第零章　绪论

计算方法概述

实际问题　　现实中，具体的科学、工程问题的解决：

↓

物理模型

↓

数学模型

↓

数值方法

计算方法是一种研究并解决数学问题的数值近似解方法

↓

计算机求结果

计算方法概述

输入复杂问题或运算

$$\sqrt{x},\quad a^{x},\quad \ln x,\quad A\vec{x}=\vec{b},$$

$$\int_a^b f(x)dx,\quad \frac{d}{dx}f(x),\quad \ldots\ldots$$

→ 数值分析 → $+\ -\ \times\ \div$ → 计算机 → 近似解

- Numerical analysis: involves the study, development, and analysis of *algorithms* for obtaining *numerical solutions* to various mathematical problems
- Scientific computing: solving *mathematical* problems *numerically* on the *computer* is scientific computing
- Numerical analysis is called the mathematics of scientific computing

计算方法课程的特点

- 理论性：数学基础
- 实践性：算法实现
- 计算方法是连接模型到结果的重要环节
- 科学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
- 理论方法＋实验方法＋科学计算方法
- 本课仅限介绍最常用数学模型的最基本数值求解方法

数值计算方法的基本内容

- **数值逼近** – 数学分析中的数值求解，如微分、积分等

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

- **数值代数** – 线性代数的数值求解，如解线性方程组、逆矩阵、特征值、特征向量

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

100亿/秒，算3,000年，而Gauss消元法2660次

- **微分方程数值解** – 常微分方程，积分方程，偏微分方程等，如*Runge-Kutta*法、打靶法，有限差分法，有限元法，有限体积法，边界元法，谱方法等

误差

- **绝对误差**：设 x^* 为精确值，x 为近似值，$e = x^* - x$ 为误差或绝对误差
- **例如**：

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

有限计算，截断误差

$$\pi = 3.14159265358979323846264338327 95\ldots$$
$$\approx 3.14159265358979$$

有限精度，舍入误差

误差

- **相对误差**

$$e_r = \frac{e}{x^*} = \frac{x^* - x}{x^*}$$ 称为相对误差

- **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

有效位数

- 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的有效位数
- 有效位的多少直接影响到近似值的绝对误差和相对误差
- 例：

π的近似值3.141具有几位有效位数?

π的近似值3.142具有几位有效位数?

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
- 测量误差－观测误差（测量引起的误差）
- 方法误差－截断误差（算法本身引起）
- 计算误差－舍入误差（计算机表示数据引起）

误差的估计

■ **绝对误差估计**

$$e(x_1 \pm x_2) \approx e(x_1) \pm e(x_2)$$

$$e(x_1 \cdot x_2) \approx x_2 e(x_1) + x_1 e(x_2)$$

$$e\left(x_1 / x_2\right) \approx e(x_1) / x_2 - e(x_2) \cdot x_1 / x_2^2$$

小数作除数，绝对误差增大

误差的估计

■ 相对误差估计

$$e_r(x_1+x_2)\approx\frac{x_1}{x_1+x_2}e_r(x_1)+\frac{x_2}{x_1+x_2}e_r(x_2)$$

$$e_r(x_1-x_2)\approx\frac{x_1}{x_1-x_2}e_r(x_1)-\frac{x_2}{x_1-x_2}e_r(x_2)$$

$$e_r(x_1\cdot x_2)\approx e_r(x_1)+e_r(x_2)$$

$$e_r(x_1/x_2)\approx e_r(x_1)-e_r(x_2)$$

两相近数相减，相对误差增大

避免误差危害的若干原则

- 选择收敛、稳定的算法
- 提高数值计算精度
  - 尽可能使用double等高精度的表示
  - 需要在内存、计算时间与计算精度之间作出平衡
- 尽可能避免两个相近的数相减
- 尽可能避免绝对值很小的数作除数
- 尽可能避免大数“吃”小数的现象

向量范数

- 定义：对任一向量 $\mathbf{x}\in\mathbb{R}^n$，按照一个规则确定一个实数与它对应，记该实数为 $\|\mathbf{x}\|$，即映射：$\|\cdot\|: R^n \to R^+ \cup \{0\}$ 。若 $\|\mathbf{x}\|$ 满足下面三个性质：

  （1）非负性 $\forall \mathbf{x}\in\mathbb{R}^n,\ \|\mathbf{x}\|\ge 0,\ \text{且}\ \|\mathbf{x}\|=0 \Leftrightarrow \mathbf{x}=0$

  （2）齐次性 $\forall a\in\mathbb{R},\ \|a\mathbf{x}\|=|a|\cdot\|\mathbf{x}\|$

  （3）三角不等式 $\forall \mathbf{x},\mathbf{y}\in\mathbb{R}^n,\ \|\mathbf{x}+\mathbf{y}\|\le\|\mathbf{x}\|+\|\mathbf{y}\|$

  则称 $\|\mathbf{x}\|$ 为向量范数。

- 作用：
  - 向量之间的距离、误差
  - 向量序列的收敛性
  - 向量的邻域、开集、连续性等
  - ……

向量范数

- 常见的范数有：

$$\|\mathbf{x}\|_1=\sum_{i=1}^{n}|x_i|,\ \|\mathbf{x}\|_2=\sqrt{\sum_{i=1}^{n}x_i^{\ 2}},\ \|\mathbf{x}\|_\infty=\max\{|x_i|\}$$

$$\|\mathbf{x}\|_p=\left(\sum_{i=1}^{n}|x_i|^p\right)^{1/p},\ \mathbf{x}=\{x_1,x_2,\cdots,x_n\}$$

- （范数等价性）设 $\|\mathbf{x}\|_p$ 和 $\|\mathbf{x}\|_q$ 为任意两向量范数，则存在与 x 无关的正常数 c_1 与 c_2 使得

$$c_1\|\mathbf{x}\|_q\le\|\mathbf{x}\|_p\le c_2\|\mathbf{x}\|_q,\ \forall\mathbf{x}\in\mathbb{R}^n$$

- 一些常用范数的等价关系：

$$\|\mathbf{x}\|_2\le\|\mathbf{x}\|_1\le\sqrt{n}\|\mathbf{x}\|_2$$

$$\|\mathbf{x}\|_\infty\le\|\mathbf{x}\|_2\le\sqrt{n}\|\mathbf{x}\|_\infty$$

$$\|\mathbf{x}\|_\infty\le\|\mathbf{x}\|_1\le n\|\mathbf{x}\|_\infty$$

向量范数

- **定义：设 $\mathbf{x}^{(1)},\mathbf{x}^{(2)},\ldots,\mathbf{x}^{(m)},\ldots$ 为 $\mathbb{R}^n$ 空间内的向量序列，若存在向量 $\boldsymbol{\alpha}\in\mathbb{R}^n$ 使得 $\lim\limits_{m\to\infty}\|\mathbf{x}^{(m)}-\boldsymbol{\alpha}\|=0$，则称向量序列 $\mathbf{x}^{(m)}$ 是收敛的，向量 $\boldsymbol{\alpha}$ 称为向量序列 $\mathbf{x}^{(m)}$ 的极限**
- **性质：**
  - 范数 $\|\mathbf{x}\|=\|(x_1,x_2,\ldots,x_n)\|$ 是关于坐标 $x_1,x_2,\ldots,x_n$ 的连续函数
  - 向量序列 $\mathbf{x}^{(m)}=(x_1^{(m)},x_2^{(m)},\ldots,x_n^{(m)})^T$ 收敛的充分必要条件为序列的每个分量收敛，即 $\alpha_i=\lim\limits_{m\to\infty}x_i^{(m)},\ i=1,2,\ldots,n$

矩阵范数

- 定义：诱导矩阵范数（或导出范数，或从属范数）

$$\|\mathbf{A}\| = \sup_{\mathbf{x}\in\mathbb{R}^n, \mathbf{x}\neq 0} \frac{\|\mathbf{Ax}\|}{\|\mathbf{x}\|} = \sup_{\mathbf{x}\in\mathbb{R}^n, \|\mathbf{x}\|=1} \|\mathbf{Ax}\|$$

- 不难验证：诱导矩阵范数满足非负性、齐次性、三角不等式
- 性质：

（相容性） $\forall \mathbf{A} \in \mathbb{R}^{n\times n}, \mathbf{x} \in \mathbb{R}^n,\ \|\mathbf{Ax}\| \le \|\mathbf{A}\| \cdot \|\mathbf{x}\|$

（可乘性） $\forall \mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n},\ \|\mathbf{AB}\| \le \|\mathbf{A}\| \cdot \|\mathbf{B}\|$

矩阵范数

- **常用矩阵范数**

$$\|\mathbf{A}\|_1 = \max_{1\le j\le n}\sum_{i=1}^{n}|a_{ij}|$$ 列和的最大值

$$\|\mathbf{A}\|_\infty = \max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}|$$ 行和的最大值

$$\|\mathbf{A}\|_2 = \sqrt{\rho(\mathbf{A}^T\mathbf{A})}$$ 谱半径：$\rho(\mathbf{A}) = \max_{1\le r\le n}|\lambda_r|$

$$\|\mathbf{A}\|_F = \left(\sum_{i,j=1}^{n}|a_{ij}|^2\right)^{1/2}$$ 非诱导矩阵范数，但与$\|\mathbf{x}\|_2$相容

- **定理：若λ为矩阵A的特征值，则$|\lambda|\le\|\mathbf{A}\|$**
- **推论：矩阵A的任一（相容）矩阵范数均不小于A的谱半径，即$\rho(\mathbf{A})\le\|\mathbf{A}\|$**
- **定理：对任意$\varepsilon>0$，则存在一个矩阵相容范数$\|\cdot\|$，使得$\|\mathbf{A}\|\le\rho(\mathbf{A})+\varepsilon$，且$\|\mathbf{I}\|=1$**

矩阵范数

- **定义：** 设 $\{A^{(k)}, k=1,2,\ldots\}$ 为 $\mathbb{R}^{n\times n}$ 上的矩阵序列，若存在 $A\in\mathbb{R}^{n\times n}$ 使得
$$\lim_{k\to\infty}\left\|A^{(k)}-A\right\|=0,$$
则称序列序列 $\{A^{(k)}, k=1,2,\ldots\}$ 是收敛的，并称 A 为该序列的极限
- **定理：** $\lim\limits_{k\to\infty} A^k=0$ 的充分必要条件是 $\rho(A)<1$
- **推论：** 若存在相容的矩阵范数使得 $\|A\|<1$，则 $\lim\limits_{k\to\infty} A^k=0$

矩阵的条件数

- 定义：若矩阵 A 非奇异，称

$$\mathrm{Cond}_p(\mathbf{A}) = \|\mathbf{A}\|_p \cdot \|\mathbf{A}^{-1}\|_p$$

为A的条件数，其中 $\|\cdot\|_p$ 表示矩阵的某种范数

- 条件数反应了矩阵对误差的放大率
- 矩阵系数扰动分析：

$$(\mathbf{A}+\delta\mathbf{A})(\mathbf{x}+\delta\mathbf{x}) = \mathbf{b} \Rightarrow \delta\mathbf{A}(\mathbf{x}+\delta\mathbf{x}) = \mathbf{b}-\mathbf{A}\mathbf{x}-\mathbf{A}\delta\mathbf{x} = -\mathbf{A}\delta\mathbf{x}$$

$$\Rightarrow \delta\mathbf{x} = (\mathbf{A}+\delta\mathbf{A})^{-1}(-\delta\mathbf{A}\mathbf{x}) = [\mathbf{A}(\mathbf{I}+\mathbf{A}^{-1}\delta\mathbf{A})]^{-1}(-\delta\mathbf{A}\mathbf{x})$$

$$= (\mathbf{I}+\mathbf{A}^{-1}\delta\mathbf{A})^{-1}\mathbf{A}^{-1}(-\delta\mathbf{A}\mathbf{x})$$

$$\Rightarrow \|\delta\mathbf{x}\| \le \|(\mathbf{I}+\mathbf{A}^{-1}\delta\mathbf{A})^{-1}\| \cdot \|\mathbf{A}^{-1}\| \cdot \|\delta\mathbf{A}\| \cdot \|\mathbf{x}\|$$

$$\Rightarrow \frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \le \frac{\mathrm{Cond}(\mathbf{A})\dfrac{\|\delta\mathbf{A}\|}{\|\mathbf{A}\|}}{1-\mathrm{Cond}(\mathbf{A})\dfrac{\|\delta\mathbf{A}\|}{\|\mathbf{A}\|}}$$

矩阵的条件数

- **右端项系数扰动分析：**

$$\mathbf{A}(\mathbf{x}+\delta\mathbf{x})=\mathbf{b}+\delta\mathbf{b}$$

$$\Rightarrow \mathbf{A}\delta\mathbf{x}=\delta\mathbf{b}$$

$$\Rightarrow \delta\mathbf{x}=\mathbf{A}^{-1}\delta\mathbf{b}$$

$$\Rightarrow \frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|}\le\|\mathbf{A}^{-1}\|\cdot\|\mathbf{A}\|\cdot\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}$$

- **整体系数扰动分析：**

$$(\mathbf{A}+\delta\mathbf{A})(\mathbf{x}+\delta\mathbf{x})=\mathbf{b}+\delta\mathbf{b}$$

$$\Rightarrow \frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|}\le\frac{\mathrm{Cond}(\mathbf{A})}{1-\mathrm{Cond}(\mathbf{A})\dfrac{\|\delta\mathbf{A}\|}{\|\mathbf{A}\|}}\left(\frac{\|\delta\mathbf{A}\|}{\|\mathbf{A}\|}+\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}\right)$$

矩阵的条件数

- 一般判断矩阵是否病态，并不计算 A^{-1}，而由经验得出，譬如：
  - 行列式很大或很小（如某些行、列近似相关）；
  - 元素间相差大数量级，且无规则；
  - 主元消去过程中出现小主元；
  - 特征值相差大数量级。
- 定理：设矩阵 A 非奇异，则

$$\min\left\{\frac{\|\delta\mathbf{A}\|_2}{\|\mathbf{A}\|_2} : \mathbf{A}+\delta\mathbf{A} \text{ is singular}\right\} = \frac{1}{\|\mathbf{A}^{-1}\|_2 \cdot \|\mathbf{A}\|_2} = \frac{1}{\mathrm{Cond}_2(\mathbf{A})}$$

- 病态问题：增加计算精度会改善解的精度，但是不能消除（除非进行符号计算）。目前，即使最出色的数值线代数软件对病态问题也束手无策！

矩阵的条件数

- 例：Hilbert矩阵

$$\mathbf{H}=(h_{ij})_{n\times n},\ h_{ij}=\frac{1}{i+j-1}=\int_0^1 x^{i+j-2}dx$$

$$\mathbf{H}^{-1}=(g_{ij})_{n\times n}, g_{ij}=(-1)^{i+j}(i+j-1)\binom{n+i-1}{n-j}\binom{n+j-1}{n-i}\binom{i+j-2}{i-1}^2$$

$$\mathrm{Cond}(\mathbf{H})=O\left((1+\sqrt{2})^{4n}/\sqrt{n}\right)$$

$$\mathbf{H}=\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} & \frac{1}{8} \\ \frac{1}{5} & \frac{1}{6} & \frac{1}{7} & \frac{1}{8} & \frac{1}{9} \end{bmatrix}$$

一些基本数学定理

- （介值定理）设 $f(x)$ 是区间 $[a,b]$ 上的一个连续函数，那么 $f(x)$ 取到 $f(a)$ 与 $f(b)$ 之间的任何一个值，即如果 y 是 $f(a)$ 与 $f(b)$ 之间的一个数，那么存在一个数 $c\in(a,b)$ 使得 $f(c)=y$

- （中值定理）设 $f(x)$ 是区间 $[a,b]$ 上的一个连续函数，且 $f(x)$ 在 (a,b) 上可微那么在 a 和 b 之间存在一个数 c，使得

$$f'(c)=\frac{f(b)-f(a)}{b-a}$$

一些基本数学定理

- （积分中值定理）设 $f(x)$ 是区间 $[a,b]$ 上的连续函数，$g(x)$ 是可积函数，并且在 $[a,b]$ 上不变号，那么在 $[a,b]$ 内存在一个数 c 使得

$$\int_a^b f(x)g(x)dx = f(c)\int_a^b g(x)dx$$

- （Rolle定理）设 $f(x)$ 是区间 $[a,b]$ 上的可微函数，并且 $f(a)=f(b)$，那么在 a 和 b 之间存在一个数 c，使得

$$f'(c)=0$$

一些基本数学定理

- （带余项的Taylor定理）设 x 和 x_0 是实数，$f(x)$ 在区间 $[x_0, x_1]$（或 $[x, x_0]$）上 $k+1$ 次连续可微，那么在 x 与 x_0 之间存在一个数 c，使得

$$f(x) = f(x_0) + (x - x_0) f'(x_0) + \frac{(x - x_0)^2}{2!} f''(x_0) + \cdots + \frac{(x - x_0)^k}{k!} f^{(k)}(x_0)$$

$$+ \frac{(x - x_0)^{k+1}}{(k+1)!} f^{(k+1)}(c)$$