数值计算方法

主讲：纪庆革副教授

中山大学数据科学与计算机学院

E-Mail: 1024180018@qq.com

前期课程：

高等数学

高等代数

计算机应用基础

Matlab, C, C++等

绪　论

- 计算方法（数值分析）的研究**对象**
- 计算方法（数值分析）的**特点**
- 学习的目的、要求
- 浮点数》》
- 误差
- 误差传播
- 设计算法的注意事项
- 例子

计算方法的研究对象

实际问题　　具体科学、工程问题的解决

↓

数学模型　　介绍最常用数学模型的最基本数值计算方法（随着计算机的飞速发展，数值计算方法已深入到计算物理、计算力学、计算化学、计算生物学、计算经济学等各个领域。

↓

数值方法　　**研究并解决数学问题的数值近似解方法**（连接模型到结果的重要环节）

↓

程序设计

↓

上机计算

计算方法的特点

- 面向计算机（加、减、乘、除和逻辑运算）
- 有可靠的理论分析(逼近、精度、收敛性、数值稳定性、误差分析)
- 要有好的计算复杂性
- 要有数值实验

学习的目的、要求

- 会套用、修改、创建公式
- 部分定理证明
- 编制程序，完成计算

课程评分方法

总分 (100) ＝ 平时考核（含平时作业、单元测试和课堂提问）（40） +期末考试 (60)

◆上机作业要求

1、 建议使用C++或Matlab编程；

注：不允许使用内置函数完成主要功能

2、以E-Mail形式交实验报告：

主题： 班级+姓名（小组）+学号+第几次作业

文件名:班级+姓名（小组）+学号+第几次作业.rar

内容:附件是压缩文件，含实验报告（附上运行结果）和源代码。

留作业&实验题目

QQ群：1059065332

交作业&实验报告

Email:

内容

1、数值逼近与数值微积分——数学分析中的数值求解，如数值微分、数值积分

$$\int_a^b f(x)dx = F(b) - F(a)$$

2、数值代数——线性代数的数值求解，如解线性方程组等

$$Ax = b \quad \Rightarrow \quad x_i = D_i / D \; n = 20, 9.7 \times 10^{20}$$

100亿/秒，算3,000年，而Gauss消元法2660次

3、微分方程—常微分，如**Runge-Kutta**法等

浮点数（1/7）（》》）

- 定点数

设**r**为大于1的正整数，**a_i**为0, 1, … , **r**-1中的某一个，位数有限的**r**进制正数可以写为$x \overset{\Delta}{=} a_{l-1}a_{l-2}\cdots a_0.a_{-1}a_{-2}\cdots a_{-m}$

x有**l**位整数，有**m**位小数。因为进制基数是**r**，故

$$x = a_{l-1}r^{l-1} + a_{l-2}r^{l-2} + \cdots + a_0 r^0 + a_{-1}r^{-1} + \cdots + a_{-m}r^{-m}$$

浮点数(2/7)

当*l*=4, *m*=4, *r*=10时，

109.312，0.4375，4236

分别表示为 0109.3120，0000.4375，4236.0000

定义：这种把小数点永远固定在指定位置上位数有限的数称为定点数，*n*=*l*+*m*称为字长。

当*l*=*m*=4, *r*=10时，8位定点非零数中绝对值最小和最大的数分别为 0000.0001，9999.9999

注：定点数表示的数的范围较小

浮点数(3/7)

定点数运算系统中的溢出情况：

(1) 表示范围的溢出；

(2) 运算结果的溢出。

$l=0$时，0.5+0.6=0.1,产生上溢出。

在16位二进制系统的计算机上计算 $\omega=\dfrac{2^{-7}\times 2^{-9}}{2^{-10}}$

利用算法 $\omega=\dfrac{(2^{-7}\times 2^{-9})}{2^{-10}}=\dfrac{0}{2^{-10}}=0$ 产生下溢出，

而利用算法 $\omega=\left(\dfrac{2^{-7}}{2^{-5}}\right)\times\left(\dfrac{2^{-9}}{2^{-5}}\right)=2^{-2}\times 2^{-4}=2^{-6}$

注：在编制定点运算程序时，要避免运算结果的上下溢出，要慎重选择计算次序。

浮点数(4/7)

- 浮点数

特点：用浮点方式表示的数有比较大的取值范围，且浮点运算有较高的计算精度。

定义：设 **s** 是 **r** 进制数，**p** 是 **r**进制正负整数或零， **r** 进制数 **x** 可以用**s** 和 **r^p** 的乘积表示为

$$x = s \times r^{p} \qquad (1.2.3)$$

再设**s**的整数部分等于零，即**s**满足条件

$$-1 < s < 1 \qquad (1.2.4)$$

则形如**(1.2.3)**而满足条件**(1.2.4)**的**r**进制数**x**称为**r**进制浮点数。**s**和**p**分别称为尾数和阶数。

如果尾数的小数位数等于有限正整数 **t**，则把 **x** 称为**t**位浮点数。

浮点数(5/7)

如果要求尾数s小数点后第一位数字不等于零，即：

$$r^{-1} \le s < 1 \qquad (1.2.5)$$

则形如（1. 2. 3）而满足条件（1. 2. 5）的浮点数称为r进制规格化浮点数。

如十进制数：

0.003012, 0.3217, 283.4

其规格化浮点数分别为：

$$0.3012 \times 10^{-2}, \quad 0.3217 \times 10^{0}, \quad 0.2834 \times 10^{3}$$

二进制数：**1001.101, 0.10101, 0.00101**

其规格化浮点数分别为：

$$\mathbf{0.1001101\times2^{4},\ 0.10101\times2^{0},\ 0.101\times2^{-2}}$$

注：只要数$x\neq0$，则x一定可以表示为规格化浮点数。

这样做有什么好处？（好处：一个数的数量级一目了然）

浮点数(6/7)

- **计算机数系**

定义： 进位制为 r，*阶数* p 满足条件

$$l \le p \le u \qquad (1.2.6)$$

其中 *l, u* 为整数（由计算机用多少位数表示阶数所确定）

如果尾数的小数位数为t，则计算机数系由一切满足(1.2.6)的t位r进制浮点数的集合F组成。

F中的浮点数具有具有以下形式

$$x = \pm\left(\frac{d_1}{r} + \frac{d_2}{r^2} + \cdots + \frac{d_t}{r^t}\right) \cdot r^p \qquad (1.2.7)$$

$$\overset{\Delta}{=} \pm 0.d_1 d_2 \cdots d_t \times r^p$$

浮点数(7/7)

其中 $d_1, d_2, \cdots, d_t$ 为整数，满足如下关系

$$0 \le d_i \le r-1, i=1,2,\cdots, \qquad (1.2.8)$$

若对$x\neq0$,规定(1.2.7)中$d_1\neq0$,则F为规格化的浮点数。

F 中的浮点个数：

$$2(r-1)r^{t-1}(u-l+1)+1 \qquad (1.2.9)$$

最小数和最大数：

-0.9999×10^{99}，0.9999×10^{99}

当$r=10$，$t=4$，$l=-99$，$u=99$时，

-0.0001×10^{-99}，0.0001×10^{-99}是数系F中绝对值最小的非零数。

误差(1/3)

简记为e^*

- 绝对误差

设x为精确值，x^*为近似值，$e(x^*)\overset{\Delta}{=}x^*-x$为绝对误差，简称误差

例如：$f(x)=\ln(x+1)$ 作**Taylor**展开，

舍弃，即为误差

$$=\sum_{i=1}^{n}\frac{(-1)^{i-1}}{i}x^i+\frac{(-1)^n x^{n+1}}{(n+1)(1+\theta x)^{n+1}}\ ,\ 0<\theta<1$$

如果存在正数$\varepsilon^*=\varepsilon(x^*)$，使得绝对误差$\left|e^*\right|=\left|x^*-x\right|\le\varepsilon^*$，

则称**ε***为**x***近似**x**的一个绝对误差限，简称 误差限。

此时有$x\in\left[x^*-\varepsilon^*,x^*+\varepsilon^*\right]$，可用 $x=x^*\pm\varepsilon^*$ 表示。

误差(2/3)

强近似值：若$e(x^*)$为正时，近似值x^*偏大，叫做强近似值。

弱近似值：若$e(x^*)$为负时，近似值x^*偏小，叫做弱近似值。

误差(3/3)

- 相对误差

$$e_r\left(x^*\right)=\frac{x^*-x}{x}$$ 称为x*近似x的相对误差

例如：150分满考139，100分满考90，两者的绝对误差分别为11和10，优劣如何？

前者相对误差(150－139)/150=0.073，

后者相对误差(100-90)/100=0.100

称 $\left|e_r^*\right|$ 的上界为相对误差限。

误差来源

- 原始误差—模型误差（忽略次要因素，如空气阻力）物理模型，数学模型
- 观测误差—观测产生的误差
- 方法误差—截断误差（近似算法本身引起）
- 计算误差—舍入误差（计算机表示数据引起）

误差的运算(1/3)

1、 $(x^* \pm y^*) - (x \pm y) = e_x \pm e_y$

$$\frac{e_x \pm e_y}{x^* \pm y^*}$$

两相近数相减，相对误差增大

2、 $(x^* \cdot y^*) - (x \cdot y) = x^*(y^* - y) + y(x^* - x)$

$$= ye_x + x^* e_y$$

$$= \max\{|x^*|, |y|\}(|e_x| + |e_y|)$$

误差的运算(2/3)

3、$\left|\frac{x^*}{y^*}-\frac{x}{y}\right| = \left|\frac{x^* y - y^* x}{yy^*}\right|$

$$= \left|\frac{-x^*(y^*-y)+y^*(x^*-x)}{yy^*}\right|$$

$$= \left|\frac{-x^* e_y + y^* e_x}{yy^*}\right|$$

小数作除数，绝对误差增大

误差的运算(3/3)

例子

求根 $x^2 - 2000x + 1 = 0$

$$x_{1,2} = \frac{2000 \pm \sqrt{2000^2 - 4}}{2} \Rightarrow x_2 = 0.0005$$

$$\begin{cases} x_1 = \dfrac{2000 + \sqrt{2000^2 - 4}}{2} \\ x_2 = \dfrac{1}{x_1} \end{cases} \Rightarrow x_2 = 0.00050005$$

另：参照 例2 P11

有效数字(有效位数)(1/6)

- 当x的误差限为某一位的半个单位，则这一位到第一个非零位的位数称为x的有效数字。

例如，$x=\pi=3.14159265\cdots$, 取 $x^*=3.14$ 时，

$$\left|x^*-x\right|\le 0.002\le 0.005$$

$x^*=3.14$ 作为π的近似值时，就有3位有效数字。

有效位的多少直接影响到近似值的绝对误差和相对误差

有效数字(2/6)

在r进制中，设近似值x^*可表示为

$$x^* = \pm\left(a_1 r^{-1} + a_2 r^{-2} + \cdots + a_n r^{-n}\right) \times r^m \qquad (1.3.11)$$

$a_1 \neq 0,$ 且

$$\left|x^* - x\right| \leq \frac{1}{2} r^{m-n}$$

则x^*有n位有效数字。

有效数字(3/6)

例1 按四舍五入原则，写出下列各数具有5位有效数字的近似数

187.9325， 0.03785551， 8.000033， 2.7182818

解：187.93，0.037856，8.0000，2.7183

有效数字(4/6)

有效位数与小数点的位置无关，具有***n***位有效数字的近似数***x****其误差限为

$$\varepsilon\left(x^*\right)=\frac{1}{2}\times r^{m-n}$$

在***m***相同的条件下，有效位数越多，则绝对误差限越小。

有效数字(5/6)

定理1 若用(1.3.11)式表示的近似数x^*具有n位有效数字，则其相对误差限为

$$\left|e_r\left(x^*\right)\right| \le \frac{1}{2a_1} \times r^{-(n-1)}$$

定理2 若用(1. 3. 11)式表示的近似数 x^* 满足

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

则x^*至少有n位有效数字。

有效数字 (6/6)

- 例：为使 π^* 的相对误差小于0.001%, 至少应取几位有效数字？

解：假设 π^* 取到 n 位有效数字，则其相对误差上限为 $\varepsilon_r^* \le \frac{1}{2a_1} \times 10^{-n+1}$

要保证其相对误差小于0.001%，只要保证其上限满足 $\varepsilon_r^* \le \frac{1}{2a_1} \times 10^{-n+1} < 0.001\%$

已知 $a_1 = 3$，则从以上不等式可解得 $n > 6 - \log 6$，即 $n \ge 6$，应取 $\pi^* = 3.14159$。

函数的误差估计(1/7)

对于 $y = f(x)$，若用 x^* 取代 x，将对y 产生什么影响？

$$e^*(y) = f(x^*) - f(x)$$

$$= f'(\xi)(x^* - x) = f'(\xi)\, e^*(x)$$

x^* 与 x 非常接近时，可认为 $f'(\xi) \approx f'(x^*)$，则有：

$$|e^*(y)| \approx |f'(x^*)| \cdot |e^*(x)|$$

结论：x^* 产生的误差经过 f 作用后被 放大 / 缩小了 $|f'(x^*)|$ 倍。故称$|f'(x^*)|$为 放大因子 或 绝对条件数

函数的误差估计(2/7)

$$|e_r^*(y)| = \left|\frac{e^*(y)}{f(x^*)}\right| \qquad |e_r^*(x)| = \left|\frac{e^*(x)}{x^*}\right|$$

$$= \left|\frac{f(x^*) - f(x)}{x^* - x} \cdot \frac{x^*}{f(x^*)} \cdot \frac{x^* - x}{x^*}\right|$$

$$\approx \left|\frac{x^* \cdot f'(x^*)}{f(x^*)}\right| \cdot |e_r^*(x)|$$

相对误差条件数

f 的条件数在某一点是小\大，则称 f 在该点是好条件的\坏条件的

函数的误差估计(3/7)

例:计算 $y=\ln x$。若 $x\approx 20$，则取 x 的几位有效数字可保证 y 的相对误差 $<0.1\%$?

解：设截取 n 位有效数字后得 $x^*\approx x$，则

$$|e_r^*(y)|\approx\left|\frac{x^*y'(x^*)}{y(x^*)}\right|\cdot|e_r^*(x)|=\frac{|e_r^*(x)|}{\ln x^*}$$

估计 x 和 y 的相对误差上限满足近似关系

$$\varepsilon_r^*(x)\approx\ln x^*\cdot\varepsilon_r^*(y)$$

$$\Rightarrow\quad \frac{1}{2a_1}\times 10^{-n+1}<\ln x^*\cdot 0.1\%$$

$$\Rightarrow\quad n\geq 4$$

例：计算 $\ln\left(20\frac{8}{9}\right)$，取 4 位有效，即 $\ln(20.89)$, 则相对误差

$$\left|\frac{\ln(20.89)-\ln\left(20\frac{8}{9}\right)}{\ln\left(20\frac{8}{9}\right)}\right|<2.0\times10^{-5}<0.1\%$$

函数的误差估计(4 / 7)

例子：

1、

$$I_n=\int_0^1\frac{x^n}{x+5}dx$$，则我们有

$$I_n+5I_{n-1}=\int_0^1\frac{x^n+5x^{n-1}}{x+5}dx=\int_0^1 x^{n-1}dx=\frac{1}{n}$$

构造方法：1) $I_n=\frac{1}{n}-5I_{n-1}$ ，$I_0=\ln\frac{6}{5}$ $\qquad \tilde{I}_n$

2) $I_{n-1}=\frac{1}{5}\left(\frac{1}{n}-I_n\right)$ ，$I_8=0.019$ $\qquad \bar{I}_n$

事先估计值

函数的误差估计(5 / 7)

n	I_n	$\tilde{I}_n$	$\bar{I}_n$
0	0.182	0.182	0.182
1	0.088	0.090	0.088
2	0.058	0.050	0.058
3	0.0431	0.083	0.0431
4	0.0343	-0.165	0.0343
5	0.0284	1.025	0.0284
6	0.024	-4.958	0.024
7	0.021	24.933	0.021
8	0.019	-124.540	0.019

函数的误差估计(6 / 7)

原因：对格式**1**，如果前一步有误差，
则被放大**5**倍加到这一步

称为不稳定格式

稳定格式，对舍入误差有抑制作用

函数的误差估计(7 / 7)

2、有时候，模型本身就是病态的
（系数引入小变化，解产生大变化）

$$\begin{cases} x + ay = 1 \\ ax + y = 0 \end{cases}$$

$a = 0.99$ $x = 50.25$

$a = 0.991$ $x = 55.81$

设计算法的注意事项

- 应尽量减少运算次数（为了防止舍入误差的大量积累）
- 多个数相加最好应按照绝对值从小到大的顺序执行（为了防止大数吃掉小数）
- 要避免两相近的数相减(避免造成有效数字的严重损失)
- 应避免除数绝对值远远小于被除数绝对值的除法运算(为了防止舍入误差)
- 应选择良性的数学模型(从源头上避免对算法稳定性造成大的影响)

补充材料1/2

- ***Reading:***
  - Page 10 – Horner’s method
  - Page 25 – significant digits（有效数字）
  - Page 31 – big Oh

补充材料2/2

- ***Home assignment:***
  - Page 13 – exercise 14
  - Page 37 – exercises 1, 3, 9

部分参考书目1/2

1. 同济大学计算数学教研室编著，现代数值计算，人民邮电出版社，2014.9第二版，2018.9 河北第七次印刷

2. 黄云清等编著，数值计算方法，科学出版社，2009.1第一版，2017.1第六次印刷

3. 魏毅强等，数值计算方法，科学出版社，北京，2006

4. 喻文健编著，数值分析与算法，清华大学出版社，2012年

部分参考书目2/2

5. 李庆扬，王能超，易大义编，数值分析（第5版），清华大学出版社，2008年出版，2014年第11次印刷

6. Pallab Ghosh著，徐士良等译，数值方法，清华大学出版社，2008

7. John H. Mathews, Kurtis D. Fink, Numerical Methods Using MATLAB, Fourth Edition，电子工业出版社，2004，2010改编版

本课件是在中国科技大学张瑞老师课件基础上修改而成的，在此表示感谢！