由于 ξ 值是未知的，这个截断误差的真正大小不可能知道。因此，截断误差的余项表达式 $R_n(x)$ 只有理论上的意义。那么，如何来估计截断误差呢？数值方法中有两种方法来估计截断误差。

①若已知存在正数 $M>0$，有 $|f^{(n+1)}(x)|\leq M, x\in[a,b]$，则有估计

$$|R_n(x)| \leq \frac{M}{(n+1)!}(x-x_0)^{n+1}$$

一般情况下，很难确切地知道 M 值。

②若 $n\to\infty$ 时 $R_n(x)\to 0$，则 Taylor 展开的级数收敛于 $f(x)$。此时 $x-x_0$ 是当 $x\to x_0$ 时的无穷小量，记 $h=x-x_0$，则 $R_n(x)=\dfrac{f^{(n+1)}(\xi)}{(n+1)!}h^{n+1}$ 是一个与 h^{n+1} 同阶的无穷小量，记 $R_n(x)=O(h^{n+1})$（注：这里是大"O"）。也可以说 $R_n(x)$ 是比 $h=x-x_0$ 高 $(n+1)$ 阶的无穷小（当 $h\to 0$ 时），称用 $P_n(x)$ 代替 $f(x)$ 的截断误差是 $O(h^{n+1})$，它表示用近似方法计算 $f(x)$ 的精确度是 $(n+1)$ 阶的。于是，Taylor 展开逼近的阶可表示为 $f(x)=P_n(x)+O(h^{n+1})$。

这里补充说明关于记号小"o"和大"O"的用法。小"o"是无穷小量记号，它们称为蓝道（Landau）记号，是数值分析中估计数值方法截断误差常用的记号。它们都是用来表示两个变量在极限过程中变化状态的相对关系。

2. 函数 $f(x)$，$g(x)$ 在 $x\to x_0$ 时的变化状态

① $\lim\limits_{x\to x_0} f(x)=0$，记 $f(x)=o(1)$，是 $x\to x_0$ 时的无穷小量；

$|f(x)|\leq A>0$，记 $f(x)=O(1)$，是 $x\to x_0$ 时的有界变量。

② $\lim\limits_{x\to x_0}\dfrac{f(x)}{g(x)}=0$，记 $f(x)=o(g(x))$， $f(x)$ 是比 $g(x)$ 高阶的无穷小量；

$\lim\limits_{x\to x_0}\dfrac{f(x)}{g(x)}=c\neq 0$，记 $f(x)=O(g(x))$， $f(x)$ 与 $g(x)$ 是同阶无穷小。

$|f(x)|\leq c\,|g(x)|, c>0$，也记 $f(x)=O(g(x))$， $\dfrac{f(x)}{g(x)}$ 是一个有界变量，常简记 $f=O(g)$。这个等式所表示的实际上是一个不等式 $|f(x)|\leq c\,|g(x)|$ （$c>0$，与自变量 x 无关的常数）。

3. Landau 记号可以运算，按定义可以推出以下关系式

① $o(1)\cdot O(f)=o(f)$，$O(1)\cdot O(f)=O(f)$，$o(1)+O(f)=O(f)$；

② $O(f)\cdot O(g)=O(f\cdot g)$；

③ $kO(f)=O(f)$，$O(kf)=O(f)$，k 是正常数，说明 Landau 记号与常数无关；

④ $(O(f))^m = O(f^m)$，m 是正整数；

⑤ $O(f)+O(g)=O(f+g)$，结合③有 $O(f)+O(f)=O(f)$；

⑥ $O(f^m)+O(f^{m+1})=O(f^m)$，m 是正整数。

例 1－4 当 $h \to 0$ 时，有如下 Taylor 展开式

$$e^h = 1 + h + \frac{h^2}{2!} + O(h^3), \sin(h) = h - \frac{h^3}{3!} + O(h^5)$$

试确定 $e^h + \sin(h)$ 和 $e^h \sin(h)$ 的截断误差。

解：① $e^h + \sin(h) = 1 + 2h + \frac{h^2}{2!} - \frac{h^3}{3!} + O(h^3) + O(h^5)$

因为 $\frac{h^3}{3!} + O(h^3) = O(h^3), O(h^3) + O(h^5) = O(h^3)$

于是 $e^h + \sin(h) = 1 + 2h + \frac{h^2}{2!} + O(h^3)$

② $e^h \sin(h) = (1 + h + \frac{h^2}{2!} + O(h^3))(h - \frac{h^3}{3!} + O(h^5))$

$= (1 + h + \frac{h^2}{2!})(h - \frac{h^3}{3!}) + (1 + h + \frac{h^2}{2!})O(h^5) + (h - \frac{h^3}{3!})O(h^3) + O(h^3) \cdot O(h^5)$

因为 $(1 + h + \frac{h^2}{2!})O(h^5) = O(h^5), (h - \frac{h^3}{3!})O(h^3) = o(1)$

$\frac{h^5}{2!\,3!} + O(h^5) = O(h^5), O(h^3) \cdot O(h^5) = O(h^8)$

于是 $e^h \sin(h) = h + h^2 + \frac{h^3}{2!} - \frac{h^3}{3!} - \frac{h^4}{3!} + O(h^5) + O(h^8) = h + h^2 + \frac{h^3}{3} - \frac{h^4}{6} + O(h^5)$

三、舍入误差分析

计算中的舍入误差和“数的误差”及“数值运算误差传播”密切相关。

1. 数的误差限和有效数字

(1)绝对误差和相对误差

定义 设数 a 是准确值，x 是 a 的一个近似值，记

$$e = |a - x|$$

$$e_r = |a - x| / |a| = e / |a|$$

称 e 为近似值 x 的绝对误差，称 e_r 为近似值 x 的相对误差。

由于准确值 a 总是无法知道的，所以数 a 的绝对误差 e 的准确值也不可能知道。但是，根据测量工具的精确度或测量计算的具体情况，可以事先估计出误差的绝对值不会超过某一个正数 δx，即

$$e = |a - x| \leqslant \delta x$$

称 δx 为数 a 的近似值 x 的绝对误差上界，或称为绝对误差限(简称误差限)。

另外，由于准确值 a 未知，按以上定义的相对误差 e_r 也无法知道。在实际计算时可改用

$$e_r = |a - x| / |x| = e / |x|$$

作为近似值 x 的相对误差。于是，当近似值与 δx 确定后，便得到相对误差的一个上界

$$\delta_r x = \delta x / |x|$$

称 $\delta_r x$ 为近似值 x 的相对误差限，显然有 $e_r = |a - x| / |x| \leqslant \delta_r x$。

(2)有效数字

对一个数的近似值，在计算中要关心的是哪几位数字是准确可靠的。而数的近似值可以看成是对准确值截取前面若干位得到的一个数，当截取的位数限定后，还要通过四舍五入才能得到绝对误差最小的近似值。

例1－5 已知准确值 $a=\pi=3.14159265\cdots$ 是一个无限不循环小数，求截取不同位数后的近似值及误差界。

解：截取一位 $x_1=3$，将 a 向右数第二位舍去，$e_1=|\pi-x_1|\approx 0.14<\frac{1}{2}\times 10^0$；

截取三位 $x_2=3.14$，将 a 向右数第四位舍去，$e_2=|\pi-x_2|\approx 0.0016<\frac{1}{2}\times 10^{-2}$；

截取五位 $x_3=3.1416$，将 a 向右数第六位舍去进一位，$e_3=|\pi-x_3|\approx 0.00001<\frac{1}{2}\times 10^{-4}$。

以上这三个 π 的近似值的绝对误差都不超过本身末位数的半个单位，是截取相应位数后所得到的近似数中绝对误差的最小值。对近似数的有效数字就是由此而定义的。工程上，数的规范化形式是用小数形式表示，如

$$x=\pm 0.a_1a_2\cdots a_n\times 10^p$$

其中，$p\in Z,n\in N,a_i(i=1,2,\cdots,n)$ 是 $0,1,2,\cdots,9$ 中任一数，但 $a_1\neq 0$（注：Z 表示整数集，N 表示正整数集）。

定义 设近似数 $x=\pm 0.a_1a_2\cdots a_n\times 10^p$ 的绝对误差限是第 n 位的半个单位，则数 x 有 n 位有效数字。

按此定义，若 x 有 n 位有效数字，则

绝对误差限由 $e=|a-x|\leqslant\frac{1}{2}\times 10^{p-n}$， 得 $\delta x=\frac{1}{2}\times 10^{p-n}$；

相对误差限由 $e_r=\frac{e}{|x|}\leqslant\frac{\delta x}{|x|}\leqslant\frac{1}{|x|}\times\frac{1}{2}\times 10^{p-n}\leqslant\frac{1}{2}\times 10^{p-n}/10^{p-1}=5\times 10^{-n}$，

得 $\delta_r x=5\times 10^{-n}$（这是偏于保守的估计式）。

但是，按此定义无法确定一个数的有效位数。因为任何数都只能是近似数，无法知道精确值，也就无法确定其绝对误差是小于哪一个位上的半个单位。实际计算中遵循两条原则：

① 已知精确值的一些常数，如 π，e，$\sqrt{2}\cdots$ 等，用四舍五入确定有效位数；

② 计算中的原始数据，认为其给出的位数都是有效位数。此原则只限于原始数据，不包括用原始数据作计算后得到的计算结果。

一个近似数的有效位数越多，则其相对误差越小，近似数也越精确。所以，在数值计算中要尽量设法多保留几位有效数字。另外，一个数的末位或末几位是零作为有效数字是有意义的。

如 $x=23.5=0.235\times 10^2,n=2,\quad \delta x=\frac{1}{2}\times 10^{2-3}=0.05$

而 $x=23.500=0.23500\times 10^2,n=5,\quad \delta x=\frac{1}{2}\times 10^{2-5}=0.0005$

2. 数值运算中的误差估计

若干近似数进行运算之后的结果也是近似数。这个计算结果的精确程度和参与运算的数的误差有关，也和运算过程有关。对运算结果进行误差分析是很复杂而困难的问题，简单的情况可以通过微分法则做出估计。

设有 n 个近似数 $x_1,x_2,\cdots,x_n$，其精确值虽未知但可设为 $a_1,a_2,\cdots,a_n$，把这 n 个近似数之