目　录

第一章 绪 论

§1.1 数值计算方法的任务与算法的概念

快速电子计算机的发展，为科学技术中提出的各种数学问题求得数值解答提供了有力的工具。随着计算机广泛地用于进行大型的科学计算，计算数学方法已经成为应用数学的一个重要分支。

运用计算机解决科学计算问题需要经历以下几个主要的步骤：提出实际问题 ⟶建立数学模型⟶选用数值计算方法⟶程序设计⟶上机计算得出数值结果。"计算方法"的任务是研究适用于科学计算的数值计算方法及有关的数学理论，它是程序设计和对数值结果进行分析的依据和基础，是用计算机进行科学计算全过程的一个重要环节。

大家知道，电子计算机具有极高的运算速度，但它只能根据给定的指令完成加、减、乘、除等算术运算和一些逻辑运算。因此，要使用计算机来求解各种数学问题，诸如方程求根、微分和积分的计算、微分方程的求解等等，必须把求解过程归结为按一定的规则进行一系列四则算术运算的结果。计算机只能机械地执行人们所给定的指令，而不会主动地思维，去进行创造性的工作。交给计算机执行的解题方法的每一步骤都必须加以准确的规定。我们把对数学问题的解法归结为有加、减、乘、除等基本运算并有确定的运算顺序的完整而准确的描述，称为**数值算法**或简称**算法**。"计算方法"就是以数学问题为对象，研究各种数值算法及其有关理论的一门学科。

数值算法的优劣，需要考虑其计算工作量、计算程序的存贮量和逻辑结构的复杂程度，还要考察收敛性和稳定性来加以鉴定。

§1.2 误差知识

（一）误差的来源

科学计算中所处理的数据和计算的结果，通常都是在一定精度范围内的近似数值，它们与实际的精确值之间总存在着误差。引起误差的因素是多方面的，但主要来源有下面几种：

1．数学描述与实际问题之间的误差

各种实际问题的数学模型，总是在一定条件下抓住主要因素，忽略次要因素加以抽象得出的结果，是一种理想化了的数学描述，它与实际问题之间总存在误差，这种误差称为**“描述误差”**或**“模型误差”**。

2．观测误差

在各种计算公式中包含着一些已知数量（称为原始数据），这些数量往往是由观测和实验得到的，它们和实际的大小之间有误差，这种误差称为**“观测误差”**。

3．在计算中常遇到超越运算或极限运算，它们需要通过无穷过程才能求得精确的结果。但在实际上人们只能进行有限的算术运算，用有限的步骤来求得近似的结果。例如常用收敛的无穷级数的部分和来代替无穷级数本身以求得其近似值。这样由于截断一个无穷过程而引起的误差，称为**“截断误差”**。

4．舍入误差

对于参与计算的数，用有限小数代替无限小数，或用位数较少的有限小数代替位数较多的有限小数，例如用3.14代替π，用1.414代替$\sqrt{2}$等等，这样所产生的误差称为**“舍入误差”**。

误差的来源虽然有以上种种，且了解这些对于数值计算都是有帮助的，但是前两种误差往往不是计算工作者所能独立解决的。因此，在计算方法课程中通常只讨论截断误差和舍入误差。

（二）绝对误差·相对误差·有效数字

为了从不同的侧面表示近似数的精确程度，通常运用绝对误

差、相对误差和有效数字的概念。

设x^*代表准确值x的一个近似值。我们定义

$$e^* = x - x^* \tag{1.2.1}$$

为近似值x^*的**绝对误差**。通常无法准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大小的范围，即估计出$|e^*|$的上界。设存在一个正数ε^*，使

$$|e^*| = |x - x^*| \leqslant \varepsilon^* \tag{1.2.2}$$

则称ε^*是近似值x^*的**绝对误差限**，以后也把它简称为绝对误差。

由（1.2.2）式可知

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

即可知道准确值x所在的范围：必落在区间$[x^* - \varepsilon^*,\ x^* + \varepsilon^*]$内。在应用上常采用下面的写法

$$x = x^* \pm \varepsilon^*$$

来表示近似值x^*的精确度或准确值所在的范围。例如用有毫米刻度的米尺测量不超过1 m的长度x，读数法如下：如果长度接近于毫米刻度x^*，就读出那个刻度数x^*作为该长度的近似值。显然这个近似值的绝对误差为0.5mm。由这个例子还可看到绝对误差是有量纲单位的。

绝对误差的大小不能完全刻画近似值的准确程度。例如假设测量1km的长度产生了5 m的误差，而测量100m的长度产生了1 m的误差。就绝对误差而言，前者为后者的五倍，这是否说明后一种测量更精确呢？其实不然，如果考虑到被测量的数量本身的大小，前者在1 km＝1000 m中误差5 m，误差所占的比例为$\frac{5}{1000}=0.5\%$，后者的误差所占的比例为1%。显然应该说前一种测量要精确些。因此，刻画近似值的精确度，不仅要看绝对误差的大小，还要考虑近似数本身的大小，这就有必要引入相对误差的概念。

我们定义x的近似值x^*的**相对误差**为

$$e_r^* = \frac{e^*}{x^*} = \frac{x - x^*}{x^*}. \tag{1.2.3}$$

相对误差说明了x^*的绝对误差与x^*本身比较起来所占的比例，它可以反映一个近似数的精确程度。相对误差是一个无量纲量，通常可用百分数表示。

一般我们不能定出e_r^*的准确值，而只能估计它的大小的范围。如果

$$|e_r^*| = \left| \frac{x - x^*}{x^*} \right| \leqslant \varepsilon_r^* \tag{1.2.4}$$

我们就把这个上界ε_r^*称为x^*的**相对误差限**，或简称为相对误差。

为了给出一种近似数的表示法，使之既能表示其大小，又能表示其精确程度，我们来引进有效数字的概念。在计算中常按四舍五入原则得到数 x 的前几位近似值x^*，例如设

$$x = \pi = 3.1415926\cdots$$

经四舍五入，若取三位得 $x^* = 3.14$，$\varepsilon^* \leqslant 0.002$；若取五位得 $x^* = 3.1416$，$\varepsilon^* \leqslant 0.000008$。它们的绝对误差都不超过末位数字的半个单位，即

$$|\pi - 3.14| \leqslant \frac{1}{2} \times 10^{-2}$$

$$|\pi - 3.1416| \leqslant \frac{1}{2} \times 10^{-4}$$

如果近似值 x^*的绝对误差限是某一位的半个单位，从该位到 x^*的第一位非零数字共有 n 位数，我们就说x^*有 n 位**有效数字**。用严格的数学语言还可以如下定义：

将数 x 的近似值x^*写成形式

$$x^* = \pm 10^m (x_1 \times 10^{-1} + x_2 \times 10^{-2} + \cdots + x_n \times 10^{-n}) \tag{1.2.5}$$

若其绝对误差限满足

$$|x - x^*| \leqslant \frac{1}{2} \times 10^{m-n} \tag{1.2.6}$$

则称近似数x^*具有n**位有效数字**，其中x_1，x_2，…，x_n都是 x^* 的有效数字，x^*又称为有 n 位有效数字的有效数。这里 n 是正整

数，m是整数，x_1，x_2，…，x_n都是0到9中的一个数字，且假定$x_1 \neq 0$。根据定义易知π的近似值3.14具有三位有效数字，而近似值3.1416具有五位有效数字。注意有效数0.0023与0.002300是不同的，前者具有两位有效数字，其绝对误差不超过$\frac{1}{2}\times 10^{-4}$，后者具有四位有效数字，其绝对误差不超过$\frac{1}{2}\times 10^{-6}$。

上面的叙述说明了有效数字与绝对误差之间的关系。形如(1.2.5)的数，若具有n位有效数字，由(1.2.6)式表明：当m一定时，有效位数越多，则绝对误差越小。至于有效数字与相对误差的关系有如下的结论：

形如(1.2.5)的近似数x^*，若具有n位有效数字，则其相对误差

$$|e_r^*| \leqslant \frac{1}{2x_1}\times 10^{-(n-1)} \tag{1.2.7}$$

反之，若x^*的相对误差$|e_r^*| \leqslant \frac{1}{2(x_1+1)}\times 10^{-(n-1)}$，则$x^*$至少具有$n$位有效数字。

事实上，由式(1.2.5)知

$$x_1\times 10^{m-1} \leqslant |x^*| \leqslant (x_1+1)\times 10^{m-1}$$

所以

$$|e_r^*| = \left|\frac{x-x^*}{x^*}\right| \leqslant \frac{\frac{1}{2}\times 10^{m-n}}{x_1\times 10^{m-1}} = \frac{1}{2x_1}\times 10^{-(n-1)}$$

反之，由 $|x-x^*| = |x^*||e_r^*| \leqslant (x_1+1)\times 10^{m-1}\times \frac{1}{2(x_1+1)}$

$$\times 10^{-(n-1)} = \frac{1}{2}\times 10^{m-n}$$

故x^*具有n位有效数字。

这就表明：有效位数越多，相对误差越小。

注意从式(1.2.7)并不能保证x^*一定具有n位有效数字。例

如：$A=\sin 29^\circ 20'=0.4900$，设其近似值 $a=0.484$，其相对误差为

$$\frac{0.4900-0.484}{0.484}=0.012397<0.0125=\frac{1}{2\times 4}\times 10^{-1}$$

我们不能由此推出 a 有两位有效数字，这是因为

$$A-a=0.4900-0.484=0.0060>0.005$$

即可知近似值 a 并不具有两位有效数字。

（三）误差的传播

数值运算中的误差估计情况比较复杂，通常可以利用台劳展开式的方法估计误差。设想把所有的原始数据 $a_1,\ a_2,\ \cdots,\ a_n$ 看成是互相独立的自变量，而把通过某个算式所得的结果看作是这些原始数据的函数，表示为

$$A=f\,(a_1,\ a_2,\ \cdots,\ a_n) \tag{1.2.8}$$

设 $a_1^*,\ a_2^*,\ \cdots,\ a_n^*$ 是 $a_1,\ a_2,\ \cdots,\ a_n$ 的近似值，则 A 的近似值为

$$A^*=f\,(a_1^*,\ a_2^*,\ \cdots,\ a_n^*) \tag{1.2.9}$$

假定多元函数 f 在点 $a=(a_1^*,\ a_2^*,\ \cdots,\ a_n^*)$ 处可微，于是函数值 A^* 的绝对误差可利用台劳展开得到

$$\begin{aligned} e^*(A)&=A-A^*=f\,(a_1,\ a_2,\ \cdots,a_n)-f\,(a_1^*,a_2^*,\cdots,a_n^*)\\ &\approx\left(\frac{\partial f}{\partial a_1}\right)_a(a_1-a_1^*)+\cdots+\left(\frac{\partial f}{\partial a_n}\right)_a(a_n-a_n^*)\\ &=\sum_{i=1}^{n}\left(\frac{\partial f}{\partial a_i}\right)_a e^*(a_i) \end{aligned} \tag{1.2.10}$$

函数值 A^* 的相对误差为

$$e_r^*(A)=\frac{e^*(A)}{A^*}\approx\sum_{i=1}^{n}\left(\frac{\partial f}{\partial a_i}\right)_a\frac{a_i^*e_r^*(a_i)}{A^*} \tag{1.2.11}$$

我们可以利用（1.2.10）与（1.2.11）来估计按函数 f 的计算误差，给定函数 f 的具体形式，就可以得出具体的估计。特别地可以得到和、差、积、商的误差估计。

例 和的误差估计

设 $$A=f(a_1,\ \cdots,\ a_n)=a_1+\cdots+a_n$$

这里 $\dfrac{\partial f}{\partial a_i}=1$，所以有

$$e^*(A)=e^*(a_1)+\cdots+e^*(a_n)$$

$$|e^*(A)|\leqslant\sum_{i=1}^{n}|e^*(a_i)| \tag{1.2.12}$$

即和的绝对误差不超过各加数的绝对误差之和。为考察相对误差，再假设$a_i^*>0$，可得

$$|e_r^*(A)|\leqslant\max_i|e_r^*(a_i)| \tag{1.2.13}$$

即和的相对误差不超过相加各项中最不准确的一项的相对误差。

(四) 选用算法的若干注意之点

解决一个计算问题往往有多种算法，用不同算法计算的结果其精确度是不同的。人们自然希望选用那些计算量小而精度又高的算法。但如何给出最好的算法没有什么固定的办法，我们只是按照一般的情况，指出选用算法的若干注意之点。

1. 简化计算步骤以减少运算次数

减少运算次数，既可以提高解题速度，又有可能使计算中的误差积累降低。

例如要计算多项式

$$p_n(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0 \tag{1.2.14}$$

的值。如果直接计算a_kx^k再逐项相加，一共需做$\dfrac{1}{2}n(n+1)$次乘法和n次加法。但若按下列递推算法（称为秦九韶算法）：

$$\begin{cases}u_n=a_n\\ u_k=xu_{k+1}+a_k\quad(k=n-1,\ n-2,\cdots,2,1,0)\\ p_n(x)=u_0\end{cases} \tag{1.2.15}$$

则只要n次乘法和n次加法就可算出$p_n(x)$的值。

又如要计算和式 $\sum_{n=1}^{1000} \frac{1}{n(n+1)}$ 的值，如果直接逐项求和，其运算次数多且误差积累也不小。但和式可以简化：

$$\sum_{n=1}^{1000} \frac{1}{n(n+1)} = \sum_{n=1}^{1000} \left(\frac{1}{n} - \frac{1}{n+1} \right)$$

$$= 1 - \frac{1}{1001} \tag{1.2.16}$$

则整个计算只要一次求倒数和一次减法了。

2. 尽量避免相近的数相减

如果要计算 $y = x - A$，假定 A 为准确数值，运算时不发生误差，而 x 的近似值为 x^*。易知 y 的相对误差为

$$e_r^*(y) = \frac{e^*(x)}{x^* - A} \tag{1.2.17}$$

当 x^* 很接近于 A 时，$|e_r^*(y)|$ 就会变得很大。相对误差很大，有效数字位数就很少。为了避免这种情形出现，常常可以采用改变原来算式的办法来改善计算的效果。例如：

当 x 充分大时变换

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

当 x 充分大时变换

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

当 x 接近于 0 时变换

$$\frac{1-\cos x}{\sin x} = \frac{\sin x}{1+\cos x}$$

等等。

3. 应该选用数值稳定性好的计算公式

例如要计算积分：

$$I_n = e^{-1} \int_0^1 x^n e^x dx \tag{1.2.18}$$

由分部积分法可以得到计算I_n的递推公式：

$$\begin{cases} I_0 = e^{-1}\int_0^1 e^x dx = 1 - e^{-1} \approx 0.6321 \\ I_n = 1 - nI_{n-1} \quad (n = 1, 2, \cdots) \end{cases} \tag{1.2.19}$$

利用上式，用四位小数计算依次得到：

0.6321，0.3679，0.2642，0.2074，0.1704，

0.1480，0.1120，0.2160，−0.7280，7.5520

由此看到I_8为负值，显然与一切$I_n > 0$矛盾。事实上由估计式：

$$\frac{e^{-1}}{n+1} = e^{-1}\left(\min_{0 \leqslant x \leqslant 1} e^x\right)\int_0^1 x^n dx < I_n$$

$$< e^{-1}\left(\max_{0 \leqslant x \leqslant 1} e^x\right)\int_0^1 x^n dx = \frac{1}{n+1} \tag{1.2.20}$$

必有$I_7 < \frac{1}{8} = 0.125$，但上面算得$I_7 = 0.2160$已经连一位有效数字也没有了。发生这个现象的原因是，因为I_0带有不超过$\frac{1}{2} \times 10^{-4}$的误差，虽然在以后的递推计算中没有新的舍入误差，但这个初始数据的误差在以后的每次计算时顺次乘以$n = 1, 2, 3, \cdots$而传播积累到I_n中，使得算到I_7时就完全不精确了。

如果将递推算式（1.2.19）改写为

$$I_{n-1} = (1 - I_n)/n \tag{1.2.21}$$

再由$\frac{e^{-1}}{10} < I_9 < \frac{1}{10}$，粗略取$I_9 \approx \frac{1}{2}\left(\frac{1}{10} + \frac{e^{-1}}{10}\right) = 0.0684$，按算式（1.2.21）对$n = 9, 8, \cdots, 1$倒推计算，计算中小数点后第5位按四舍五入，得$I_9$，$I_8$，…，$I_0$顺次为：

0.0684，0.1035，0.1121，0.1268，0.1455，

0.1709，0.2073，0.2642，0.3679，0.6321

可见$I_0 = 0.6321$全部为有效数字。这样计算的结果如此精确的

原因是，I_9的误差传播到I_8时要乘以$\frac{1}{9}$，直到计算I_0时，I_9的误差已缩小为$\frac{1}{9!}$倍。

我们把运算过程中舍入误差不增长的计算公式称为**数值稳定**的，否则称为**不稳定**的。例如递推算式（1.2.21）是数值稳定的，而（1.2.19）是不稳定的算式。我们自然应当选用数值稳定性好的计算公式。

习 题 一

1．下列各近似值的误差均不超过末位数的半个单位，试求其和并估计其和的绝对误差：

136.421，28.3，231，62.243，17.482

2．设有三个数：$a=325\pm 1$，$b=3.25\pm 0.01$，$c=0.00325\pm 10^{-5}$。其近似值$a^*=325$，$b^*=3.25$，$c=0.00325$，哪个近似值的精确度高？为什么？

3．设下列各对近似值均为有效数，问它们是否一样？若不一样有何区别？

（1）45800与458×10^2

（2）0.00438与0.04380×10^{-1}

（3）0.4015×10^2与0.04015×10^3

（4）9800与98×10^2

（5）0.8070与0.807

4．设下列近似值均为有效数，试求各数的绝对误差、相对误差和有效数字的位数。

（1）3580　　（2）0.00476

（3）0.1430×10^3　　（4）2958×10^{-2}

（5）10.1010×10

5．取3.14，3.15，$\frac{22}{7}$，$\frac{355}{113}$作为π的近似值，求各自的绝对误差、相对误差和有效数字的位数。

6．求$\sqrt{10}$的近似值a^*，使相对误差不超过0.1%。

7．要使$\sqrt{20}$的近似值的相对误差限小于0.1%，要取几位有效数字？

8．利用（1.2.10）、（1.2.11）式导出二数之积、商的误差公式。

9．设已测得某长方形场地长l的值为$l^*=110$米，宽d的值$d^*=80$米。若已知$|l-l^*|\leqslant 0.2$米，$|d-d^*|\leqslant 0.1$米，试求其面积的绝对误差与相对误差。

10．设$s=\frac{1}{2}gt^2$，假定g是准确的，而对t的测量有± 0.1秒的误差，证明当t增大时，s的绝对误差增大而相对误差却减小。

11．设$a=1000$，取四位数字进行运算，求近似值x^*：

（1）按算式$x=\sqrt{a+1}-\sqrt{a}$计算

（2）按算式$x=\dfrac{1}{\sqrt{a+1}+\sqrt{a}}$计算

将二结果与准确值$x=0.015807437\cdots$比较，各有多少位有效数字。

第二章　方程的近似解法

当 $f(x)$ 是多项式或超越函数时，则称 $f(x)=0$ 为代数方程或超越方程。理论上已经证明，不高于四次的代数方程的根，可以用根式来表示，但高于四次的代数方程的根一般已不能用根式表示，即不存在根的解析表达式。对于超越方程，一般更不存在根的解析表达式。另一方面，在实际应用中，并不一定要求得到根的解析表达式，而只要求获得具有一定准确程度的近似值就可以了。本章我们讲述求方程根的近似值的一些实用方法。

求方程 $f(x)=0$ 的根可分两步来做：首先求一个比较粗糙的近似值，称为初始近似值；然后按某种方法将近似值逐步精确化，直到满足所要求的精确度。为了确定根的初始近似值，通常是设法寻找一个区间 (a, b)，使其中恰恰只含有方程 $f(x)=0$ 的一个根，这个步骤叫做**“根的隔离”**，这样的区间 (a, b) 叫做**隔根区间**。只要隔根区间适当小，则该区间内的任一点都可作为根的初始近似值。

将根隔离确定初始近似值的简单方法有：

1. **描图法**：画出 $y=f(x)$ 的略图，观察曲线与 x 轴交点的位置。

2. **试算法**：适当试算 $f(x)$ 的一些值，观察其符号改变的情况。

现在举例说明如下：

〔**例 1**〕　$f(x)=x^3-3x^2+4x-3=0$

先作 $y=f(x)$ 的略图。由于 $f'(x)=3x^2-6x+4=3(x-1)^2+1>0$，$f''(x)=6(x-1)$，故知 $y=f(x)$ 单调增大，而在 $x<1$ 时是凸的，在 $x>1$ 时是凹的，$x=1$ 时 $y=f(x)$ 有拐点。故其粗略的图形如图 2-1。

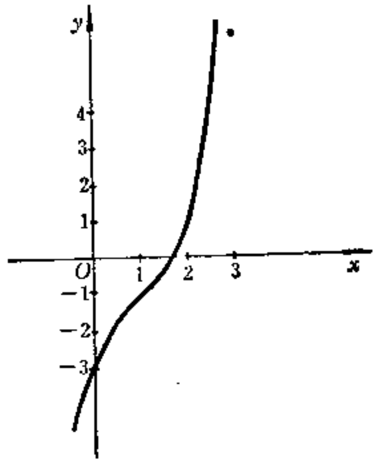

图 2-1

又因$f(0)=-3$，$f(1)=-1$，$f(2)=1$，故知在$(1,2)$内有$f(x)=0$的一个根α。从图上大体看出：

$$1.5<\alpha<2$$

〔例2〕 $f(x)=3x-1-\cos x=0$

因$f'(x)=3+\sin x>0$，且$x\to-\infty$时，$f(x)\to-\infty$，$x\to+\infty$时$f(x)\to+\infty$。故$f(x)=0$存在唯一的一个根α。

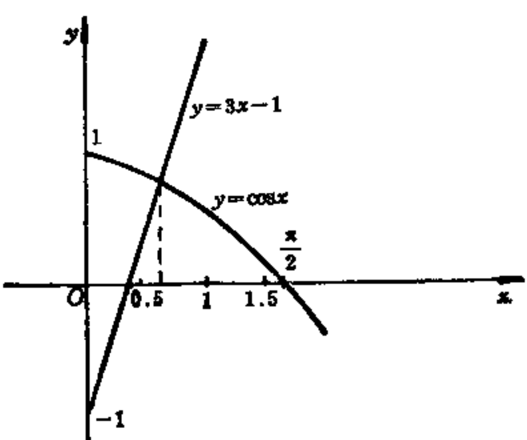

图 2-2

为确定根的位置，最好把这个方程改写为

$$3x-1=\cos x$$

因为 $y=3x-1$ 与 $y=\cos x$ 的图形是大家所熟知的。画出这两个函数的略图如图 2-2，其两曲线的交点的横坐标即为方程 $f(x)=3x-1-\cos x=0$ 的根。从图上可以看出，交点的横坐标约为 0.6。

下面介绍几种使根逐步精确化的方法。

§2.1 对 分 法

求解方程 $f(x)=0$。如果函数 $f(x)$ 在 $[a, b]$ 上单调连续，且在区间的两端点的函数值 $f(a)$ 和 $f(b)$ 异号（这些假设可由“根的隔离”实现），那末方程 $f(x)=0$ 在 (a, b) 内的唯一实根 α，可以用下述方法来逐次逼近。

把区间 $[a, b]$ 二等分，分点为 $\frac{1}{2}(a+b)$，计算函数值 $f\left(\frac{a+b}{2}\right)$，如碰巧 $f\left(\frac{a+b}{2}\right)=0$，就得到方程的实根 $\alpha=\frac{a+b}{2}$。否则 $f\left(\frac{a+b}{2}\right)$ 或者与 $f(a)$ 异号，或者与 $f(b)$ 异号。前一种情形取 $a_1=a$，$b_1=\frac{a+b}{2}$（即丢弃原来区间的右半部分）；后一种情形取 $a_1=\frac{a+b}{2}$，$b_1=b$（即丢弃原来区间的左半部分）。于是得到一个长度只有原来一半的区间 $[a_1, b_1]$，即 $b_1-a_1=\frac{1}{2}(b-a)$，而在这个新区间两端的函数值仍异号。再把区间 $[a_1, b_1]$ 二等分，计算函数值 $f\left(\frac{a_1+b_1}{2}\right)$，等等，重复上述过程，经过 n 步后得到区间 $[a_n, b_n]$，方程的实根 α 一定在这个区间中。取区间 $[a_n, b_n]$ 的中点 $\frac{1}{2}(a_n+b_n)$ 作为方程的根 α 的近似值，则误差不超过原区间长