2018-2019学年第一学期

计 算 方 法

第二讲：误差分析简介

第一章 §4

主讲人：张治国

zgzhang@szu.edu.cn

深圳大学 医学部 生物医学工程学院

SHENZHEN UNIVERSITY HEALTH SCIENCE CENTER SCHOOL OF BIOMEDICAL ENGINEERING

上节课回顾

- 介绍了本课程的目标和意义
  - 计算方法是一种研究并解决数学问题的数值近似解方法，是一种在计算机上使用的解决数学问题的方法。
- 研究数值方法的基本内容和重点：
  - 数值方法的基本内容：将计算机上不能执行的运算化为在计算机上可执行的运算；
  - 算法及其设计：针对所求解的数值问题，研究在计算机上可执行的且有效的计算公式；
  - 误差分析：即数值问题的性态及数值方法的稳定性。

上节课回顾

- 介绍了科学计算的过程：将数学模型变成数值问题，研究求解数值问题的数值方法，进而设计有效的数值算法
  - 数值问题： 输入数据与输出数据之间函数关系的一个确定而无歧义的描述
  - 数值方法：指求解数值问题在计算机上可执行的一系列运算和数据
  - 数值算法：有步骤地完成求解数值问题的过程
- 介绍了算法设计的目的和表示方法（自然语言法与图示法）

本节课内容

§4 误差分析简介

4-1 误差的基本概念

- 舍入误差和截断误差
- 绝对误差和相对误差
- 有效数字

4-2 浮点基本运算的误差

4-3 数值方法的稳定性与算法设计原则

误差的基本概念

- 数值计算中的误差来源有两种
  1) 舍入误差：由于计算机的字长有限，原始数据在计算机中的表示、运算产生的误差；
  2) 截断误差：由数学问题化成数值问题产生的误差

- 例（舍入误差）：
  若计算机仅能表示6位十进制数，则将 π 表示为

$$\pi^* = 3.141\,59\ ,\ \text{误差}\quad R_1 = \pi - \pi^* = 0.000\ \ 0026\cdots$$

若将其与数 9.210 00 进行加法运算，得

$$s = 3.141\ \ 59 + 9.210\ \ 00 \approx 1.235\ \ 16 \times 10$$

$$R_2 = 12.351\ \ 6 - 12.351\ \ 59 = 0.000\ \ 01$$

R_1 和 R_2 都是舍入误差。

误差的基本概念

- 例（截断误差）：
  计算 e^x 的数值。因

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

由算法的有限性，故利用截断部分和

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

来近似代替，由此产生的误差即为截断误差，为

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

误差的基本概念

- 例（截断误差）：
  计算 e^x 的数值。因

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

MATLAB代码：取 $x = 2, n = 10$

```
x = 2;
S = 1;
for n = 1:10
     S =  S + 1/factorial(n) * x^n;
end
```

Matlab 运行结果 7.3890，而 $e^x = 7.3891$。

若取 $n = 5$，运行结果 7.2667 。

绝对误差与相对误差

- 绝对误差：设 x 为准确值, x^*为 x 的一个近似值，称

$$E(x^*) = x^* - x$$

为近似值 x^*的（绝对）误差，简记为 E。

- 例：圆周率的近似值 $a = 3.14$, 绝对误差 $E = 0.00159\cdots$

- 一般来说，E 的准确值很难求出，只能估计出 $|E|$ 的某个上界 $\varepsilon(x^*)$，即

$$|E| = |x - x^*| \le \varepsilon(x^*)$$

$\varepsilon(x^*)$ 称为近似值 x^* 的（绝对）误差限，简记为 ε 。
于是 $x^* - \varepsilon \le x \le x^* + \varepsilon$ 可表示成：$x = x^* \pm \varepsilon$ 。

绝对误差与相对误差

- 相对误差：近似值 x^* 的误差与准确值 x 之比

$$E_r(x^*) = \frac{E(x^*)}{x} = \frac{x^* - x}{x}$$

称为近似值 x^*的相对误差，简记为 E_r 。

- E_r 绝对值的任一上界 $\varepsilon_r(x^*)$，称为相对误差限 ε_r，

$$|E_r(x^*)| = \left|\frac{E(x^*)}{x}\right| = \left|\frac{x^* - x}{x}\right| \le \varepsilon_r(x^*)$$

- 由于 x 的准确值难以确定，通常利用

$$E_r^* = E_r^*(x^*) = \frac{E(x^*)}{x^*} \text{ 代替 } E_r(x^*)$$

- 注：绝对误差限与相对误差限不唯一，它们越小越好。

有效数字

- 对于准确值 x 取近似值最常用的方法是采用"四舍五入"的原则。由此产生了一个专有名词—有效数字。
- 有效数字：如果近似值 x^* 的误差的绝对值不超过某一位数字的半个单位，且该位数字到 x^* 的第一位非零数字共有 n 位，则称用 x^*（近似 x 时）具有 n 位有效数字。（最大的 n 值）

- 例如：$\pi^* = 3.1416$ 有 5 位有效数字，
而 $\pi^* = 3.1415$ 有 4 位有效数字。

有效数字

- 确定有效数字的等价方法：数的规格化表示

$$x = \pm 10^k \times 0.\alpha_1 \alpha_2 \cdots \alpha_n \alpha_{n+1} \cdots \qquad (4.2)$$

其中，k 是某个范围内的整数，$\alpha_1 \neq 0$，α_i 为 0, 1, … , 9 中的数字。

- 若将近似值 x^* 表示成 (4.2) 式，且 x^* 满足不等式

$$|x^* - x| \leq 0.5 \times 10^{k-n} \quad （最大的 n 值） \qquad (4.5)$$

则称 x^* 具有 n 位有效数字。

有效数字

- 例：$\pi^* = 3.1416 = 10 \times 0.31416$，有 $|\pi^* - \pi| \le 0.5 \times 10^{1-5}$，因此其有五位有效数字。

- 特殊情况：有效数字不唯一。如：

  $x_1^* = 4.0$ 和 $x_2^* = 3.9$ 都是 $x = 3.95$ 的2位有效数字。

- 一般地，近似值的有效位数越多，误差的绝对值越小。但也有个别例外，如对于 $x = 1000$：

  $$x_1^* = 999.9\text{，}\quad x_2^* = 1000.1$$

  分别有3和4位有效数字。

有效数字

- 有效数字位数与小数点的位置无关。只有经过四舍五入写成如(4.2)的规格化形式后，小数点后的数字位数才能反映其有效位数的多少。

- 有效数字小数点后面的零不能随便增减。

- 通过对某个数进行四舍五入取近似值可以得到它的有效数字。不是通过四舍五入得到的近似数，它的数字并不都是有效数字。

有效数字

- 近似值的有效数字与相对误差之间的关系：
  定理4.1 设 x^*是 x 的近似值，它的表达式为
  $$x^* = \pm 10^k \times 0.\alpha_1\alpha_2\cdots\alpha_n$$
  则 x^*的有效位数与 x^*的相对误差之间有如下关系：
  1) 若 x^*具有 n 位有效数字，则 x^* 的相对误差 E_r^* 满足
  $$|E_r^*| \le \frac{1}{2}\times 10^{-n+1} \tag{4.6}$$
  2) 若 x^* 的相对误差 E_r^* 满足 $|E_r^*| \le \frac{1}{2}\times 10^{-n}$ (4.7)
  则 x^*至少具有 n 位有效数字。

  证明：p19

- 近似值的有效位数越多，相对误差越小；反之，相对误差越小，近似值有效数字的位数就可能越多。

有效数字

- 例： 设 $x = 1.986$，$x^* = 1.98$，则

$$|E_r^*| = \frac{|x^* - x|}{|x^*|} = \frac{|1.986 - 1.98|}{1.98} \le \frac{1}{2} \times 10^{-2}$$

故，x^*至少具有2位有效数字。

- 因 $|x^* - x| = 0.006 \le 0.5 \times 10^{-1}$ ，由(4.5)，$k - n = -1$， $k = 1$，$n = k + 1 = 2$，所以 x^*具有2位有效数字。

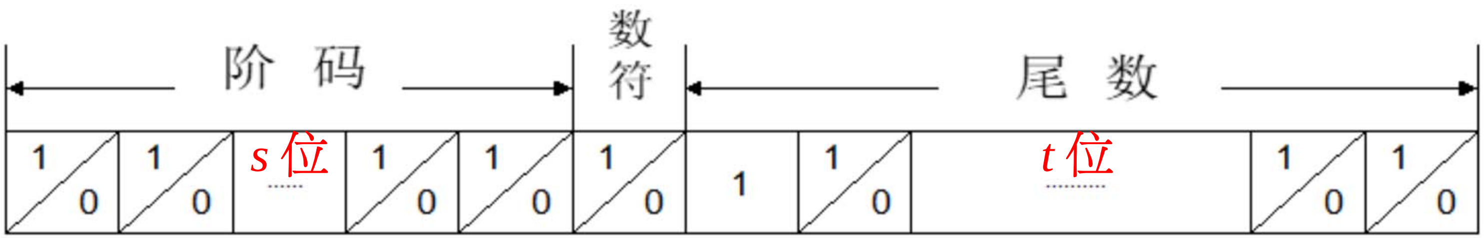
- 另一方面： x^*具有2位有效数字，故由(4.6)，得

$$|E_r^*| \le 0.5 \times 10^{-2+1}$$

注：上界 $0.5 \times 10^{-2+1}$ 并不是最小的。

浮点基本运算的误差

- 在计算机中，每个数都用有限位二进制数表示，其规格化浮点数形式由三部分组成：
  1) 阶码：确定小数点的位置；决定了数的取值范围
  2) 数符：表示正、负号；‘0’表示正，‘1’表示负
  3) 尾数：表示机器数字长；长度与精度有关



- 形式为：$x^* = \pm 2^{\alpha} \times 0.\beta_1\beta_2\cdots\beta_t$ ，其中，2称为浮点数的基，$\alpha \in [-2^s, 2^s - 1]$，$\beta_1 = 1$，$\beta_2, \beta_3, \cdots, \beta_t$是0或1；称 x^*为 x 的浮点数，记为 $x^* = fl(x)$。

浮点基本运算的误差

- 设原始数据 $x = 2^{\alpha} \cdot \beta$ ，其尾数

$$\beta = \pm 0.\beta_1 \beta_2 \cdots \beta_t \cdots$$

若浮点数 $fl(x)$ 的尾数 β^* 是 t 位：$\beta^* = \pm 0.\beta_1 \beta_2 \cdots \beta_t$

则 $$fl(x) = 2^{\alpha} \cdot \beta^*$$

$$0.10\cdots 0 \le |\beta^*| \le 0.11\cdots 1$$

$$2^{\alpha}(\beta^* - 2^{-t}) \le x \le 2^{\alpha}(\beta^* + 2^{-t})$$

于是绝对误差 $E = |x - 2^{\alpha}\beta^*| = |x - fl(x)| \le 2^{\alpha - t}$

- $fl(x)$ 的相对误差

因 $|x| = |2^{\alpha} \cdot \beta| \ge 2^{\alpha} \cdot 0.10\cdots = 2^{\alpha-1}$，故 $|\varepsilon| = |E/x| \le 2^{-t+1}$ 。

浮点基本运算的误差

一. 浮点数的四则运算的误差

- 浮点数 x 和 y 的四则运算结果的浮点表示：

$$fl(x \pm y) = (x \pm y)(1 + \varepsilon_{1,2})$$

$$fl(x \cdot y) = (x \cdot y)(1 + \varepsilon_3)$$

$$fl\left(\frac{x}{y}\right) = \left(\frac{x}{y}\right)(1 + \varepsilon_4)$$

其中，$|\varepsilon_i| \le 2^{-t+1}$， $i = 1,2,3,4$ 。

浮点基本运算的误差

二. 连加和连乘的误差

- 设 $x_1, x_2, \cdots, x_n$ 为规格化浮点数，求

  $$fl(x_1 + x_2 + \cdots + x_n) \text{ 和 } fl(x_1 \cdot x_2 \cdot \cdots \cdot x_n)$$

  计算从左向右进行，每次运算都进行截取，再作下一次运算。

- 连和运算而言：
  - 各相对误差限的大小与运算先后有关；
  - 先运算的数，误差也较大；
  - 运算应先安排小数参加运算，可防止“大数吃小数”。
- 连乘运算而言：
  - 各相对误差限的大小与运算先后无关。

数值方法的稳定性

- 设计和选择算法首要关心的问题是精度要求，要建立一些定性分析准则用于判断结果的可靠性，这是数值稳定性问题。
- 对于一个数值方法，若对原始数据或某一步有舍入误差，在执行过程中，这些误差能得到控制，则称该数值方法稳定。否则，称为不稳定的。
- 利用两种数值方法A和B解输入和舍入误差规则相同的同一问题，若用A法比B法得到的计算解精度更高，则称A法比B法具有较大的稳定性。

数值方法的稳定性

- 例：计算积分 $I_n = \frac{1}{e}\int_0^1 x^n e^x dx$， $n = 0,1,2,\cdots,7$

解：由于 $eI_n = x^n e^x |_0^1 - n\int_0^1 x^{n-1} e^x dx = e - neI_{n-1}$

故 $I_n = 1 - nI_{n-1}$ (4.16)

方法一：利用(4.16)式，先计算I_0，然后计算$I_1, \ldots, I_7$。设计算值 I_i^* 的误差为$\varepsilon(I_i^*)$，若 I_0时误差为δ，则

$$\varepsilon(I_1^*) = \delta, \varepsilon(I_2^*) = 2!\delta, \cdots, \varepsilon(I_7^*) = 7!\delta$$

方法二：利用公式 $I_{n-1} = \dfrac{1 - I_n}{n}$ 先计算I_7，再计算 $I_6, I_5, \ldots, I_0$。计算 I_7时产生误差 δ，那么计算 I_0时的误差为 $\varepsilon(I_0^*) = \dfrac{\delta}{7!}$ 。

数值方法的稳定性

- 例:求线性方程组

$$\begin{cases} x_1 + \frac{1}{2}x_2 + \frac{1}{3}x_3 = \frac{11}{6} \\ \frac{1}{2}x_1 + \frac{1}{3}x_2 + \frac{1}{4}x_3 = \frac{13}{12} \\ \frac{1}{3}x_1 + \frac{1}{4}x_2 + \frac{1}{5}x_3 = \frac{47}{60} \end{cases}$$

其准确解为: $x_1 = x_2 = x_3 = 1$

但是，若对数据取3位有效数字，利用Gauss消去法（见第二章）求解，则得到：$x_1 \approx 1.09,\ x_2 \approx 0.488,\ x_3 \approx 0.491$

数值方法的稳定性

- 同样对方程 $\begin{cases} 2x_1 - x_2 + 3x_3 = 1 \\ 4x_1 + 2x_2 + 5x_3 = 4 \\ 2x_1 \qquad\quad + 2x_3 = 6 \end{cases}$

  利用Gauss消去法，取3位有效数字，则得到准确解：

$$x_1 = 9, \quad x_2 = -1, \quad x_3 = -6$$

- 数值问题计算解的精度，不但与数值方法的稳定有关，而且还与数值问题的性态好坏有关。
- 在数值问题中，若输出数据对输入数据的扰动（误差）很敏感（小的变化会引起较大的变化），称这类数值问题为病态问题；否则称为良态问题。

算法设计原则

一. 四则运算中的稳定性问题

1) 防止大数吃小数

$$0.3684676+10^7\times0.6327544+0.4932001+0.4800100$$

$$=10^7\times0.6327544$$

- 预防方法：先加小数，由小到大逐次相加。如

$$0.3684676+0.4800100+0.4932001+10^7\times0.6327544$$

$$=10^7\times0.6327545$$

算法设计原则

一. 四则运算中的稳定性问题

2) 要避免两个相近数相减

相近数相减会严重损失有效数值的位数

$1-\cos(x)$ 可改为 $2\sin^2(x/2)$

3) 避免小数作除数和大数作乘法

这样可避免误差放大，因

$$|E(x_1x_2)| \le |x_1|\cdot|E_2| + |x_2|\cdot|E_1|$$

$$|E(x_1/x_2)| \le \{|x_1|\cdot|E_2| + |x_2|\cdot|E_1|\}/|x_2|^2$$

算法设计原则

二. 提高算法效率问题

1) 尽量减少运算次数：

$$x^{255}=x\cdot x^{2}\cdot x^{4}\cdot x^{8}\cdot x^{16}\cdot x^{32}\cdot x^{64}\cdot x^{128}$$

254次乘法 ⟶ 14次乘法

- 例：计算 $p_n(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$ 的值。
  - 直接计算需要 $n(n+1)/2$ 次乘法和 n 次加法。
  - 若将公式变成递推公式：

$$\begin{cases} s_n=a_n \\ s_k=xs_{k+1}+a_k,\quad k=n-1,n-2,\cdots,1,0 \\ p_n(x)=s_0 \end{cases}$$

后，再计算 $p_n(x)$，只需做 n 次乘法和 n 次加法。

  - 充分利用递推公式，可提高算法效率。

算法设计原则

二. 提高算法效率问题

2) 充分利用耗时少的运算

运算时间：加法/减法 < 乘法 < 除法/平方等

例如：$k+k$ 比 $2k$，$a*a$ 比 a^2，$b*0.25$ 比 $b/4$ 等节省运算时间。

3) 充分利用存贮空间

- 节省原始数据的存贮单元；
- 节省工作单元。

三. 着眼于高质量的软件

- 可移植性，与具体计算机有关的问题
- 易用性，易读性，易维护性

本节课小结

- 介绍了误差分析的基本概念：
  - 舍入误差和截断误差
  - 绝对误差和相对误差
  - 有效数字
- 讨论了浮点数的形式和其带来的基本运算（四则运算，连加连乘）的误差
- 讨论了数值方法的稳定性与算法设计原则

作业

习题一：2，12-(1)(2)(4)，

4【选做，MATLAB编程， 20分】

作业上交日期：2018年9月18日

下节课内容

第二章 解线性方程组的直接法

§1 直接法与三角形方程组的解法

§2 Gauss列主元素消去法