第一章

数值计算的误差

内容提要

- 误差基本概念
  - 误差的来源
  - 绝对误差与相对误差
  - 有效数字
  - 误差估计
- 数值计算的稳定性

误差的来源

误差 是用来描述数值计算中近似解的精确程度，是科学计算中的一个十分重要的概念。

误差的来源

- 从实际问题中抽象出数学模型 —— 模型误差
- 通过测量和实验得到模型中的各种数据 —— 观测误差
- 数学模型的数值求解 —— 截断误差（方法误差）
- 机器字长有限 —— 舍入误差

在数值分析中，主要研究 截断误差 和 舍入误差 对计算结果的影响

误差举例

例：计算 $\int_0^1 e^{-x^2}\mathrm{d}x$ 的近似值。

解：将 e^{-x^2} 作 Taylor 展开，然后积分

$$\int_0^1 e^{-x^2}\mathrm{d}x = \int_0^1 \left(1 - x^2 + \frac{x^4}{2\,!} - \frac{x^6}{3\,!} + \frac{x^8}{4\,!} - \cdots\right)\mathrm{d}x$$

$$= \underbrace{1 - \frac{1}{3} + \frac{1}{2!}\times\frac{1}{5} - \frac{1}{3!}\times\frac{1}{7}}_{S_4} + \underbrace{\frac{1}{4!}\times\frac{1}{9} - \cdots}_{R_4}$$

取 S_4 作为定积分的近似值，即 $\int_0^1 e^{-x^2}\mathrm{d}x \approx S_4$

则 $R_4 = \frac{1}{4\,!}\times\frac{1}{9} - \frac{1}{5\,!}\times\frac{1}{11} + \cdots$ 就是 截断误差

误差举例

$$\int_0^1 e^{-x^2}\,\mathrm{d}x \approx S_4 = 1 - \frac{1}{3} + \frac{1}{10} - \frac{1}{42}$$

保留小数点后 4 位数字

$$\approx 1 - \boxed{0.3333} + 0.1000 - \boxed{0.0238}$$

$$= 0.7429$$

舍入误差

绝对误差

定义：设 x 为精确值，$\tilde{x}$ 为它的一个近似值，则称

$$\epsilon = \tilde{x} - x$$

为近似值 $\tilde{x}$ 的 **绝对误差**，有时简称 **误差**。

记号说明

x — 精确值

$\tilde{x}$ — 近似值

注记

- 在实际应用中，我们所知道的通常是 $\tilde{x}$
- 绝对误差可正可负
- 由于精确值 x 往往是无法获得的，所以绝对误差通常是不可知的

绝对误差限

定义： 存在一个正数 ε，使得，

$$|\tilde{x} - x| \le \varepsilon$$

则称 ε 为 绝对误差限，简称 误差限。

工程上通常记为： $x = \tilde{x} \pm \varepsilon$

例如： $x = 0.5 \pm 0.01$，表示 x 大约为 0.5，误差不超过 0.01

注记

- 通常所说的误差，一般指的是 误差限
- 衡量一个值的精确程度，不仅仅看绝对误差，更重要是看相对误差

相对误差/相对误差限

定义：设 x 为精确值（非零），$\tilde{x}$ 为它的一个近似值，则称

$$\epsilon_r = \frac{\tilde{x} - x}{x}$$

为近似值 $\tilde{x}$ 的 **相对误差**

若存在正数 ε_r，使得 $|\epsilon_r| \le \varepsilon_r$，则称 ε_r 为 **相对误差限**

由于精确值难以获得，相对误差通常也采用下面的定义：

$$\epsilon_r = \frac{\tilde{x} - x}{\tilde{x}}$$

有效数字

例：已知 $\pi = 3.14159265\cdots$

则近似值 $x_1 = 3.14$ 有 3 位有效数字

近似值 $x_2 = 3.1416$ 有 5 位有效数字

我们在计算有效数字的个数时，是从最小的有效数字位开始往前数，直至第一个非零数字为止，如果总共有 n 个数字，那么我们就称有 n 位有效数字。

数学描述

设 $\tilde{x}$ 为 x 的近似值，将 $\tilde{x}$ 表示为

$$\tilde{x} = \pm a_1.a_2 \cdots a_n \cdots \times 10^m$$

其中 a_i 是 0 到 9 中的数字且 $a_1 \neq 0$，且有

$$0.5 \times 10^{k-1} < |\tilde{x} - x| \leq 0.5 \times 10^k$$

则 $\tilde{x}$ 恰好有 $m-k+1$ 位有效数字。

$\tilde{x}$ 有 n 位有效数字 $\Longleftrightarrow$ $0.5 \times 10^{m-n} < |\tilde{x} - x| \leq 0.5 \times 10^{m-n+1}$

推论：若 $|\tilde{x} - x| \leq 0.5 \times 10^{m-n+1}$，则 $\tilde{x}$ 至少有 n 位有效数字。

有效数字举例

例：$\pi = 3.14159265\cdots$，近似值 $x_1 = 3.1415 = 3.1415 \times 10^0$，则

$$x_1 - \pi = -0.92\ldots \times 10^{-4} \Rightarrow 0.5 \times 10^{-4} < |x_1 - \pi| \le 0.5 \times 10^{-3}$$

$\Rightarrow$ 所以 x_1 有 **4** 位有效数字

例：$\pi = 3.14159265\cdots$，近似值 $x_2 = 3.1416 = 3.1416 \times 10^0$，则

$$x_2 - \pi = 0.73\ldots \times 10^{-5} \Rightarrow 0.5 \times 10^{-5} < |x_2 - \pi| \le 0.5 \times 10^{-4}$$

$\Rightarrow$ 所以 x_2 有 **5** 位有效数字

有效数字与相对误差限

定理： 设近似值 $\tilde{x}$ 可表示为

$$\tilde{x} = \pm a_1.a_2 \cdots a_n \cdots \times 10^m \quad (a_1 \neq 0),$$

若 $\tilde{x}$ 具有 n 位有效数字，则其相对误差限满足

$$\left|\frac{\tilde{x}-x}{\tilde{x}}\right| \le \frac{1}{2a_1} \times 10^{-(n-1)}$$

反之，若

$$\left|\frac{\tilde{x}-x}{\tilde{x}}\right| \le \frac{1}{2(a_1+1)} \times 10^{-(n-1)}$$

则 $\tilde{x}$ 至少有 n 位有效数字。

证明：课后阅读

† 直观理解：有效数字越多，相对误差限越小，反之亦然。

如何进行误差估计

误差估计是指计算误差限或相对误差限

- **四则运算**的误差估计

$$\varepsilon(\tilde{x}_1 \pm \tilde{x}_2) \le \varepsilon(\tilde{x}_1) + \varepsilon(\tilde{x}_2)$$

这里用 $\varepsilon(\tilde{x})$ 表示 $\tilde{x}$ 的误差限

$$\varepsilon(\tilde{x}_1\tilde{x}_2) \le |\tilde{x}_2|\varepsilon(\tilde{x}_1) + |\tilde{x}_1|\varepsilon(\tilde{x}_2) + \varepsilon(\tilde{x}_1)\varepsilon(\tilde{x}_2)$$

$$\approx |\tilde{x}_2|\varepsilon(\tilde{x}_1) + |\tilde{x}_1|\varepsilon(\tilde{x}_2)$$

$$\varepsilon\left(\frac{\tilde{x}_1}{\tilde{x}_2}\right) \le \frac{|\tilde{x}_2|\varepsilon(\tilde{x}_1) + |\tilde{x}_1|\varepsilon(\tilde{x}_2)}{|\tilde{x}_2|^2} + \cdots \approx \frac{|\tilde{x}_2|\varepsilon(\tilde{x}_1) + |\tilde{x}_1|\varepsilon(\tilde{x}_2)}{|\tilde{x}_2|^2}$$

一元函数误差估计

一元可微函数 $f(x)$ 求值

设 $\tilde{x}$ 为 x 的近似值，则有

$$f(x)-f(\tilde{x})=f'(\tilde{x})(x-\tilde{x})+\frac{f''(\xi)}{2}(x-\tilde{x})^2$$

假定用 $\tilde{x}$ 非常接近 x，且 $|f''(\tilde{x})|$ 不是很大

$$\Rightarrow \quad |f(\tilde{x})-f(x)| \approx |f'(\tilde{x})| \times |\tilde{x}-x|$$

$$\Rightarrow \quad \boxed{\varepsilon\big(f(\tilde{x})\big) \approx |f'(\tilde{x})|\,\varepsilon(\tilde{x})}$$

多元函数误差估计

- **多元可微函数 $f(x)$ 求值**，其中 $x=(x_1, x_2, \ldots, x_n)$

  设 $\tilde{x}=(\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n)$ 为 x 的近似值，则

$$\boxed{\varepsilon\big(f(\tilde{x})\big) \approx \sum_{k=1}^{n} \left| \frac{\partial f(\tilde{x})}{\partial x_k} \right| \varepsilon(\tilde{x}_k)}$$

误差估计

例： 测得某场地的长 x 和宽 y 分别为：$\tilde{x}=110\text{m}, \tilde{y}=80\text{m}$，测量误差限分别为 0.2m 和 0.1m。试求面积 S 的误差限和相对误差限。

解：板书

解. 由于 $\varepsilon(\tilde{x})=0.2\text{m}, \varepsilon(\tilde{y})=0.1\text{m}$, 故

$$\begin{aligned}\varepsilon(\tilde{S}) &\approx \left|\frac{\partial S(\tilde{x},\tilde{y})}{\partial x}\right|\varepsilon(\tilde{x})+\left|\frac{\partial S(\tilde{x},\tilde{y})}{\partial y}\right|\varepsilon(\tilde{y}) \\ &= |\tilde{y}|\cdot\varepsilon(\tilde{x})+|\tilde{x}|\cdot\varepsilon(\tilde{y}) \\ &= 80\times 0.2+110\times 0.1=27(\text{m}^2).\end{aligned}$$

相对误差限

$$\varepsilon_r(\tilde{S})=\frac{\varepsilon(\tilde{S})}{|\tilde{S}|}\approx\frac{27}{110\times 80}\approx 0.0031.$$

误差分析

误差分析

- 数值计算中的误差分析很重要，但也很复杂
- 在计算过程中，误差会 传播、积累、对消
- 误差分析的核心是研究原始数据的误差和计算中产生的误差对最终计算结果的影响

† 算法有“优劣”之分，问题有“好坏”之别，即使不能定量地估计出最终误差，但是若能判别计算过程中误差不会被任意放大，那就能放心地实施计算，这就是误差分析的初衷。

内容提要

- 数值计算的稳定性
  - 算法的稳定性
  - 病态问题与条件数
  - 数值计算几点注意事项

算法的稳定性

例：近似计算 $S_n = \int_0^1 \frac{x^n}{x+5}\mathrm{d}x$ ，其中 n=1, 2, ..., 8

解：$S_n + 5S_{n-1} = \int_0^1 \frac{x^n + 5x^{n-1}}{x+5}\mathrm{d}x = \int_0^1 x^{n-1}\mathrm{d}x = \frac{1}{n}$

$$\Rightarrow \boxed{S_n = \frac{1}{n} - 5S_{n-1}}$$

该公式精确成立

直接计算可得： $S_0 = \ln 6 - \ln 5 \approx 0.182$

保留 3 位有效数字

通过递推公式可得（每次都保留 3 位有效数字）

$$\tilde{S}_1 = 0.090,\quad \tilde{S}_2 = 0.050,\quad \tilde{S}_3 = 0.0833,\quad \tilde{S}_4 \overset{?}{=} -0.166,$$
$$\tilde{S}_5 \overset{?}{=} 1.03,\quad \tilde{S}_6 \overset{?}{=} -4.98,\quad \tilde{S}_7 \overset{?}{=} 25.0,\quad \tilde{S}_8 \overset{?}{=} -125.$$

$$\frac{1}{6(n+1)} = \int_0^1 \frac{x^n}{6}dx \le \int_0^1 \frac{x^n}{x+5}dx \le \int_0^1 \frac{x^n}{5}dx = \frac{1}{5(n+1)}$$

算法的稳定性

考察第 n 步的误差：

$$\tilde{S}_n - S_n \approx \left(\frac{1}{n} - 5\tilde{S}_{n-1}\right) - \left(\frac{1}{n} - 5S_{n-1}\right) = -5(\tilde{S}_{n-1} - S_{n-1})$$

即有

$$\boxed{\left|\tilde{S}_n - S_n\right| \approx 5\left|\tilde{S}_{n-1} - S_{n-1}\right| \approx \cdots \approx 5^n \left|\tilde{S}_0 - S_0\right|}$$

误差以 5 倍 的速度增长！

这表明计算过程中误差得不到有效控制，

计算过程（算法）是不稳定的，需要更换新算法！

算法的稳定性

解法二： $S_n = \dfrac{1}{n} - 5S_{n-1}$ $\Rightarrow$ $\boxed{S_{n-1} = \dfrac{1}{5n} - \dfrac{1}{5}S_n}$

新思路：先计算 S_8，再计算 S_7，S_6，. . .，S_0

$$\left|\tilde{S}_n - S_n\right| \approx 5\left|\tilde{S}_{n-1} - S_{n-1}\right|$$

$$S_8 \approx \frac{1}{2}\left(\frac{1}{5(8+1)} + \frac{1}{6(8+1)}\right) \approx 0.0204 \quad \leftarrow \quad \frac{1}{6(n+1)} \le S_n \le \frac{1}{5(n+1)}$$

$$\tilde{S}_8 = 0.0204,\quad \tilde{S}_7 = 0.0209,\quad \tilde{S}_6 = 0.0244,\quad \tilde{S}_5 = 0.0285,$$
$$\tilde{S}_4 = 0.0343,\quad \tilde{S}_3 = 0.0431,\quad \tilde{S}_2 = 0.0580,\quad \tilde{S}_1 = 0.0884.$$

ex11.m

可以验证，计算结果都具有 3 位有效数字。

也可以考虑先计算更大的 S_n，如 S_{10}，以获得更高精度。

† 数值计算中，误差不可避免，算法的稳定性是一个非常重要的性质。

算法的稳定性

算法的稳定性

在计算过程中，如果误差不增长或能得到有效控制，则称该算法是稳定的，否则为不稳定的。

在数值计算中，不要采用不稳定的算法！

问题的性态

如果输入数据的微小扰动会引起计算结果的很大变化，则称该数学问题是 **病态** 的，否则就是 **良态** 的。

病态问题举例

例：解线性方程组 $\begin{cases} x+\alpha y=1 \\ \alpha x+y=0 \end{cases}$

解：当 $\alpha=1$ 时，无解

当 $\alpha\neq 1$ 时，解为：$x=\dfrac{1}{1-\alpha^2}$，$y=\dfrac{-\alpha}{1-\alpha^2}$

若 $\alpha\approx 1$，则 α 的微小误差可能会引起解的很大变化

比如 $\alpha=0.9990$ 时，$x\approx 500.25$;

如果输入的数据有 0.0001 的误差，即 $\tilde{\alpha}=0.9991$，则 $\tilde{x}\approx 555.81$，误差约为 55.56！

因此，当 $\alpha\approx 1$ 时，该线性方程组就是病态的！

思考：当 α 远离 1 呢？

怎么判断一个问题是否病态

输入数据的微小扰动是否会引起计算结果的很大变化

考虑一元函数情形

设一元函数 $f(x)$ 可微，$\tilde{x}$ 为 x 的近似值，则有

$$f(\tilde{x}) - f(x) = f'(x)(\tilde{x} - x) + \frac{f''(\xi)}{2}(\tilde{x} - x)^2$$

$$\Rightarrow \quad \varepsilon_r\left(f(\tilde{x})\right) = \left|\frac{f(\tilde{x}) - f(x)}{f(x)}\right| \approx \underbrace{\left|\frac{xf'(x)}{f(x)}\right|}_{C_p} \times \underbrace{\left|\frac{\tilde{x} - x}{x}\right|}_{\varepsilon_r(\tilde{x})}$$

C_p 越大，计算结果的误差就可能会越大。

我们称 C_p 就称为 $f(x)$ 的**条件数**，它是判断 f 是否病态的一个重要依据。

几点注记

- 病态是问题本身固有的性质，与数值算法无关
- 对于病态问题，选择数值算法时需要更加谨慎

- 线性方程组的条件组将在后面的章节中介绍。

数值计算注意事项

为了尽可能地减小各种误差对最终计算结果的危害，提高计算结果的可靠性，在数值计算过程中，下面几点建议须牢记！

(1) 避免两个相近的数相减

(2) 避免数量级相差很大的数相除

(3) 避免大数吃小数

(4) 简化计算，减少误差积累

数值计算注意事项

(1) 避免两个相近的数相减，否则会损失有效数字

例：$a_1 = 0.12345$，$a_2 = 0.12346$，都有 5 位有效数字。

而 $a_2 - a_1 = 0.00001$，只剩下 1 位有效数字。

数值计算注意事项

例：计算 $y=\left(\dfrac{\sqrt{101}-10}{\sqrt{101}+10}\right)^2$

我们分别用三种方法计算：

方法一 分母有理化 $y=\left(\dfrac{\sqrt{101}-10}{\sqrt{101}+10}\right)^2=80801-8040\sqrt{101}$

方法二 分子有理化 $y=\left(\dfrac{\sqrt{101}-10}{\sqrt{101}+10}\right)^2=\dfrac{1}{80801+8040\sqrt{101}}$

方法三 直接代入计算

分别取近似值 $\sqrt{101}\approx 10.04$ 和 $\sqrt{101}\approx 10.06$ 进行计算（3 位有效数字）

$\left(\sqrt{101}=10.0498756\ldots\right)$

数值计算注意事项

计算结果如下：

y=6.188042227…e-6

	发法一	方法二	方法三
$\sqrt{101}\approx 10.04$	79.4	6.1911e-06	3.9841e-06
$\sqrt{101}\approx 10.06$	-81.4	6.1849e-06	8.9462e-06

法一和法三的结果都不对，

法二却都具有 3 位有效数字！

数值计算注意事项

几种经验性避相近的数相减的方法

基本思路：通过恒等变换，避免相近的数相减

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

$$\ln(x+\varepsilon)-\ln x=\ln\left(1+\frac{\varepsilon}{x}\right)$$

$$1-\cos x=2\sin^2\frac{x}{2} \quad (x\to 0)$$

$$e^x-1=x\left(1+\frac{1}{2}x+\frac{1}{6}x^2+\dots\right) \quad (x\to 0)$$

举例

例：计算 $y=\ln 2$

方法一 $\ln(1+x)=x-\frac{1}{2}x^2+\frac{1}{3}x^3-\frac{1}{4}x^4+\frac{1}{5}x^5+\cdots+\frac{(-1)^{n+1}}{n}x^n+\cdots$

将 $x=1$ 代入，计算结果如下：

n	1	2	3	4	5	10	20	50	100
S_n	1	0.5	0.833	0.583	0.783	0.646	0.669	0.683	0.688
$\lvert S_n-\ln 2\rvert$	3.1E-01	1.9E-01	1.4E-01	1.1E-01	9.0E-02	4.8E-02	2.4E-02	9.9E-03	5.0E-03

计算前 **100** 项后，误差仍有 **0.05**

举例

方法二 $$\ln\left(\frac{1+x}{1-x}\right)=2\left(x+\frac{1}{3}x^3+\frac{1}{5}x^5+\cdots+\frac{1}{2n-1}x^{2n-1}+\cdots\right)$$

将 $x=1/3$ 代入，计算结果如下：

n	1	2	3	4	5	10
S_n	0.667	0.691	0.693	0.693	0.693	0.693
$\lvert S_n-\ln 2\rvert$	2.6E-02	1.8E-03	1.4E-04	1.2E-05	1.1E-06	1.0E-11

ex12.m

计算前 **10** 项后，误差已经小于 10^{-10}！

算法的选取很重要！

数值计算注意事项

(2) 避免数量级相差很大的数相除

避免可能会产生的溢出情形（超出计算机所能表示的范围）

避免放大误差（很小的数做分母）

为了尽可能地保证计算过程的稳定性，通常两个数相除时，最好使得分母的绝对值不小于分子的绝对值

数值计算注意事项

(3) 避免大数吃小数

例：计算 $\dfrac{10^9+10^{-9}-10^9}{10^{-9}}$

数值计算结果为 0 而不是1

例：分别按从小到大、以及从大到小的顺序分别计算

$$S=1+2+3+\ldots+100+10^{16}$$

ex13.m

求和时 **从小到大** 相加，可使结果更精确

数值计算注意事项

(4) 简化计算，减少误差积累

例：已知 $\boldsymbol{p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n}$，计算 $\boldsymbol{p(2)}$ 的值。

秦九韶算法（西方也称为 Horner 算法）

$$p(x) = a_0 + x(a_1 + x(a_2 + x(a_3 + x(\cdots x(a_{n-1} + a_n x)\cdots))))$$

直接计算，需要 $\frac{n(n+1)}{2}$ 次乘法和 $\boldsymbol{n}$ 次加法；用秦九韶算法则只需 $\boldsymbol{n}$ 次乘法和 $\boldsymbol{n}$ 次加法，运算量大大减小！