科学版

研究生教学丛书

数值分析原理

吴勃英 主编

吴勃英　王德明　丁效华　李道华 编

崔明根 主审

科学出版社
www.sciencep.com

科学版研究生教学丛书

数值分析原理

吴勃英 主编
吴勃英　王德明　丁效华　李道华 编
崔明根 主审

科学出版社
北京

内 容 简 介

本书介绍了常用数值计算方法的构造和使用，内容包括线性代数方程、非线性方程和方程组、常微分方程和方程组的数值解法，插值法与数值逼近，数值积分，矩阵的特征值和特征向量的计算等. 同时，对数值计算方法的计算效果、稳定性、收敛性、误差分析、适用范围及优缺点也作了必要的分析与介绍.

本书可作为高等院校各类工科专业研究生和数学系各专业本科生教材或参考用书，也可供从事科学与工程计算的科研工作者参考.

图书在版编目(CIP)数据

数值分析原理/吴勃英主编. —北京：科学出版社，2003
科学版研究生教学丛书
ISBN 7-03-011483-3

Ⅰ.数… Ⅱ.吴… Ⅲ.数值计算-研究生-教材 Ⅳ.O241

中国版本图书馆 CIP 数据核字(2003)第 036136 号

策划编辑：杨 波 吕 虹/文案编辑：邱 璐 贾瑞娜/责任校对：朱光光
责任印制：安春生/封面设计：黄华斌

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
源海印刷有限责任公司 印刷
科学出版社发行 各地新华书店经销
*
2003年8月第 一 版 开本：B5(720×1000)
2003年8月第一次印刷 印张：20 3/4
印数：1—4 000 字数：355 000

定价：30.00 元

(如有印装质量问题，我社负责调换〈环伟〉)

前　言

随着科学技术的发展，科学与工程计算愈来愈显示出其重要性，与实验、理论三足鼎立，成为科学实践的三大手段之一，其应用范围渗透到所有的科学活动领域. 作为科学与工程计算的数学工具，“数值分析”从 20 世纪 80 年代起，就相继成为各高等院校工科硕士研究生学位公共必修课.

本教材考虑到工科各专业对数值分析的实际需要，重点突出学以致用的原则，着重介绍在计算机上常用的数值计算方法的构造和使用，同时对数值计算方法的计算效果、稳定性、收敛性、误差分析、适用范围及优缺点也作了必要的分析与介绍. 教材中每章都配有难易程度不等的习题，有些习题必须通过上机实践来完成. 这样，能让学生通过习题来消化课堂内容，结合实验课中上机实习的要求，可使学生对所学数值方法有更深刻确切的理解.

由于编者水平所限，教材中难免有不妥之处，恳请读者指正，以便今后做进一步的修改.

编　者

2003 年 3 月

目　录

绪　　论

§0.1　研究数值分析的必要性

随着科学技术的发展,科学与工程计算已被推向科学活动的前沿. 科学与工程计算的范围扩大到了所有科学领域,并与实验、理论三足鼎立,相辅相成,成为人类科学活动的三大方法之一. 因此,熟练地运用计算机进行科学计算,已成为科技工作者的一项基本技能,这就要求人们去研究和掌握适用于计算机上使用的数值计算方法. 而数值分析就是研究用计算机解决数学问题的数值计算方法及有关理论.

一般地,用计算机进行科学与工程计算时要经历如下过程:

实际问题⇒数学建模⇒数值分析⇒算法研制⇒软件实现⇒程序的执行、分析⇒验证及结果的可视化

可见,数值分析是科学与工程计算过程中必不可少的环节. 它以纯数学为基础,但不只研究数学本身的理论,而着重研究解决问题的数值方法及效果,如怎样使计算速度最快、存储量最少等问题,以及数值方法的收敛性、稳定性、误差分析. 虽然有些方法在理论上还不够完善与严密,但通过对比分析、实际计算和实践检验等手段,被证明是行之有效的方法,也可采用. 因此,数值分析这门课程既带有纯数学高度抽象性和严密科学性的特点,又具有应用的广泛性和实际试验的高度技术性特点,是一门与计算机密切相连的实用性很强的计算数学课程.

例如,用 Cramer 法则解一个 n 阶线性代数方程组需要计算 $n+1$ 个 n 阶行列式. 不计加减运算,求解总共需要 $n!(n+1)+n$ 次乘除法. 当 n 很大时,这个计算量是相当惊人的. 比如一个不算太大 20 阶的方程组,大约要做 9.7×10^{20} 次乘除法,显然,这样的方法是毫无意义的. 然而,如果采用数值分析中介绍的任何一种解线性方程组的数值方法,比如 Gauss 消元法,乘除法次数不超过 3000 次,即使在微型计算机上,也只需几秒钟时间就能很容易地完成. 这个例子说明了研究实用的数值方法是非常有必要的. 而数值分析研究的正是在计算效率上最佳的或近似最佳的方法,而不是像 Cramer 法则这样的方法.

§0.2　误差来源与误差概念

对数学问题进行数值求解,求得的结果一般都包含有误差. 即数值计算绝大

多数情况是近似计算,因此,误差分析和估计是数值计算过程中的主要内容,通过它们可以确切地知道误差的性态和误差的界.

0.2.1 误差来源

数值结果中的误差通常来自固有误差与计算误差,如下面所示

$$
\text{误差来源}\begin{cases}\text{固有误差}\begin{cases}\text{模型误差}\\\text{观测误差}\end{cases}\\\text{计算误差}\begin{cases}\text{截断误差}\\\text{舍入误差}\end{cases}\end{cases}
$$

固有误差的一个来源是由求解问题的数学模型本身所固有的**模型误差**,它包括对实际物理过程进行近似的数学描述时所引进的误差. 另一个来源是由于物理数据的不精确性,这些数据往往是由实验观测而得到的,从而带有**观测误差**. 这些都不是数值分析所研究的内容.

计算误差主要有两个来源. 一个是由于在求解某一个已公式化的数学问题时,不是对其本身求解而是对它的某一个近似问题求解而造成的,称这类误差为**截断误差**或方法误差. 这类误差往往是由有限过程逼近一个无限过程时产生的. 比如,函数 e^x 可展开为幂级数形式

$$
e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots, \tag{0.2.1}
$$

如果用式(0.2.1) 右边前 $n+1$ 项

$$
S_n(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} \tag{0.2.2}
$$

来近似 e^x 的无穷多项的和,所产生的误差就是这一问题的截断误差,为

$$
e^x - S_n(x) = \frac{x^{n+1}}{(n+1)!}e^{\theta x}, \qquad 0 < \theta < 1. \tag{0.2.3}
$$

再比如序列

$$
\begin{cases} x_0 = 3, \\ x_{n+1} = \dfrac{1}{2}\left(x_n + \dfrac{5}{x_n}\right), \end{cases} \qquad n = 0,1,2,\cdots, \tag{0.2.4}
$$

因为 $\lim\limits_{n\to+\infty} x_n = \sqrt{5}$,所以,可以用无限迭代过程式(0.2.4) 的有限次结果来得到$\sqrt{5}$ 的近似值,而产生的误差也是截断误差.

通常这类误差的精确值是不能求得的(当然,若能知道误差的确切值,便能将

其消除). 所以,人们一般只研究这类误差的某一估计值或它的某一个界.

产生计算误差的另一重要来源,是由于算术运算几乎不可能在计算机上完全精确地进行. 首先,由于计算机所能表示的数字的位数有限(即字长有限),在进行计算时,对超过计算机所能表达的位数的,数字就要进行舍入;其次,尽管有些数据可以精确地由计算机表达,但是,当进行乘除运算时,常常也要对其运算的结果进行舍入,如计算$\frac{1}{7}$. 通过对某一个数进行舍入而产生的误差称为**舍入误差**.

一般情况,每一步的舍入误差是微不足道的,但是经过计算过程的传播和积累,舍入误差可能会对真值产生很大的影响. 甚至在一些情况中,一次的舍入就会大大改变计算结果.

0.2.2 绝对误差与相对误差

【定义 0.1】 设 $\tilde{x}$ 代表精确值 x 的一个近似值,称

$$E(\tilde{x}) = x - \tilde{x} \tag{0.2.5}$$

为近似值 $\tilde{x}$ 的**绝对误差**,或简称**误差**.

显然,绝对误差依赖于量纲,通常无法精确地算出绝对误差的真值,只能根据具体测量或计算的情况估计它的绝对值的范围,也就是去估计 $|E(\tilde{x})|$ 的上界. 若

$$|E(\tilde{x})| = |x - \tilde{x}| \leqslant \varepsilon, \tag{0.2.6}$$

称 ε 为 $\tilde{x}$ 的**绝对误差界**,或简称**误差界**.

在工程技术上,常将不等式(0.2.6) 表示成

$$x = \tilde{x} \pm \varepsilon.$$

绝对误差的大小,在许多情况下还不能完全刻画一个近似值的精确程度. 如有两个数

$$x = 10 \pm 0.1, \qquad y = 10^{15} \pm 10^{6},$$

这里 y 的绝对误差是 x 的 10^7 倍,但是不能就此断定近似值 $\tilde{x} = 10$ 一定比近似值 $\tilde{y} = 10^{15}$ 精确程度高. 若考虑到精确值本身的大小,在 10^{15} 内差 10^6 显然比在 10 内差 0.1 更精确些. 这说明一个近似值的精确程度,除了与绝对误差有关,还与精确值本身有关. 为此引入相对误差概念.

【定义 0.2】 设 $\tilde{x}$ 是精确值 x 的一个近似值,称

$$RE(\tilde{x}) = \frac{E(\tilde{x})}{x} = \frac{x - \tilde{x}}{x} \tag{0.2.7}$$

为近似值 $\tilde{x}$ 的相对误差.

相对误差是无量纲的,通常用百分数表示. 与绝对误差类似,我们只能估计相对误差绝对值的某一个上界. 当

$$|RE(\tilde{x})| \leqslant \varepsilon_r, \tag{0.2.8}$$

则称 ε_r 为近似值 $\tilde{x}$ 的**相对误差界**.

由于

$$\frac{x-\tilde{x}}{\tilde{x}} - \frac{x-\tilde{x}}{x} = \frac{(x-\tilde{x})^2}{\tilde{x}x} = \left(\frac{E(\tilde{x})}{\tilde{x}}\right)^2 \left[\frac{1}{1+\dfrac{E(\tilde{x})}{\tilde{x}}}\right],$$

当

$$\left|\frac{E(\tilde{x})}{\tilde{x}}\right| \leqslant \frac{1}{2},$$

有

$$\left|1+\frac{E(\tilde{x})}{\tilde{x}}\right| \geqslant 1 - \left|\frac{E(\tilde{x})}{\tilde{x}}\right| \geqslant \frac{1}{2},$$

从而

$$\left|\frac{x-\tilde{x}}{\tilde{x}} - \frac{x-\tilde{x}}{x}\right| \leqslant 2\left(\frac{E(\tilde{x})}{\tilde{x}}\right)^2.$$

当 $\left|\dfrac{E(\tilde{x})}{\tilde{x}}\right|$ 很小时,$\dfrac{x-\tilde{x}}{\tilde{x}}$ 与 $\dfrac{x-\tilde{x}}{x}$ 的差是 $\dfrac{E(\tilde{x})}{\tilde{x}}$ 的平方量级,可以忽略不计. 因此,在实际计算中,常取

$$RE(\tilde{x}) = \frac{x-\tilde{x}}{\tilde{x}}. \tag{0.2.9}$$

0.2.3　有效数字

我们表示一个近似数时,为了能反映它的精确程度,常常用到"有效数字"的概念.

【定义 0.3】　若 x 的某一近似值 $\tilde{x}$ 的绝对误差界是某一位的半个单位,则从这一位起直到左边第一个非零数字为止的所有数字都称为 $\tilde{x}$ 的有效数字.

具体地说,对于数 x,经四舍五入之后,得到它的近似值

$$\tilde{x} = \pm(x_1 \cdot 10^{-1} + x_2 \cdot 10^{-2} + \cdots + x_n \cdot 10^{-n}) \cdot 10^m, \tag{0.2.10}$$

其中，$x_1, x_2, \cdots, x_n$ 都是 0,1,2,…,9 这十个数字之一，$x_1 \neq 0$，n 是正整数，m 是整数. 如果 $\tilde{x}$ 的绝对误差满足

$$|x - \tilde{x}| \leqslant \frac{1}{2} \times 10^{m-n}, \tag{0.2.11}$$

我们称 $\tilde{x}$ 为 x 具有 n 位有效数字的近似值，也可以说它精确到第 n 位. 其中 $x_1, x_2, \cdots, x_n$ 都是 $\tilde{x}$ 的有效数字. 如果表示一个数的数字全是有效数字，则称此数为有效数.

【**例 0.1**】 按四舍五入原则分别写出数 0.03783551，e = 2.718281828…，0.002030002 具有 5 位有效数字的近似数.

【**解**】 按有效数字定义，上述各数具有 5 位有效数字的近似数分别是

$$0.037836, 2.7183, 0.0020300.$$

需要注意的是，有效数 0.00203 与 0.0020300 是不同的，前者具有 3 位有效数字，其绝对误差不超过 $\frac{1}{2} \times 10^{-5}$，而后者具有 5 位有效数字，其绝对误差不超过 $\frac{1}{2} \times 10^{-7}$.

由式(0.2.11)可见，n 越大，绝对误差界越小. 即**有效数字越多**，数字越准确，**绝对误差越小**.

下面再来讨论有效数字与相对误差的关系. 一个具有 n 位有效数字的近似值 $\tilde{x}$ 的相对误差可如下估计，由式(0.2.10)知

$$x_1 \cdot 10^{m-1} \leqslant |\tilde{x}| \leqslant (x_1 + 1) \cdot 10^{m-1},$$

所以

$$|RE(\tilde{x})| = \left|\frac{x - \tilde{x}}{\tilde{x}}\right| \leqslant \frac{\frac{1}{2} \times 10^{m-n}}{x_1 \cdot 10^{m-1}} = \frac{1}{2x_1} \cdot 10^{-(n-1)}. \tag{0.2.12}$$

这个结果说明，**有效数字越多**，**相对误差也越小**. 因此，在计算过程中，我们要尽量保留多的有效数字.

§0.3 数值计算中应注意的若干问题

我们用一些例子来说明数值计算中常遇到的一些应该注意的问题.

0.3.1 防止有效数字的损失

1. 相近两数相减有效数字会严重损失

【**例 0.2**】 用中心差商公式

$$f'(x) \approx \frac{f(x+h)-f(x-h)}{2h}, \tag{0.3.1}$$

求 $f(x)=\sqrt{x}$ 在 $x=2$ 的导数近似值.

【解】 根据所给公式

$$\frac{\mathrm{d}\sqrt{x}}{\mathrm{d}x} \approx \frac{\sqrt{x+h}-\sqrt{x-h}}{2h}, \tag{0.3.2}$$

用 5 位字长的数字计算,取 $h=0.1$ 得

$$\left.\frac{\mathrm{d}\sqrt{x}}{\mathrm{d}x}\right|_{x=2} \approx \frac{1.4491-1.3784}{0.2}=0.35350,$$

与导数精确值 $\frac{1}{2\sqrt{2}}=0.353553\cdots$ 比较,计算结果是可接受的. 然而,若取 $h=0.0001$,则由

$$\left.\frac{\mathrm{d}\sqrt{x}}{\mathrm{d}x}\right|_{x=2} \approx \frac{1.4142-1.4142}{0.0002}=0,$$

算出的结果完全失真. 出现这种现象的原因是由于计算机上数的表示受机器字长的限制. 当 h 很小时,发生两个值相近的数相减,损失了有效数字,甚至在计算机字长范围内,有效数字损失殆尽. 为避免损失有效数字,可将式(0.3.2) 改写成

$$\frac{\mathrm{d}\sqrt{x}}{\mathrm{d}x} \approx \frac{1}{\sqrt{x+h}+\sqrt{x-h}},$$

用这个公式,仍取 $h=0.0001$,计算出的值为 0.35356. 显然,这一结果有 4 位有效数字.

表达式 $\frac{\sqrt{x+h}-\sqrt{x-h}}{2h}$ 与 $\frac{1}{\sqrt{x+h}+\sqrt{x-h}}$ 从纯数学的角度,两者完全等价,没有任何差异. 造成上面计算效果的不同,完全是由于数值计算中的舍入误差. 再分析 Taylor 公式

$$f'(x)=\frac{f(x+h)-f(x-h)}{2h}+\frac{h^2}{6}f''(\zeta), \qquad x-h<\zeta<x+h,$$

即用式(0.3.1) 近似 $f'(x)$,其截断误差为 $\frac{h^2}{6}f''(\zeta)$,当 h 不太小时,近似计算的误差主要取决于截断误差,从理论上看,h 越小,截断误差也越小,逼近程度应该越好. 之所以在上面出现与这个结论矛盾的结果,是因为当 h 很小时,截断误差将变得微乎其微. 对结果的影响占主导地位的是舍入误差,而不是截断误差. 在一般

的数值计算中，截断误差与舍入误差之间常常处于这种矛盾之中，要解决它们之间的这种矛盾，通常的作法是，在满足给定的截断误差范围内，尽量选取大的步长 h.

2. 大数可能"吃掉"小数而使有效数字损失

【例 0.3】　在计算机上求二次方程 $ax^2+bx+c=0$ 的根.

【解】　由求根公式，得

$$x_1=\frac{-b+\sqrt{b^2-4ac}}{2a},\qquad x_2=\frac{-b-\sqrt{b^2-4ac}}{2a},$$

如果 $b^2\gg|ac|$，则$\sqrt{b^2-4ac}\approx|b|$，若用上面公式计算 x_1 和 x_2，其中之一将会损失有效数字. 原因就是由于在 b^2-4ac 中，大数 b^2"吃掉了"小数 $4ac$，并且公式之一中出现两个值相近的数相减. 如果改用公式

$$x_1=\frac{-b-\mathrm{sign}(b)\sqrt{b^2-4ac}}{2a},\qquad x_2=\frac{c}{ax_1}$$

就可以得到好的结果. 其中 sign(b) 是 b 的符号函数.

出现大数吃小数的现象，主要是参与计算的数之间数量级相差太大造成的. 在有些情形，大数吃掉小数不会引起结果的太大变化，如在计算 x_1 时，这些情形允许大数吃掉小数. 但是，在另一些情形则不允许. 为避免大数"吃掉"小数，一定要注意安排计算次序，使计算始终在数量级相差不大的数之间进行.

0.3.2　减少计算次数

对于 n 次多项式 $P_n(x)=a_nx^n+\cdots+a_1x+a_0$，要计算在某一点 x_0 的值 $P_n(x_0)$. 如果直接计算需要计算 $n+n-1+\cdots+1=\dfrac{n(n+1)}{2}$ 次乘法和 n 次加法. 如果把它写成

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

记 $s_n=a_n, s_k=s_{k+1}x_0+a_k, k=n-1,n-2,\cdots,1,0$，则只需要做 n 次乘法和 n 次加法就可计算出 $s_0=P_n(x_0)$，这大大减少了计算次数. 这就是计算多项式的著名的**秦九韶算法**.

【例 0.4】　利用 $\ln(1+x)=\sum\limits_{n=1}^{\infty}(-1)^{n+1}\dfrac{x^n}{n}$ 计算 ln2，要求精确到 10^{-5}.

【解】　如果直接计算，这需要计算 10 万项求和，才能达到精度要求，不仅计算量很大，而且舍入误差的积累也十分严重. 如果改用级数

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

取 $x = \frac{1}{3}$，只须计算前 9 项，截断误差便小于 10^{-10}.

0.3.3　避免使用不稳定的数值方法

一个数值方法如果输入数据有扰动(即误差)，而在计算过程中由于舍入误差的传播，造成计算结果与真值相差甚远，则称这个数值方法是**不稳定的**或是**病态的**. 反之，在计算过程中舍入误差能够得到控制，不增长，则称该数值方法是**稳定的**或**良态的**.

【例 0.5】　计算 $y_n = 10y_{n-1} - 1, n = 1,2,\cdots$，并估计误差. 其中 $y_0 = \sqrt{3}$.

【解】　由于 $y_0 = \sqrt{3}$ 是无限不循环小数，计算机只能截取其前有限位数，这样得到 y_0 经机器舍入的近似值 $\tilde{y}_0$，记 $\tilde{y}_n$ 为利用初值 $\tilde{y}_0$ 按所给公式计算的值，并记 $e_n = y_n - \tilde{y}_n$，则

$$y_n = 10^n y_0 - 10^{n-1} - 10^{n-2} - \cdots - 1,$$

$$\tilde{y}_n = 10^n \tilde{y}_0 - 10^{n-1} - 10^{n-2} - \cdots - 1,$$

$$e_n = y_n - \tilde{y}_n = 10^n(y_0 - \tilde{y}_0) = 10^n e_0.$$

这个结果表明，当初始值存在误差 e_0 时，经 n 次递推计算后，误差将扩大为 10^n 倍，这说明计算是不稳定的. 这种不稳定现象在数值分析中也经常会遇到，特别是在微分方程的差分计算中. 我们在实际应用中要选择稳定的数值方法，不稳定的数值方法是不能使用的.

在实际计算中，对任何输入数据都是稳定的数值方法，称为**无条件稳定**；对某些数据稳定，而对另一些数据不稳定的数值方法，称为**条件稳定**.

第一章　非线性方程和方程组的数值解法

§1.1　基本问题

本章讨论非线性方程组 $\boldsymbol{F}(\boldsymbol{x}) = 0$ 的数值求解方法. 其中 $\boldsymbol{F}$ 和 $\boldsymbol{x}$ 以及 $\boldsymbol{0}$ 都是 n 维向量. 当 $n = 1$ 时就是单个方程

$$f(x) = 0. \tag{1.1.1}$$

下面先讨论单个方程,对方程组情况在后面讨论.

在方程(1.1.1)中,若函数 $f(x)$ 是 n 次多项式,则称方程(1.1.1)为 n 次多项式方程或代数方程;若 $f(x)$ 是超越函数,则称方程(1.1.1)为超越方程. 理论上已证明,当次数 $n \leqslant 4$ 时,多项式方程的根可用求根公式表示,而次数 $n \geqslant 5$ 时,它的根一般已不可能用公式表示,也即不能用解析形式表示. 然而,对于函数方程(1.1.1),在实际应用中,一般并不需要得到求根的解析表达式,只要得到满足一定精度要求的根的近似值就可以了. 现在,我们主要讨论求方程(1.1.1)的单实根 α 问题,若对多个根都有兴趣,则可根据已求得的根 $\alpha_1, \cdots, \alpha_m$,利用

$$f_m(x) = \frac{f(x)}{\prod_{i=1}^{m}(x - \alpha_i)}$$

就能求出第 $m + 1$ 个根 α_{m+1}.

对方程(1.1.1)求根大致分三个步骤:

(1) 根的存在性:方程是否有根?如果有,有几个根?对于多项式方程,n 次方程有 n 个根.

(2) 根的隔离:把有根区间分成较小的子区间,每个子区间或者有一个根,或者没有根,这样可以将有根子区间内的任一点都可看成该根的一个近似值.

(3) 根的精确化:对根的某个近似值设法逐步精确化,使其满足一定的精度要求.

求根方法中最简单最直观的方法是**二分法**.

若函数 $f(x)$ 在$[a, b]$上连续,且 $f(a)f(b) < 0$. 为了讨论方便,不妨假设 $f(a) < 0, f(b) > 0$. 根据连续函数根的存在性定理可知,方程(1.1.1)在(a, b)区间上一定有实根. 称(a, b)是方程的有根区间,这里假设(a, b)区间内只有一个根 α,如图 1.1 所示.

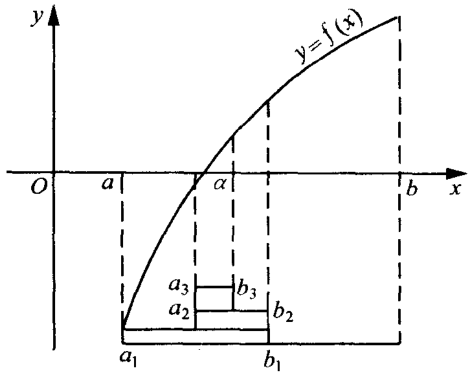

图 1.1　二分法

二分法的计算过程为:

(1) 取区间$[a,b]$的中点 $x_0=\dfrac{a+b}{2}$,并计算中点函数值 $f(x_0)$,判断:

若 $f(a)f(x_0)<0$,则有根区间为$[a,x_0]$,取 $a_1=a,b_1=x_0$,即新的有根区间为$[a_1,b_1]$;

若 $f(a)f(x_0)=0$,则 x_0 即为所求的根 α;

若 $f(a)f(x_0)>0$,则有根区间为$[x_0,b]$,取 $a_1=x_0,b_1=b$,即新的有根区间为$[a_1,b_1]$.

(2) 取区间$[a_1,b_1]$的中点 $x_1=\dfrac{a_1+b_1}{2}$,并计算中点函数值 $f(x_1)$,判断:

若 $f(a_1)f(x_1)<0$,则有根区间为$[a_1,x_1]$,取 $a_2=a_1,b_2=x_1$,即新的有根区间为$[a_2,b_2]$;

若 $f(a_1)f(x_1)=0$,则 x_1 即为所求的根 α;

若 $f(a_1)f(x_1)>0$,则有根区间为$[x_1,b_1]$,取 $a_2=x_1,b_2=b_1$,即新的有根区间为$[a_2,b_2]$.

此过程可以一直进行下去,则可得到一系列有根区间

$$[a,b]\supset[a_1,b_1]\supset[a_2,b_2]\supset\cdots\supset[a_n,b_n]\supset\cdots,$$

显然,$[a_n,b_n]$的区间长度为

$$b_n-a_n=\frac{b_{n-1}-a_{n-1}}{2}=\cdots=\frac{b-a}{2^n},$$

如图 1.1 所示. 这时,我们取最后一个区间的中点 x_n 作为方程(1.1.1)的根的近

似值

$$x_n = \frac{a_n + b_n}{2}, \qquad x_n \in [a_n, b_n], \tag{1.1.2}$$

其误差估计式为

$$|\alpha - x_n| \leqslant \frac{b_n - a_n}{2} = \frac{b - a}{2^{n+1}}, \tag{1.1.3}$$

当 $n \to \infty$ 时，$|\alpha - x_n| \leqslant \frac{b-a}{2^{n+1}} \to 0$，即 $x_n \to \alpha$.

对给定的小数 $\varepsilon > 0$，要使

$$|\alpha - x_n| < \varepsilon,$$

只须令

$$\frac{b-a}{2^{n+1}} < \varepsilon,$$

即

$$2^{n+1} > \frac{b-a}{\varepsilon},$$

所以

$$n = [(\ln(b-a) - \ln\varepsilon)/\ln 2], \tag{1.1.4}$$

其中，[ ] 表示取整数.

利用式(1.1.4)，对给定的精度，可预先确定出二分的次数.

二分法的优点是计算过程简单，收敛性可保证，对函数的性质要求低，只要求连续就可以了；它的缺点是计算收敛的速度慢，不能求偶数重根，也不能求复根和虚根. 特别是函数值 $f(a_k), f(b_k)(k = 0,1,2,\cdots)$ 每次均已计算出来，但没有利用上，只利用了它们的符号，显然是一种浪费.

§1.2 迭 代 法

在实际应用中，求方程(1.1.1)的根的主要方法是迭代法. 它的基本思想是，通过构造一个递推关系式，即迭代格式，计算出一个根的近似值序列，并希望该序列能收敛于方程(1.1.1)的根.

一般来讲，我们总可将方程(1.1.1)化成等价方程