Numerical Analysis

数值分析

同学们好！

任课教师：　马宁

办公室　　：　综合楼A座820

联系方式：　bjmaning@163.com

基本要求

♦两周交一次作业。

♦掌握Matlab语言，实现算法的程序应用。

考试评分

总成绩=平时成绩*30%+期末成绩*70%。

教材

张明主编，应用数值分析（第四版）,石油工业出版社,2013

课程相关资料可在网络课程平台获取

第一章 绪论

第一节 数值分析的研究对象和特点

数值分析实际上就是介绍在计算机上解决数学问题的数值计算方法及其理论。这门课程又称为数值计算方法.

概率积分 $$\frac{2}{\sqrt{\pi}}\int_0^t e^{-x^2}dx \qquad t \in [0,+\infty)$$

求数学问题的数值解称为数值问题.

数值分析的特点

计算机
- 硬件　功能 算术与逻辑运算
- 软件　核心算法
  - 数值算法
  - 非数值算法

计算机硬件的特点是快，软件就是利用计算机高速的简单运算去实现各种复杂的功能。

现代科学的三个组成部分:

科学理论,科学实验,科学计算

科学计算的核心内容是以现代化的计算机及数学软件为工具，以数学模型为基础进行模拟研究。

计算数学,计算物理学,计算力学,计算化学,计算生物学,计算地质学,计算经济学,等等

科学计算的步骤:实际问题→数学模型→数值方法→程序设计→上机计算→分析结果。

1、建立数学模型（实际问题数学化）

2、设计计算方案（数学问题数值化）

$$\begin{cases} \dfrac{dy}{dx} = f(x, y) & x > x_0 \\ y(x_0) = y_0 \end{cases}$$

3、程序设计（数值问题机器化）

4、上机计算并分析结果

数值分析是计算数学的基础课，主要内容包括

数值代数：线性代数方程组的数值解法，求矩阵特征值的数值解法，非线性方程组的数值解法.

数值逼近：函数插值，函数逼近，数值积分，数值微分，常微分方程的数值解法.

数值分析的特点：

1、要根据计算机的特点设计有效算法。即算法只能包括加、减、乘、除运算和逻辑运算。因此“归纳”成了不容忽视的思维方法。

2、要有可靠的理论分析。即近似解能任意达到精度要求，近似算法要保证收敛性和稳定性。因此讨论的核心问题是“误差”。

3、要有好的计算复杂性。计算复杂性包括时间复杂性和空间复杂性。时间复杂性好是指节省时间，空间复杂性好是指节省存储量。因此这是在算法设计和程序设计中要研究的问题。

4、要有数值实验。即任何一个算法都要通过数值试验证明是行之有效的。

学习重点：

1. 构造数值方法的原理（支撑理论）

迭代法,以直代曲,化整为零,外推法

2. 评价数值方法的好坏

(研究数值方法的性态、可靠性、效率)

3. 数值方法的计算机实现（计算机实习）

要掌握高级编程语言：

Matlab，FORTRAN，C

Matlab几个显著特点

1　用Matlab处理矩阵——容易

2　用Matlab绘图——轻松

3　用Matlab编程——简洁

4　Matlab具有丰富的工具箱

本课程的基本目的，是使学生通过学习和实验，初步建立并理解数值计算，特别是科学与工程计算的基本概念，为进一步深入的学习打下坚实的基础。

第二节 数值问题与数值算法

算法： 从给定的已知量出发，经过有限次四则运算及规定的运算顺序，最后求出未知量的数值解，这样构成的完整计算步骤称为**算法**。

评价算法的两个主要标准：**速度和精度**

一个面向计算机,计算复杂性好,又有可靠理论分析的算法就是一个好算法.

计算复杂性包括时间复杂性和空间复杂性

时间复杂性即计算量：一个算法所需四则运算总次数.

一个算法所需的乘除运算总次数，单位是flop.

空间复杂性即存储量

例1 计算 x^{255}

A: $x^{255} = x\ x \overbrace{\cdots x}^{254}$

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

算法*B*(*Matlab*) （输入x，输出y）

$s = x;$

$y = x;$

$\text{for } i = 1:7$

$\quad s = s * s;$

$\quad y = y * s;$

end

计算量

$N = 14\ flop$

存储量=4

例2: 计算多项式$p(x)=3x^3+4x^2-2x+6$的值。

算法1: 由x计算出x^2, x^3后再算。

需乘法6次，加法3次，存储单元7个。

算法2: $p(x)=x[x(3x+4)-2]+6$

需乘法3次，加法3次，存储单元6个。

一般地，计算n次多项式的值

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

算法1、需乘法2n次，加法n次，存储单元n+4个。

算法2、秦九韶算法1247（又称为Horner算法1819)

$$P_n(x)=x(x(x\cdots(x(a_nx+a_{n-1})+a_{n-2}+\cdots+a_1)+a_0$$

有递推公式

需乘法n次，加法n次，存储单元n+3个。

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

算法1 (输入$a(i)(i=0,1，\dots,n)$,x;输出y)

注意

$$
\begin{aligned}
&t=1\\
&u=a(0)\\
&\text{for}\quad i=1:n\\
&\quad t=x*t\\
&\quad u=u+a(i)*t\\
&\text{end}\\
&y=u
\end{aligned}
$$

计算量

$N=2n\ flop$

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

算法2 （秦九韶算法）

(输入$a(i)(i=0,1，\ldots,n),x$;输出y)

$$p=a(n)$$

$$\text{for } \ k=n-1:-1:0$$

$$p=x*p+a(k)$$

$$\text{end}$$

$$y=p$$

注意

计算量

$N=n \ flop$

其原理为

$$(\cdots((a_n x+a_{n-1})x+a_{n-2})x+\cdots+a_1)x+a_0$$

例3 矩阵乘积AB的计算量分析

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ \cdots & \cdots & \cdots & & \cdots \\ a_{m1} & a_{m2} & a_{m3} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & b_{13} & \ldots & b_{1s} \\ b_{21} & b_{22} & b_{23} & \ldots & b_{2s} \\ \cdots & \cdots & \cdots & & \cdots \\ b_{n1} & b_{n2} & b_{n3} & \ldots & b_{ns} \end{pmatrix} = [c_{ij}]_{m\times s}$$

A　　　　　　　　　　　　B

$$C_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \quad i = 1,\cdots,m;\ \ j = 1,\cdots,s$$

A B 的计算量为N= (m ×n ×s)flop

例4： 求解n元线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=b_1 \\ \quad\vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \\ a_{n1}x_1+a_{n2}x_2+\ldots+a_{nn}x_n=b_n \end{cases} \quad (1)$$

$Ax=b$

A可逆

由线性方程组的克莱姆(Cramer)规则可知,如果方程组(1)的系数矩阵A的行列式(一般记为D=|A|)不等于零,那么,这个方程组有唯一解,而且它们可以表示为

$$x_i=D_i/D \qquad (i=1,\cdots,n)$$

这里,D_i是指D中第i列元素用右端（$b_1,\cdots b_n$）代替构成的行列式。

克莱姆算法步骤

1. $$D = \sum_{(j_1 \cdots j_n)} (-1)^t a_{1j_1} a_{2j_2} \cdots a_{nj_n}$$

2. $for \quad i = 1 \cdots n$

   2.1. $$D_i = \sum_{(i_1 \cdots i_n)} (-1)^t a_{i_1 1} \cdots b_{i_2 j} \cdots a_{i_n n}$$

   2.2. $$x_i = \frac{D_i}{D}$$

N=[(n²−1)n!+n]flop

n=20, N= 9.707×10²⁰ flop

线性方程组 $Ax=b$，当 $n=20$ 时，用克莱姆（*Cramer*）法则求解运算次数约为9.7×10^{20}，用每秒1000亿次的计算机也要算300年。

用高斯（*Gauss*）消元法求解，运算次数约为3060.

第三节 数值计算的误差分析

用计算机解决科学计算问题时，需要经历以下几个环节：

实际问题 → 建立数学模型 → 确定数值计算方法 → 编制程序上机算出结果

实际问题的精确解与用计算机计算出来的数值结果之间就有差异，这种差异在数学上称为误差。

数值结果是指在选择某种数值方法之后，编制程序正确，输入初始数据正确的情形下所获得的数值结果。

一、误差的来源

1、数学模型

数学模型是通过科学实验或者观察分析一系列数据后，用数学作为工具近似地描述客观事物的一种数学表达式。

在数学模型中，往往包含了若干参量如物体比重、阻力系数、热交换系数等，这些物理参数通常由实验仪器测得，根据仪器的精密程度，物理参数的确定也会产生一定的误差。

例1　记L_t是金属棒在温度t的长度，L_0是t_0度时金属棒的长度，求在任何温度下金属棒的长度。

根据热胀冷缩原理和物理实验，可建立如下数学模型：

$$l_t = L_0(1+\alpha t+\beta t^2)$$

这里L_0是金属棒在t_0时的长度，α、β为物理参数并有如下估计：

$$\alpha = 0.001253 \pm 10^{-6}, \beta = 0.000068 \pm 10^{-6}$$

2、四种误差

模型误差　数学模型与实际问题之间出现的这种误差称为模型误差。在例1中，$L_t - l_t$ 就是模型误差。

观测误差　通过仪器观测，确定数学模型中的参数所引起这种的误差称为观测误差。在例1中的 10^{-6} 就是观测误差。

例2 求e^x的近似值。

$$e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots + \frac{1}{n!}x^n + \cdots,$$

$$S_3(x) = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3$$

截断误差 模型的准确解与某种数值方法的准确解之间的误差称为**截断误差**或方法误差。在例2中，

$e^x - S_3(x) = \frac{1}{4!}x^4 + \cdots$就是截断误差。

舍入误差 用计算机计算，由于计算机字长有限而在数值运算的每一步所产生的误差称为**舍入误差**。用4位浮点机计算 $\frac{1}{6}$ 所产生的误差

$\frac{1}{6} - 0.1667 = -0.0000334\cdots$ 就是舍入误差。

二、截断误差分析

例1：(截断误差)

已知$e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots + \frac{1}{n!}x^n + \cdots$,

求e^{-1}的近似值，并估计误差。

解：利用展开式的前三项，取n=2,

$$e^{-1} \approx 1 + (-1) + \frac{1}{2}(-1)^2 = 0.5$$

由$Taylor$公式：$f(x) = f(x_0) + f'(x_0)(x - x_0) + \cdots +$

$$\frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}$$

$$R_n(x) = \frac{x^{n+1}}{(n+1)!}e^{\theta x}, \qquad 0 < \theta < 1$$

截断误差 $|R_2| = |e^{-1} - 0.5| \le \frac{1}{3!} < 1.7*10^{-1}$

例2：当$h \to 0$时，有如下Taylor展开式

$$e^h = 1 + h + \frac{1}{2!}h^2 + O(h^3), \sin(h) = h - \frac{1}{3!}h^3 + O(h^5)$$

试确定近似计算公式$e^h + \sin(h) \approx 1 + 2h + \frac{1}{2!}h^2 - \frac{1}{3!}h^3$的截断误差。

解：$$e^h + \sin(h) = 1 + 2h + \frac{1}{2!}h^2 - \frac{1}{3!}h^3 + O(h^3) + O(h^5)$$

$e^h + \sin(h)$的截断误差为$O(h^3)$。

三、舍入误差分析

例1：舍入误差

$$1.492 \times 1.066 = 1.590472$$

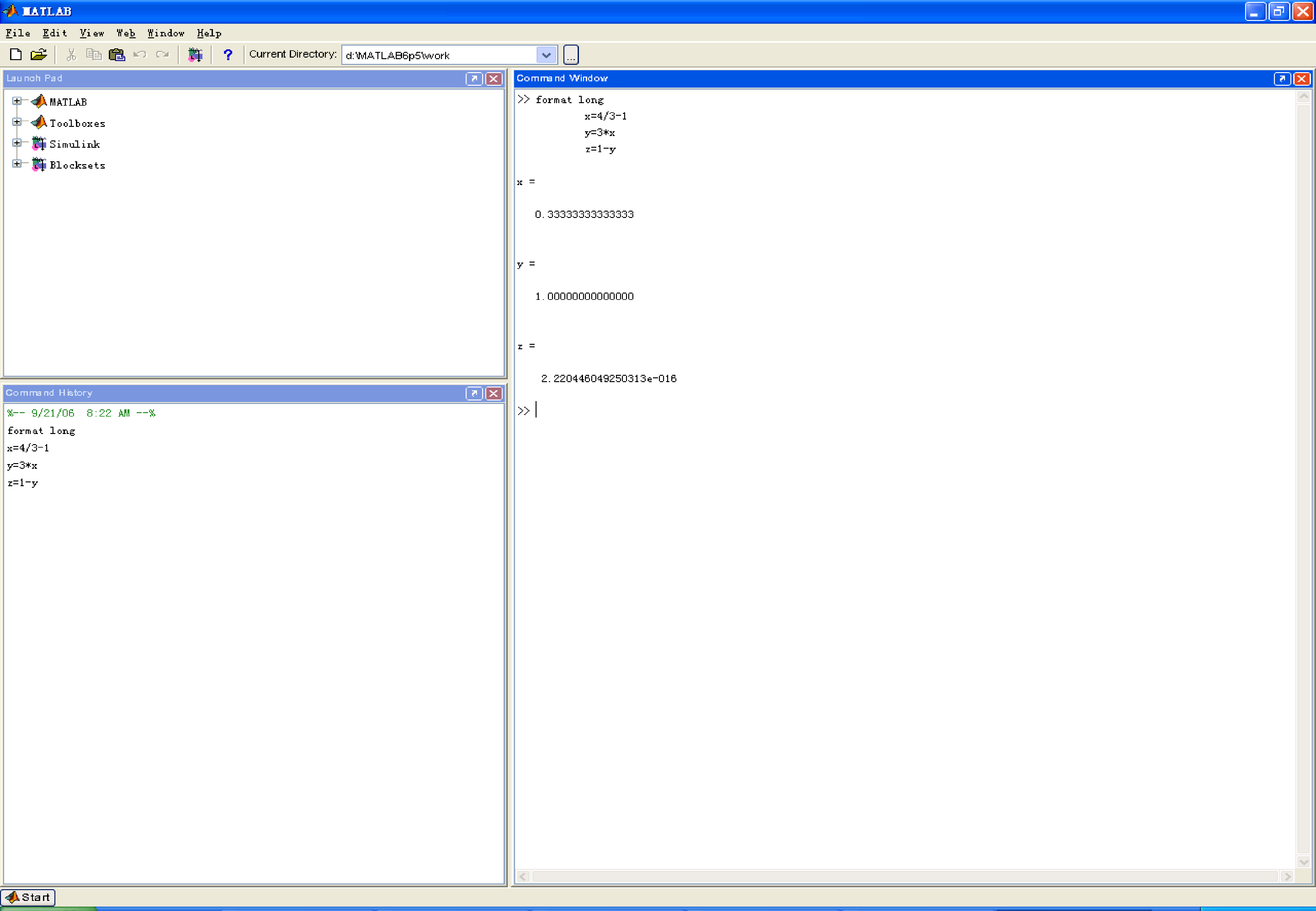
设在一台虚构的4位数字的计算机上计算

$$1.492 \times 1.066 \approx 1.590$$

舍入误差为 0.000472

例2：考虑Matlab简单程序

```
format long
x=4/3-1
y=3*x
z=1-y
```


MATLAB
File Edit View Web Window Help
Current Directory: d:\MATLAB6p5\work
Launch Pad
MATLAB
Toolboxes
Simulink
Blocksets
Command History
%-- 9/21/06 8:22 AM --%
format long
x=4/3-1
y=3*x
z=1-y
Command Window
>> format long
x=4/3-1
y=3*x
z=1-y
x =
0.33333333333333
y =
1.00000000000000
z =
2.220446049250313e-016
>>
Start

舍入误差对计算结果影响很大

例3: 考察方程组

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

其解为 $x_1 = x_2 = x_3 = 1$

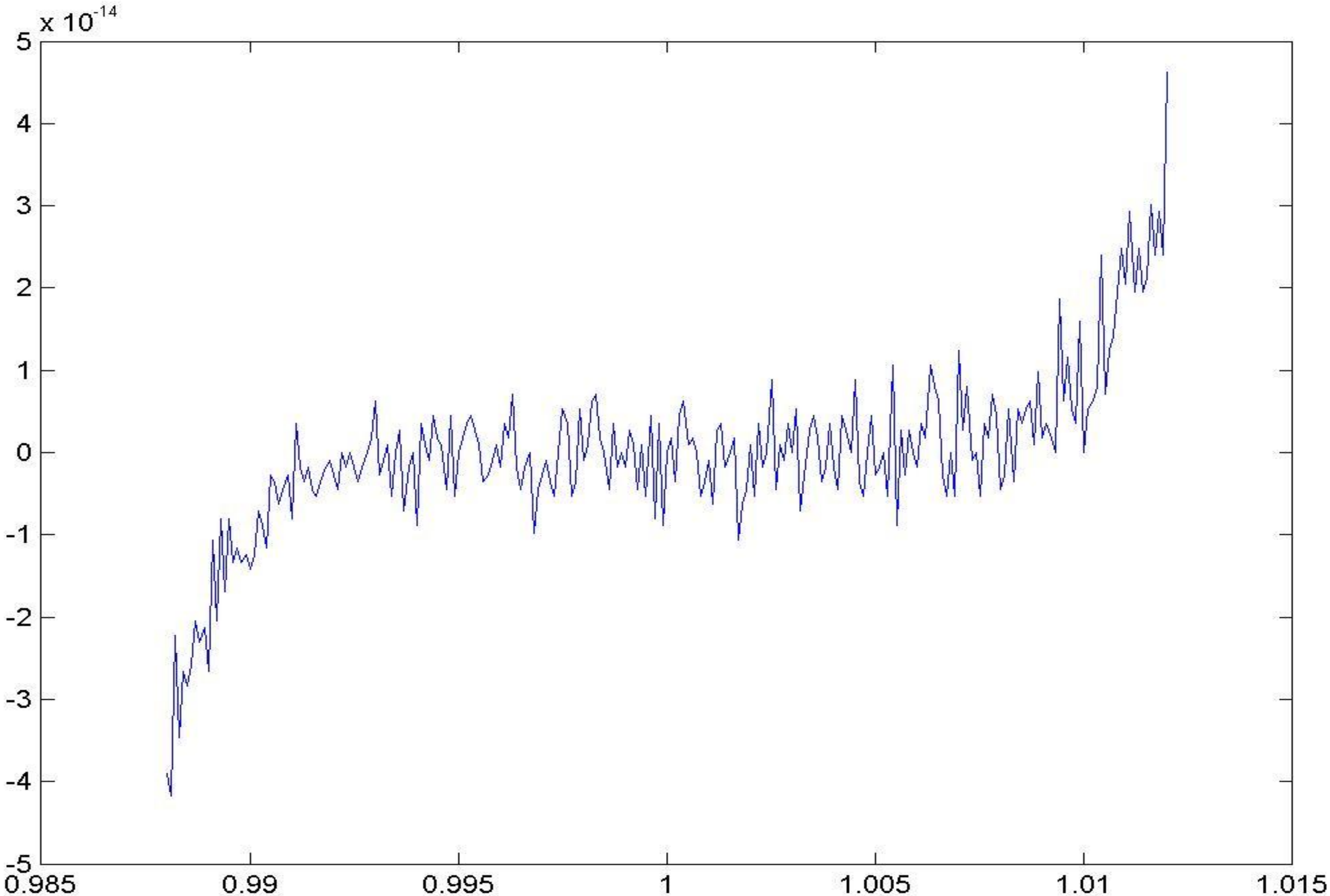
如果把系数舍入成三位数字

$$\begin{bmatrix} 1.000 & 0.500 & 0.333 \\ 0.500 & 0.333 & 0.250 \\ 0.333 & 0.250 & 0.200 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1.83 \\ 1.08 \\ 0.783 \end{bmatrix}$$

求解得 $x_1 = 1.09, \quad x_2 = 0.488, \quad x_3 = 1.49$

例4： 考虑Matlab程序

```
x=0.988:0.0001:1.012;
y=x.^7-7*x.^6+21* x.^5-35* x.^4+35* x.^3
  -21* x.^2+7* x -1;
plot(x, y)
```


x 10-14
5
4
3
2
1
0
-1
-2
-3
-4
-5
0.985
0.99
0.995
1
1.005
1.01
1.015

1、计算机数系

在实数系中，每一个实数可以有无穷位，不同的实数代表数轴上不同的点；

$$\sqrt{3} = 1.732050808 \quad \cdots$$

在计算机数系中，每一个数只有有限位，只有部分有理数能被计算机数系中的数精确表示。

浮点数：这种允许小数点位置浮动的表示法称为数的浮点形式。

实数x的十进制浮点形式为

$$x = \pm\ \underbrace{0.a_1 a_2 \ldots a_k \ldots}_{\text{尾数}} \times \underbrace{10}_{\text{基数}}^{\overbrace{c}^{\text{阶码}}}, \qquad (1)$$

$a_i \in \{0, 1, 2, \ldots, 9\}$， $c \in Z$

$a_1 \neq 0$, (1)称为x的规格化的浮点形式

x 的k位规格化十进制机器数

$y=\pm 0.a_1 a_2 \ldots a_k \times 10^c$, $y=fl(x)$

$a_i \in \{0,1,2,\ldots,9\}$, $a_1 \neq 0$，$L \leq c \leq U$,

k是机器数的字长； L、U 是常数。

x的k位十进制机器数$fl(x)$可用两种方法定义：

(1) 截断式

$x=\pm 0.a_1 a_2 \ldots a_k a_{k+1} \ldots \times 10^c$

$fl(x)=\pm 0.a_1 a_2 \ldots a_k \times 10^c$

(2) 四舍五入式

$$fl(x)=\begin{cases} \pm 0.a_1 a_2 \cdots a_k \times 10^c & a_{k+1}<5 \\ \pm(0.a_1 a_2 \cdots a_k + 10^{-k}) \times 10^c & a_{k+1} \geq 5 \end{cases}$$

一般数制情况：k位规格化机器数

$y = \pm\ 0.\ a_1\ a_2 \ldots\ a_k \times \beta^c$，　　$\beta = 2, 8, 10, 16,$

$a_i \in \{0, 1, 2, \ldots, \beta-1\}$，$L \le c \le U$, $a_1 \ne 0$

F(β,k.L,U)表示以上数集全体加数0，它是计算机中使用的有限离散数集(机器数系）。

F(β,k,L,U)中的数称为机器数。

F(10,4,-33,33)，$y = \pm\ 0.\ a_1 a_2 a_3 a_4 \times 10^c$

例5 在机器数系 F(10,4,-33,33)中表示 $fl(\pi)$.

$$\pi = 3.1415926\cdots \notin F(10,4,-33,33),$$

但是

$$0.1000\times 10^{-33} < \pi < 0.9999\times 10^{33}$$

采用截断式 $fl(\pi)=0.3141\times 10$

采用四舍五入式 $fl(\pi)=0.3142\times 10$

若浮点数的阶码不在[L,U]内，则出现上溢或下溢。

例如 在4位机器数系 F(10,4,-33,33)中输入 0.28×10^{-34} 出现下溢，输入 0.199×10^{35}出现上溢。

2、数的误差和有效数字

定义： 设数a是精确值，x是a的一个近似值，记

$$e=|a-x|$$ 称e为近似值x的绝对误差，

$$e_r=\frac{|a-x|}{|a|}=\frac{e}{|a|}$$ 称e_r为近似值x的相对误差。

$e=|a-x|\le\delta x$，称δx为数a的近似值x的绝对误差限，

$e_r=\dfrac{|a-x|}{|a|}\le\delta_r x$，称$\delta_r x$为近似值$x$的相对误差限。

一般准确值a未知,实际计算中 $e_r=|a-x|/|x|=e/|x|$

例6　已知准确值a=3. 1415926...是一个无限不循环小数，求截取不同位数后的近似值和误差界。

解：　截取一位　$x_1 = 3,$

$$e_1 = |a - x_1| \approx 0.14 < \frac{1}{2} \times 10^0;$$

截取三位　$x_2 = 3.14,$

$$e_2 = |a - x_2| \approx 0.0016 < \frac{1}{2} \times 10^{-2};$$

截取五位　$x_3 = 3.1416,$

$$e_3 = |a - x_3| \approx 0.0000074 < \frac{1}{2} \times 10^{-4}。$$

a的近似值的绝对误差都超不过本身末位数的半个单位

是截取相应位数后，所得到的近似数中绝对误差的最小值。

a的近似值的绝对误差都超不过本身末位数的半个单位
是截取相应位数后，所得到的近似数中绝对误差的最小值。

$$\text{截取五位}\, x_3 = 3.1416,$$

$$e_3 = |a - x_3| \approx 0.0000074 < \frac{1}{2}\times 10^{-4}$$

$$\text{若取}\quad x_4 = 3.1415$$

$$e_4 = |a - x_4| \approx 0.000093 < \frac{1}{2}\times 10^{-3}$$

定义 设近似数$x = \pm 0.a_1 a_2 \cdots a_n \times 10^p$的绝对误差限是第$n$位的半个单位，则数$x$有$n$位有效数字。（$a_1 \neq 0, a_i \in \{0,1,\ldots,9\}$）

即 $|x^* - x| \leq 0.5 \times 10^{p-n}$

$$|x^*-x|\le 0.5\times 10^{p-n}$$

例8：设计算机数系为F(10,t,L,U),将实数 $x=\pm 0.a_1a_2\ldots a_ta_{t+1}\ldots\times 10^c,(a_1\neq 0)$,用四舍五入法表为机器数$fl(x)$;求其有效数字、绝对误差限、相对误差限。

解：
$$fl(x)=\pm\bar{a}\times 10^c$$

$$\bar{a}=\begin{cases}0.a_1a_2\cdots a_t, & a_{t+1}<5\\ 0.a_1a_2\cdots a_t+10^{-t}, & a_{t+1}\ge 5\end{cases}$$

绝对误差为：
$$|x-fl(x)|\le 0.5\times 10^{c-t}$$

$fl(x)$有t位有效数字。

相对误差为
$$e_r=\frac{|x-fl(x)|}{|x|}\le\frac{0.5\times 10^{c-t}}{0.1\times 10^c}=5\times 10^{-t}$$

机器数的相对误差与x无关，只与字长t有关。

例7: 设 $x=\sqrt{3}=1.7320508\ldots=0.17320508\times 10^{1}$

$x_1=1.73$, $x_2=1.7321$, $x_3=1.7320$是其近似值,

问它们分别有几位有效数字?

解: $e_1=\left|\sqrt{3}-x_1\right|\approx 0.0021<\frac{1}{2}\times 10^{-2}$ 10^{1-3} 3位

$e_2=\left|\sqrt{3}-x_2\right|\approx 0.00005=\frac{1}{2}\times 10^{-4}$ 10^{1-5} 5位

$e_3=\left|\sqrt{3}-x_3\right|\approx 0.000051<\frac{1}{2}\times 10^{-3}$ 10^{1-4} 4位

当某个量的准确值很小或很大时，相对误差比绝对误差更能反映准确数与近似数的差异。

例：考虑 1. $x=11$，$a=10$， $e=1$， $e_r=0.1$

2. $x=1001$，$a=1000$，$e=1$，$e_r=0.001$

3. $x=1001$，$a=1101$，$e=100$，$e_r=0.1$

一个近似值的准确程度，不仅与绝对误差的大小有关，而且与准确值本身的大小有关。

注：

在数值计算中，尽可能多地保留近似数的有效数字，有效数字越多，相对误差越小，计算结果越精确。

3、数值运算中的误差估计

（1）数值运算的绝对误差和相对误差

$$u = f(x,y) \qquad f:\text{表示某一种确定的运算}$$

$$f(a,b) = f(x,y) + \frac{\partial f(x,y)}{\partial x}(a-x) + \frac{\partial f(x,y)}{\partial y}(b-y) + \cdots$$

二元函数的 Taylor公式

$\boxed{u\text{ 的绝对误差}}$

$$e(u) = |u^* - u| = |f(a,b) - f(x,y)|$$

$$\approx \left| \frac{\partial f(x,y)}{\partial x}(a-x) + \frac{\partial f(x,y)}{\partial y}(b-y) \right|$$

$$\le \left| \frac{\partial f(x,y)}{\partial x} \right| |e(x)| + \left| \frac{\partial f(x,y)}{\partial y} \right| |e(y)| \le \left| \frac{\partial f}{\partial x} \right| \delta x + \left| \frac{\partial f}{\partial y} \right| \delta y$$

u 的相对误差

$$e_r(u)=\frac{e(u)}{|u|}\approx\left|\frac{\partial f(x,y)}{\partial x}\right|\frac{e(x)}{|u|}+\left|\frac{\partial f(x,y)}{\partial y}\right|\frac{e(y)}{|u|}$$

$$=\left|\frac{\partial f(x,y)}{\partial x}\right|\frac{|x|}{|u|}\frac{e(x)}{|x|}+\left|\frac{\partial f(x,y)}{\partial y}\right|\frac{|y|}{|u|}\frac{e(y)}{|y|}$$

$$=\left|\frac{\partial f(x,y)}{\partial x}\right|\frac{|x|}{|u|}e_r(x)+\left|\frac{\partial f(x,y)}{\partial y}\right|\frac{|y|}{|u|}e_r(y)$$

$$\leq\left|\frac{\partial f}{\partial x}\right|\frac{|x|}{|u|}\delta_r(x)+\left|\frac{\partial f}{\partial y}\right|\frac{|y|}{|u|}\delta_r(y)$$

(2) 和、差、积、商的误差估计式

$$e(u) \le \left|\frac{\partial f}{\partial x}\right| e(x) + \left|\frac{\partial f}{\partial y}\right| e(y) \qquad e_r(u) = \left|\frac{\partial f(x,y)}{\partial x}\right| \left|\frac{x}{u}\right| e_r(x) + \left|\frac{\partial f(x,y)}{\partial y}\right| \left|\frac{y}{u}\right| e_r(y)$$

由以上公式可以推导和 差、积、商的误差估计式

(1) $e(x \pm y) \approx e(x) + e(y)$

$$e_r(x \pm y) \approx \left|\frac{x}{x \pm y}\right| e_r(x) + \left|\frac{y}{x \pm y}\right| e_r(y)$$

(2) $e(xy) \approx |y| e(x) + |x| e(y)$

$$e_r(xy) \approx e_r(x) + e_r(y)$$

(3) $e(\frac{x}{y}) \approx \left|\frac{1}{y}\right| e(x) + \left|\frac{x}{y^2}\right| e(y)$

$$e_r(\frac{x}{y}) \approx e_r(x) + e_r(y)$$

4. 数值计算中应注意的问题

(1) 防止相近的两数相减（防止严重丢失有效数字）

例1:当x较大时,计算 $\sqrt{x+1}-\sqrt{x}$

解：$\sqrt{x+1}-\sqrt{x}=\dfrac{1}{\sqrt{x+1}+\sqrt{x}}$

例2 计算 $\dfrac{1-\cos x}{\sin x}$，$x=2^{\circ}$

解：当x很小时，分子出现相近数相减，将以上算式变形

$$\frac{(1-\cos x)(1+\cos x)}{\sin x(1+\cos x)}=\frac{1-\cos^2 x}{\sin x(1+\cos x)}=\frac{\sin x}{1+\cos x}$$

(2) 防止大数吃小数

当两个绝对值相差很大的数进行加法或减法运算时,绝对值小的数有可能被绝对值大的数"吃掉"从而引起计算结果很不可靠.

先对阶, 后运算, 再舍入

例3:在F(10,5,-119,119)中，计算 23456+0.2+0.4+0.4

$$上式=0.23456\times10^5+0.000002\times10^5+0.000004\times10^5$$

$$+0.000004\times10^5=23456$$

在上式中，重新排序计算

$$上式=0.2+0.4+0.4+23456=1+23456=0.00001\times10^5+$$

$$0.23456\times10^5=23457$$ ✓

(3) 防止接近零的数做除数

分母接近零的数会产生溢出错误,因而产生大的误差,严重影响计算结果的精度，此时可以用数学公式化简后再做.

$$y=\frac{1605}{\sqrt{1001}-\sqrt{1000}}=1605(\sqrt{1001}+\sqrt{1000})$$

(4) 注意计算步骤的简化,减小运算次数

简化计算步骤是提高程序执行速度的关键，它不仅可以节省时间，还能减少舍入误差。

例4： 设A、B、C、D分别是10×20、 20×50、 50×1、1×100的矩阵，试按不同的算法求矩阵乘积E=ABCD.

解： 由矩阵乘法的结合律，可有如下算法

1. E=((AB)C)D.　　　计算量N=11500flop
2. E=A(B(CD)).　　　计算量N=125000flop
3. E=(A(BC))D.　　　计算量N=2200flop ✓

例5 计算 $\ln 2$，要精确到10^{-5}.

解： $\ln(1+x)=\sum_{n=1}^{\infty}(-1)^{n+1}\frac{x^n}{n}$ 取 $x=1$，需要计算十万项

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

取 $x=\frac{1}{3}$，只须计算前9项

(5) 不要与实数相等比较

由于十进制与二进制转换及舍入误差的影响，实数相等的比较一般是不成功的。例如，用数值方法验证等式

$$\sin^2 x + \cos^2 x = 1$$ 总是不成功.

正确的做法是设置一个很小的误差限 ε 去验证

$$\left|\sin^2 x + \cos^2 x - 1\right| \le \varepsilon$$

四、误差的定性分析

误差的定量分析法（对复杂的实际计算很难估计出切合实际的定量分析）

误差的定性分析法：即研究算法的数值稳定性。

1、**解的稳定性**，常称为解适定性（数学问题的稳定性）

解对原始数据连续依赖；原问题固有的特性

2、 **数值稳定性**：一个算法如果输入数据有扰动（即有误差），而计算过程中舍入误差不增长，则称此算法是数值稳定的，否则称此算法为不稳定的。

是对某数值方法而言的；在原始问题是良态的情形下才有意义的；误差传播能得到控制的。

例1：计算 $I_n = \frac{1}{e}\int_0^1 x^n e^x dx\,, \quad n = 0, 1, 2, \ldots\ldots$

方法一：

$$I_n = e^{-1}\left(x^n e^x \Big|_0^1 - n\int_0^1 x^{n-1} e^x dx \right)$$

此式精确成立

$$= 1 - n\, e^{-1}\int_0^1 x^{n-1} e^x dx = 1 - nI_{n-1}$$

$$I_0 = \frac{1}{e}\int_0^1 e^x dx = 1 - \frac{1}{e} \approx 0.63212056 \overset{\text{记为}}{=\!=\!=} I_0^*$$

则初始误差 $|E_0| = |I_0 - I_0^*| < 0.5\times 10^{-8}$

$$I_1^* = 1 - 1 \cdot I_0^* = 0.36787944$$

$$\cdots\ \cdots\ \cdots\ \cdots$$

$$I_{10}^* = 1 - 10 \cdot I_9^* = 0.08812800$$

$$I_{11}^* = 1 - 11 \cdot I_{10}^* = 0.03059200$$

$$I_{12}^* = 1 - 12 \cdot I_{11}^* = 0.63289600 \ ?$$

$$I_{13}^* = 1 - 13 \cdot I_{12}^* = -7.2276480 \ ??$$

$$I_{14}^* = 1 - 14 \cdot I_{13}^* = 94.959424 \quad ?\,!$$

$$I_{15}^* = 1 - 15 \cdot I_{14}^* = -1423.3914 \quad !\,!$$

考察第n步的误差 $|E_n|$

$$|E_n| = |I_n - I_n^*| = |(1 - nI_{n-1}) - (1 - nI_{n-1}^*)|$$

$$= n|E_{n-1}| = \ldots = n!\,|E_0|$$

可见初始的小扰动 $|E_0| < 0.5\times10^{-8}$ 迅速积累，误差呈递增走势。

造成这种情况的是不稳定的算法.

✍ 方法二： $I_n = 1 - n I_{n-1} \quad \Rightarrow \quad I_{n-1} = \frac{1}{n}(1 - I_n)$

方法：先估计一个I_N ,再反推要求的I_n ($n << N$)。

$$\because \frac{1}{e}\int_0^1 x^n \cdot e^0 dx < I_n < \frac{1}{e}\int_0^1 x^n \cdot e^1 dx \qquad \therefore \frac{1}{e(N+1)} < I_N < \frac{1}{N+1} \qquad (n = N)$$

可取 $$I_N^* = \frac{1}{2}\left[\frac{1}{e(N+1)} + \frac{1}{N+1}\right] \approx I_N$$

当$N \to +\infty$时, $|E_N| = |I_N - I_N^*| \to 0$

取 $$I_{15}^* = \frac{1}{2}\left[\frac{1}{e\cdot 16} + \frac{1}{16}\right] \approx 0.042746233$$

$$\Rightarrow I_{14}^* = \frac{1}{15}(1 - I_{15}^*) \approx 0.063816918$$

$$I_{13}^* = \frac{1}{14}(1 - I_{14}^*) \approx 0.066870220$$

$$I_{12}^* = \frac{1}{13}(1 - I_{13}^*) \approx 0.071779214$$

$$I_{11}^* = \frac{1}{12}(1 - I_{12}^*) \approx 0.077351732$$

$$I_{10}^* = \frac{1}{11}(1 - I_{11}^*) \approx 0.083877115$$

$$\cdots\ \cdots\ \cdots\ \cdots$$

$$I_{1}^* = \frac{1}{2}(1 - I_{2}^*) \approx 0.36787944$$

$$I_{0}^* = \frac{1}{1}(1 - I_{1}^*) \approx 0.63212056$$

考察反推一步的误差：

$$|E_{N-1}| = \left| \frac{1}{N}(1 - I_N) - \frac{1}{N}(1 - I_N^*) \right| = \frac{1}{N}|E_N|$$

以此类推，对 $n < N$ 有：

$$|E_n| = \frac{1}{N(N-1)\dots(n+1)}|E_N|$$

误差逐步递减, 这样的算法称为稳定的算法 /* stable algorithm */

在我们今后的讨论中，误差将不可回避，

算法的稳定性会是一个非常重要的话题。

无条件稳定与条件稳定

对任何输入数据都是稳定的算法称为无条件稳定。

对某些数据稳定，而对另一些数据不稳定的算法称为条件稳定。

例2: 求二次方程 $ax^2+bx+c=0$ 的根。

方法1: $$x_{1,2}=\frac{-b\pm\sqrt{b^2-4ac}}{2a},$$

当$b^2>>4ac$时算法不稳定。

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

其中 $$\mathrm{sign}(b)=\begin{cases}1 & b\geq 0\\ -1 & b<0\end{cases}$$ 此算法无条件稳定。

例3 在F(10,4,-19,19)数系中，求解二次方程：

$$x^2 - 320x + 16 = 0 \qquad b^2 >> 4ac$$

解法1　　按求根公式　$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

解得，$x_1=0.3199\times10^3$，$x_2=0.1000\times10$；

解法2

$$x_1 = \frac{-b - sign(b)\sqrt{b^2 - 4ac}}{2a}$$

$$x_2 = \frac{c}{ax_1} \qquad (x_1 x_2 = \frac{c}{a})$$

解得$x_1=0.3199\times10^3$

$x_2=0.5002\times10^{-1}$

精确解为

$x_1=319.950$

$x_2=0.0500078$

病态问题与条件数

对数学问题而言，如果输入数据有微小扰动，引起输出数据（即数学问题的解）有很大扰动，则称数学问题是病态问题，否则称为良态问题。

例1: 求$f(x)=x^2+x-1150$在$x=\frac{100}{3}$处的值。

解: $$f(\frac{100}{3})=-\frac{50}{9}\approx -5.6$$

若x有误差δx, $x+\delta x=33$,

$|\delta x|=|\frac{100}{3}-33|<0.34$变化不大, $f(33)=-28$

而$f(\frac{100}{3})$的误差$|f(x+\delta x)-f(x)|=|28-5.6|$

$=22.4$ 却很大

因此，此问题在 $x=\frac{100}{3}$ 接近 $f(x)=0$

的根处是一个病态问题。

例2: 求 $f(x)=x^2+x-1150$ 在 $x=2$处的值。

解: $f(2)=-1144$

若x有误差δx, $x+\delta x=2.4$,

$|\delta x|=0.4$　　$f(2.4)=-1141.84$

$f(2)$的误差　　$|f(x+\delta x)-f(x)|=2.16$

x在远离$f(x)=0$的根处是一个良态问题。

条件数

设两个不同的数据$\boldsymbol{x},\overline{\boldsymbol{x}}$，对应的函数值为$\boldsymbol{f(x)}$，$\boldsymbol{f(\overline{x})}$，假设$\boldsymbol{x \neq 0, f(x) \neq 0}$.

相对误差 $e_r = \dfrac{|x-\overline{x}|}{|x|}$， $R = \dfrac{|f(x)-f(\overline{x})|}{|f(x)|}$

如果能找到一个数$\boldsymbol{m}$满足$\boldsymbol{R \le m e_r}$，则称$\boldsymbol{m}$为该问题的条件数，记为$\boldsymbol{Cond((f(x))}$。

推导计算函数值的条件数

$$f(x)-f(\bar{x})=f'(\xi)(x-\bar{x})$$

当$|x-\bar{x}|$足够小时，由$f'(x)$的连续性，$f'(\xi)\approx f'(x)$

$$f(x)-f(\bar{x})\approx f'(x)(x-\bar{x})$$

于是
$$\frac{f(x)-f(\bar{x})}{f(x)}\approx x\frac{f'(x)}{f(x)}\frac{(x-\bar{x})}{x}$$

$$R\le\left|x\frac{f'(x)}{f(x)}\right|e_r$$

函数值的条件数$Cond((f(x))=\left|\dfrac{xf'(x)}{f(x)}\right|$

函数值的条件数 $Cond((f(x)) = \left| \dfrac{xf'(x)}{f(x)} \right|$

上例中

$f(x) = x^2 + x - 1150$

$x = \dfrac{100}{3}$,　　$Cond((f(x)) = 406$ →大，病态

$x = 2$,　　$Cond((f(x)) = 0.0087$ →小，良态

第四节 数学软件工具

一、几种常用的数学软件

目前流行的数学软件主要有以下几种：

符号运算软件： Mathematica, Maple

矩阵处理软件： Matlab

统计处理软件： SAS, Spss, Origin

数学CAD软件： MathCAD

1、 符号运算软件：Mathematica, Maple

一提起计算机求解，人们就会想到数值计算。但随着符号运算及计算机代数理论的日益成熟，人们又利用计算机来进行符号计算。70年代人们开发了几个较为成功的通用的计算机代数系统，如Reduce系统和MACSYMA系统。但由于速度慢，开发不方便，因此没有流行起来。在1988年推出的Mathematica软件才使得符号运算系统的使用普及起来。在这方面代表性的软件就是Mathematica和Maple。虽然在Matlab里也有符号计算系统，但它采用的是Maple内核。

1、 符号运算软件：Mathematica, Maple

Mathematica：美国Wolfram公司出品。1988 年推出第一版版本，刚推出时受到极大好评，New York 时代周刊这样评论："the importance of the program cannot be overlooked"，Business Week 把Mathematica列为当年十大最重要的新产品之一。Mathematica在技术界也被作为智力和应用的革命性产品受到极大欢迎。人们常说Mathematica的出现标志着现代技术计算的开始。

1、 符号运算软件：Mathematica, Maple

Maple：加拿大**Mathsoft**公司出品。**1980**年 加拿大**Waterloo**大学开始研究符号运算系统。取这个名字是为了表明这个软件是加拿大生产的。

Maple功能基本和**Mathematica**一致。由于这两个软件功能相似，因此统一介绍这两个软件的功能和特点。

功能：

数值计算：初等计算、线性代数计算、数值方法等

符号运算：多项式因式分解与展开、微分、积分、级数、数列等。

绘图功能：方便、强大、直观的二、三维绘图。

特点：

1) 简单易用的交互式操作方式:两种软件都使用英文单词来命名函数，非常直观、易用。对每一个命令都能给出显示结果，非常方便。

2)功能强大:强大的数值计算功能和威力无比的符号运算功能。

3)结构化的程序设计语言

4)Tex输出格式:Tex是科技文章的一种格式，科技界应用十分广泛。

2、 矩阵处理软件： Matlab

●简介：

MATLAB的含义是矩阵实验室,是Matrix Laboratory的缩写。它的前身是LINPACK（解线性方程）和EISPACK（解特征值问题）的FORTRAN子程序库。由于它把矩阵当成一个对象，因此编写程序更加直观、方便。

MATLAB具有非常强大和直观的计算功能，并且由于其有非常好的扩展性能，现在已经成为世界上应用最广泛的工程计算软件之一

● **特点:**

(1)强大的数值运算功能

在MATLAB环境中，有超过500种数学、统计、科学及工程方面的函数可使用，函数的命名表示自然，使得问题和解答像数学公式一般简单明了，让用户可全力发挥在解题方面，而非浪费在电脑操作上。

(2)数据分析和可视化功能、文字处理功能

MATLAB可以绘制二、三维图形，与Mathematic和Maple相比，它还能处理光照模型，制作出高品质的图形。功能十分强大。MATLAB Notebook为用户提供了强大的文字处理功能，并允许WORD访问MATLAB的数值计算和可视化结果，制作科学性或工程性图文并茂的文章.

(3)高级、简单、高效的程序环境

做为一种解释型的程序语言，**MATLAB**允许使用者在短时间内写完程序,所花的时间约为用 **FORTRAN** 或 **C** 的几分之一，而且不需要编译 **(compile)** 及连接 **(link)** 即能执行，同时包含了更多及更容易使用的内建功能。

(4)开放及可延伸的架构

MATLAB允许使用者接触它的大多数的数学源代码，检查运算法，更改现有函数，甚至加入自己的函数使 **MATLAB**成为使用者所需要的环境。

(5)丰富的工具箱

MATLAB的工具箱融合了套装前软体的优点，与一个灵活的开放但容易操作之环境，这些工具箱提供了使用者在特别应用领域所需的许多函数。现有工具箱有：符号运算（利用Maple V的计算核心执行）、图像处理、统计分析、信号处理、通信、线性矩阵不等式、偏微分方程、高阶谱分析、财政金融、神经网络、模拟分析、控制系统、实时控制、小波分析、最优化、模糊逻辑、μ分析及合成等30多种。

3、数学CAD软件： MathCAD

Mathcad是由MathSoft公司推出的一种交互式数值系统。在输入一个数学公式、方程组、矩阵之后，计算机能直接给出结果，而无须去考虑中间计算过程。最令人激动的是在加入软件包自带的Maple插件后能直接支持符号运算。你可以在计算机上输入数学公式、符号和等式等，很容易地算出代数、积分、三角以及很多科技领域中的复杂表达式的值，并可显示数学表格和图形,通过对图形结果的分析，使我们对问题的理解更加形象。

Mathcad的用户主要针对具备应用数学知识但并不要求具有较多的计算机知识的用户，如工程研究人员、学生等。

特点：公式编辑功能强大、直观性强。

功能：

- 公式编辑：**MathCAD**的最吸引人的地方在于编辑公式非常方便、实用，几乎没有不能表示的数学公式。
- 矩阵功能

  定义的矩阵或矢量

  矩阵操作：矩阵转置、求逆矩阵、求矩阵的行列式的值。

- 数学计算功能

解方程：一般方程、线性方程组、不定方程、
常微分方程、偏微分方程等

数理统计与数据处理：统计函数、统计分布函数、
插值预测、曲线拟合（回归）等。

积分变换：**Fourier**变换、**Laplace**变换、**Z**变换、
小波变换等。

- 符号处理：借用的是**Maple**内核。
- 图形功能：包括绘制二维和三维图形。

作业

习题一

P27----2, 3, 10, 14