



目录

- 引言
- Guass消去法
- Guass主元素消去法
- Guass消去法的变形
- 向量和矩阵的范数
- 误差分析



向量和矩阵的范数

- 为了研究线性方程组近似解的误差估计和迭代法的收敛性，需要对 \mathbf{R}^n 中向量（或 $\mathbf{R}^{n \times n}$ 中矩阵）的“大小”引进某种度量——**向量（或矩阵）范数**的概念
 - 向量范数概念是三维Euclid空间中向量长度概念的推广，在数值分析中起着重要作用

- 首先将向量长度概念推广到 \mathbf{R}^n
 - 注意到本节的内容同样可以拓展到 n 维复向量空间 \mathbf{C}^n



数量积与Euclid范数

□ 定义7.1 设

$$\mathbf{x} = (x_1, x_2, \dots, x_n)^T, \mathbf{y} = (y_1, y_2, \dots, y_n)^T \in \mathbf{R}^n$$

向量 \mathbf{x} , \mathbf{y} 的数量积（内积）定义为：

$$(\mathbf{x}, \mathbf{y}) = \mathbf{y}^T \mathbf{x} = \sum_{i=1}^n x_i y_i$$

向量 \mathbf{x} 的Euclid范数定义为：

$$\|\mathbf{x}\|_2 = (\mathbf{x}, \mathbf{x})^{\frac{1}{2}} = \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}} \geq 0$$



数量积与Euclid范数（续）

□ 下述定理可在线性代数书中找到

□ **定理7.10** 设 $\mathbf{x}, \mathbf{y} \in \mathbf{R}^n$ ，则

1. $(\mathbf{x}, \mathbf{x}) = 0$ ，当且仅当 $\mathbf{x} = \mathbf{0}$ 时成立

2. $(\alpha\mathbf{x}, \mathbf{y}) = \alpha(\mathbf{x}, \mathbf{y})$ ， α 为实数

3. $(\mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x})$

4. $(\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}) = (\mathbf{x}_1, \mathbf{y}) + (\mathbf{x}_2, \mathbf{y})$

5. （Cauchy–Schwarz不等式） $|(\mathbf{x}, \mathbf{y})| \leq \|\mathbf{x}\|_2 \cdot \|\mathbf{y}\|_2$ ，等式当且仅当 \mathbf{x} 与 \mathbf{y} 线性相关时成立

6. （三角不等式） $\|\mathbf{x} + \mathbf{y}\|_2 \leq \|\mathbf{x}\|_2 + \|\mathbf{y}\|_2$



向量范数

- 除了Euclid范数之外，还可以用其他办法来度量 \mathbf{R}^n 中向量的“大小”
 - 例如对于 $\mathbf{x} = (x_1, x_2)^T \in \mathbf{R}^n$ ，用一个 \mathbf{x} 的函数 $N(\mathbf{x}) = \max_{i=1,2} |x_i|$ 来度量 \mathbf{x} 的“大小”，而且这种度量方法计算起来比Euclid范数更方便
 - 在许多应用中，对度量 \mathbf{x} 的“大小”的函数 $N(\mathbf{x})$ 都要求是正定的、齐次的且满足三角不等式
- 下面给出向量范数的一般定义



向量范数（续）

□ **定义7.2**（向量的范数）如果向量 $\mathbf{x} \in \mathbf{R}^n$ 的某个实值函数 $N(\mathbf{x}) = \|\mathbf{x}\|$ ，满足条件：

1. $\|\mathbf{x}\| \geq 0$ ， $\|\mathbf{x}\| = 0$ 当且仅当 $\mathbf{x} = \mathbf{0}$ （正定性）
2. $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$ ， $\forall \alpha \in \mathbf{R}$ （齐次性）
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ （三角不等式）

则称 $N(\mathbf{x})$ 是 \mathbf{R}^n 上的一个**向量范数**（或**模**）

□ 由条件3可以推出不等式

$$|\|\mathbf{x}\| - \|\mathbf{y}\|| \leq \|\mathbf{x} - \mathbf{y}\|$$



常用的向量范数

- 向量的 ∞ -范数（最大范数）

$$\|\mathbf{x}\|_{\infty} = \max_{1 \leq i \leq n} |x_i|$$

- 向量的1-范数

$$\|\mathbf{x}\|_1 = |x_1| + |x_2| + \cdots + |x_n|$$

- 向量的2-范数

$$\|\mathbf{x}\|_2 = (\mathbf{x}, \mathbf{x})^{\frac{1}{2}} = (|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2)^{\frac{1}{2}}$$

- 向量的 p -范数，其中 $p \in [1, \infty)$

$$\|\mathbf{x}\|_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{\frac{1}{p}}$$

- 上述三种范数是 p -范数的特殊情况



范数的连续性

- **定义7.3** 设 $\{\mathbf{x}^{(k)}\}$ 为 \mathbf{R}^n 中一向量序列, $\mathbf{x}^* \in \mathbf{R}^n$, 记 $\mathbf{x}^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})^T$, $\mathbf{x}^* = (x_1^*, x_2^*, \dots, x_n^*)^T$ 。如果 $\lim_{k \rightarrow \infty} x_i^{(k)} = x_i^*$, $i = 1, 2, \dots, n$, 则称 $\mathbf{x}^{(k)}$ 收敛于向量 \mathbf{x}^* , 记作

$$\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{x}^*$$

- **定理7.11** ($N(\mathbf{x})$ 的连续性) 设非负函数 $N(\mathbf{x}) = \|\mathbf{x}\|$ 为 \mathbf{R}^n 上任一向量范数, 则 $N(\mathbf{x})$ 是 \mathbf{x} 分量 x_1, x_2, \dots, x_n 的连续函数



范数的等价性

□ **定理7.12**（向量范数的等价性）设 $\|\mathbf{x}\|_s, \|\mathbf{x}\|_t$ 为 \mathbf{R}^n 上向量的任意两种范数，则存在常数 $c_1, c_2 > 0$ ，使得

$$c_1 \|\mathbf{x}\|_s \leq \|\mathbf{x}\|_t \leq c_2 \|\mathbf{x}\|_s, \quad \forall \mathbf{x} \in \mathbf{R}^n$$

- 注意，定理7.12不能推广到无穷维空间
- 从上述定理可知：如果在一种范数意义下向量序列收敛，则在任何一种范数意义下该向量序列亦收敛

□ **定理7.13** $\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{x}^*$ 当且仅当 $k \rightarrow \infty$ 时， $\|\mathbf{x}^{(k)} - \mathbf{x}^*\| \rightarrow 0$ ，其中 $\|\cdot\|$ 为向量的任一种范数



矩阵范数

□ 将向量范数的概念推广到矩阵上

- 由向量的2-范数可以得到 $\mathbf{R}^{n \times n}$ 中矩阵的一种范数

$$F(\mathbf{A}) = \|\mathbf{A}\|_F = \left(\sum_{i,j=1}^n a_{ij}^2 \right)^{\frac{1}{2}}$$

称为 \mathbf{A} 的Frobenius范数

- $\|\mathbf{A}\|_F$ 显然满足正定性、齐次性及三角不等式



矩阵范数（续）

□ **定义7.4**（矩阵范数）如果矩阵 $A \in \mathbf{R}^{n \times n}$ 的某个实值函数 $N(A) = \|A\|$ ，满足条件：

1. $\|A\| \geq 0$ ， $\|A\| = 0$ 当且仅当 $A = \mathbf{0}$ （正定性）
2. $\|cA\| = |c|\|A\|$ ， c 为实数（齐次性）
3. $\|A + B\| \leq \|A\| + \|B\|$ （三角不等式）
4. $\|AB\| \leq \|A\| \|B\|$

则称 $N(A)$ 是 $\mathbf{R}^{n \times n}$ 上的一个**矩阵范数**（或模）

- 最后的条件未必需要
- 上一页定义的 $F(A) = \|A\|_F$ 就是 $\mathbf{R}^{n \times n}$ 上的一个矩阵范数



矩阵的算子范数

- 在大多数与估计有关的问题中，矩阵和向量会同时参与讨论，所以希望引进一种矩阵的范数，它是和向量范数相联系而且相容，即

$$\|A\mathbf{x}\| \leq \|A\| \|\mathbf{x}\|$$

对于任意向量 $\mathbf{x} \in \mathbf{R}^n$ 及 $A \in \mathbf{R}^{n \times n}$ 都成立

- **定义7.5**（矩阵的算子范数）设 $\mathbf{x} \in \mathbf{R}^n$ 及 $A \in \mathbf{R}^{n \times n}$ ，给出一种向量范数 $\|\mathbf{x}\|_v$ （如 $v = 1, 2$ 或 ∞ ），定义 $\mathbf{R}^{n \times n}$ 上的矩阵范数

$$\|A\|_v = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|_v}{\|\mathbf{x}\|_v}$$



矩阵的算子范数（续）

□ 可验证 $\|A\|_v$ 满足定义7.4，所以 $\|A\|_v$ 是 $\mathbf{R}^{n \times n}$ 上的一个矩阵范数，具体而言有如下定理

□ 定理**7.14** 设 $\|x\|_v$ 是 \mathbf{R}^n 上一个向量范数，则 $\|A\|_v$ 是 $\mathbf{R}^{n \times n}$ 上的矩阵范数，且满足相容条件

$$\|Ax\|_v \leq \|A\|_v \|x\|_v$$

□ 显然这种矩阵范数 $\|A\|_v$ 依赖于向量范数 $\|x\|_v$ 的具体含义



常用的矩阵范数

□ **定理7.15** 设 $x \in \mathbf{R}^n$, $A \in \mathbf{R}^{n \times n}$, 则

$$1. \|A\|_{\infty} = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| \quad (\text{行范数})$$

$$2. \|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}| \quad (\text{列范数})$$

$$3. \|A\|_2 = \sqrt{\lambda_{\max}(A^T A)} \quad (2\text{-范数})$$

其中 $\lambda_{\max}(A^T A)$ 表示 $A^T A$ 的最大特征值

- 矩阵的2-范数 $\|A\|_2$ 在计算上不方便, 但是矩阵的2-范数具有许多好的性质, 它在理论上是有用的



特征值与范数

- **定义7.6** 设 $A \in \mathbf{R}^{n \times n}$ 的特征值为 $\lambda_i, i = 1, 2, \dots, n$, 称 $\rho(A) = \max_{1 \leq i \leq n} |\lambda_i|$ 为 A 的**谱半径**
- **定理7.16** (特征值上界) 设 $A \in \mathbf{R}^{n \times n}$, 则 $\rho(A) \leq \|A\|$, 即 A 的谱半径不超过 A 的任何一种算子范数 (对于 $\|A\|_F$ 亦成立)
- **定理7.17** 如果 $A \in \mathbf{R}^{n \times n}$ 为对称矩阵, 则 $\|A\|_2 = \rho(A)$
- **定理7.18** 如果 $\|B\| < 1$, 则 $I \pm B$ 为非奇异矩阵, 且 $\|(I \pm B)^{-1}\| \leq \frac{1}{1 - \|B\|}$, $\|\cdot\|$ 为算子范数



定理7.18证明

- 用反证法，若 $I \pm B$ 是奇异矩阵，则行线性相关，则 $(I \pm B)x = 0$ 有非零解，即存在 $x_0 \neq 0$ 使

$$Bx_0 = x_0$$

- 注意到

$$\|B\| = \max_{x \neq 0} \frac{\|Bx\|}{\|x\|} \geq \frac{\|Bx_0\|}{\|x_0\|} = 1$$

与假设矛盾，因此 $I \pm B$ 为非奇异矩阵

- 由 $(I \pm B)(I \pm B)^{-1} = I$ ，有

$$(I \pm B)^{-1} = I \mp B(I \pm B)^{-1}$$

$$\Rightarrow \|(I \pm B)^{-1}\| \leq 1 + \|B\| \|(I \pm B)^{-1}\| \Rightarrow \|(I \pm B)^{-1}\| \leq \frac{1}{1 - \|B\|}$$



目录

- 引言
- Guass消去法
- Guass主元素消去法
- Guass消去法的变形
- 向量和矩阵的范数
- 误差分析



微小误差对解的影响

□ 考虑线性方程组 $Ax = b$ ，其中 A 为非奇异矩阵， x 为方程组的精确解

■ 由于 A （或 b ）中的元素是测量得到的，或者是计算的结果，所以 A （或 b ）常带有某些观测误差或者包含有舍入误差

□ 在处理实际问题 $(A + \delta A)x = b + \delta b$ 时，通常要考虑 A （或 b ）的微小误差对解的影响

■ 考虑估计 $x - y$ ，其中

$$Ax = b$$

$$(A + \delta A)y = b + \delta b$$



例7.8

□ 设有方程组 $A\mathbf{x} = \mathbf{b}$ ，其精确解为 $\mathbf{x} = (2,0)^T$

$$\begin{bmatrix} 1 & 1 \\ 1 & 1.0001 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

■ 现在考虑常数项的微小变化对方程组解的影响，

即考察
$$\begin{bmatrix} 1 & 1 \\ 1 & 1.0001 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2.0001 \end{bmatrix}$$

■ 可描述为 $A(\mathbf{x} + \delta\mathbf{x}) = \mathbf{b} + \delta\mathbf{b}$ ， $\delta\mathbf{b} = (0,0.0001)^T$

，易得
$$\mathbf{x} + \delta\mathbf{x} = (1,1)^T$$

■ 可以看到，常数项 \mathbf{b} 的第二个分量只有 $1/10000$ 的微小变化，方程组的解却变化很大，这样的方程组称为病态方程组



病态方程组

- **定义7.7** 设如果矩阵 **A** 或常数项 **b** 的微小变化，引起方程组 **$Ax = b$** 解的巨大变化，则称此方程组为**病态方程组**，矩阵 **A** 称为**病态矩阵**（相对于方程组而言），否则称方程组为**良态方程组**， **A** 称为**良态矩阵**
- 应该注意，矩阵的病态性质是**矩阵本身的特性**，我们希望可以定量地刻画矩阵的病态程度，且无需通过解具体的方程组



微小误差影响的定量分析

□ 设方程组 $A\mathbf{x} = \mathbf{b}$ ，其中 A 为非奇异阵， \mathbf{x} 为方程组的精确解

1. 设 A 是精确的， \mathbf{b} 有微小误差 $\delta\mathbf{b}$ ，解为 $\mathbf{x} + \delta\mathbf{x}$

■ 首先

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

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

■ 其次

$$A\mathbf{x} = \mathbf{b} \Rightarrow \|\mathbf{b}\| = \|A\mathbf{x}\| \leq \|A\| \|\mathbf{x}\| \Rightarrow \frac{1}{\|\mathbf{x}\|} \leq \frac{\|A\|}{\|\mathbf{b}\|}$$

其中假设 $\mathbf{b} \neq \mathbf{0}$



微小误差影响的定量分析（续）

□ 设方程组 $\mathbf{Ax} = \mathbf{b}$ ，其中 \mathbf{A} 为非奇异阵， \mathbf{x} 为方程组的精确解

1. 设 \mathbf{A} 是精确的， \mathbf{b} 有微小误差 $\delta\mathbf{b}$ ，解为 $\mathbf{x} + \delta\mathbf{x}$

■ 结合

$$\|\delta\mathbf{x}\| \leq \|\mathbf{A}^{-1}\| \|\delta\mathbf{b}\| \quad \frac{1}{\|\mathbf{x}\|} \leq \frac{\|\mathbf{A}\|}{\|\mathbf{b}\|}$$

□ 定理 7.19 设 \mathbf{A} 为非奇异阵， $\mathbf{Ax} = \mathbf{b} \neq \mathbf{0}$ ，且 $\mathbf{A}(\mathbf{x} + \delta\mathbf{x}) = \mathbf{b} + \delta\mathbf{b}$ ，则

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq \|\mathbf{A}^{-1}\| \|\mathbf{A}\| \frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}$$

■ 常数项 \mathbf{b} 的相对误差在解中可能放大 $\|\mathbf{A}^{-1}\| \|\mathbf{A}\|$ 倍



微小误差影响的定量分析（续）

□ 设方程组 $\mathbf{Ax} = \mathbf{b}$ ，其中 \mathbf{A} 为非奇异阵， \mathbf{x} 为方程组的精确解

2. 设 \mathbf{b} 是精确的， \mathbf{A} 有微小误差 $\delta\mathbf{A}$ ，解为 $\mathbf{x} + \delta\mathbf{x}$

■ 首先

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

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

■ 如果 $\delta\mathbf{A}$ 不受限制， $\mathbf{A} + \delta\mathbf{A}$ 可能奇异，注意到

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

■ 由定理 7.18 知，当 $\|\mathbf{A}^{-1}\delta\mathbf{A}\| < 1$ 时， $(\mathbf{I} + \mathbf{A}^{-1}\delta\mathbf{A})^{-1}$ 存在



微小误差影响的定量分析（续）

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

$$\begin{aligned}\Rightarrow \|\delta \mathbf{x}\| &\leq \|(I + A^{-1}\delta A)^{-1}\| \|A^{-1}\| \|\delta A\| \|\mathbf{x}\| \\ &\leq \frac{\|A^{-1}\| \|\delta A\| \|\mathbf{x}\|}{1 - \|A^{-1}\delta A\|} \leq \frac{\|A^{-1}\| \|\delta A\| \|\mathbf{x}\|}{1 - \|A^{-1}\| \|\delta A\|}\end{aligned}$$

其中假设 $\|A^{-1}\| \|\delta A\| < 1$

□ **定理7.20** 设 A 为非奇异阵, $A\mathbf{x} = \mathbf{b} \neq \mathbf{0}$, 且 $(A + \delta A)(\mathbf{x} + \delta \mathbf{x}) = \mathbf{b}$, 如果 $\|A^{-1}\| \|\delta A\| < 1$

$$\frac{\|\delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \frac{\|A^{-1}\| \|\delta A\|}{1 - \|A^{-1}\| \|\delta A\|} = \frac{\|A^{-1}\| \|A\| \frac{\|\delta A\|}{\|A\|}}{1 - \|A^{-1}\| \|A\| \frac{\|\delta A\|}{\|A\|}}$$



微小误差影响的定量分析（续）

- 若 $\delta\mathbf{A}$ 充分小，且 $\|\mathbf{A}^{-1}\|\|\delta\mathbf{A}\| < 1$ ，则矩阵 \mathbf{A} 的相对误差 $\frac{\|\delta\mathbf{A}\|}{\|\mathbf{A}\|}$ 在解中可能放大 $\|\mathbf{A}^{-1}\|\|\mathbf{A}\|$ 倍
- 从上面的讨论可知，量 $\|\mathbf{A}^{-1}\|\|\mathbf{A}\|$ 愈小，由 \mathbf{A} （或 \mathbf{b} ）的相对误差引起的解的相对误差就愈小；量 $\|\mathbf{A}^{-1}\|\|\mathbf{A}\|$ 愈大，该解的相对误差就可能愈大
- 量 $\|\mathbf{A}^{-1}\|\|\mathbf{A}\|$ 实际上刻画了解对原始数据变化的灵敏度，即刻画了方程组的病态程度



条件数

- **定义7.8** 设 A 为非奇异矩阵，称 $\text{cond}(A)_v = \|A^{-1}\|_v \|A\|_v$ ($v = 1, 2$ 或 ∞) 为矩阵 A 的**条件数**
- 矩阵的条件数与范数有关
 - 当 A 的条件数相对地大时，即 $\text{cond}(A) \gg 1$ 时，式 $Ax = b$ 是病态的（即 A 是病态矩阵，或者说 A 是坏条件的）
 - 当 A 的条件数相对地小时，式 $Ax = b$ 是良态的（或者说 A 是好条件的）
 - A 的条件数愈大，方程组的病态程度愈严重，也就愈难得到方程组的比较准确的解



常用的条件数

1. $\text{cond}(\mathbf{A})_\infty = \|\mathbf{A}^{-1}\|_\infty \|\mathbf{A}\|_\infty$

2. \mathbf{A} 的谱条件数

$$\text{cond}(\mathbf{A})_2 = \|\mathbf{A}^{-1}\|_2 \|\mathbf{A}\|_2 = \sqrt{\frac{\lambda_{\max}(\mathbf{A}^T \mathbf{A})}{\lambda_{\min}(\mathbf{A}^T \mathbf{A})}}$$

- 当 \mathbf{A} 为对称阵时, $\text{cond}(\mathbf{A})_2 = \frac{|\lambda_1|}{|\lambda_n|}$, 其中 λ_1 和 λ_n 为 \mathbf{A} 的绝对值最大和绝对值最小的特征值
- 这是最常见的条件数



条件数的性质

1. 任何非奇异矩阵 \mathbf{A} ，都有 $\text{cond}(\mathbf{A})_v \geq 1$

$$\text{cond}(\mathbf{A})_v = \|\mathbf{A}^{-1}\|_v \|\mathbf{A}\|_v \geq \|\mathbf{A}^{-1}\mathbf{A}\|_v = \|\mathbf{I}\|_v = 1$$

2. \mathbf{A} 为非奇异矩阵， c 为不等于零的常数，则

$$\text{cond}(c\mathbf{A})_v = \text{cond}(\mathbf{A})_v$$

3. 如果 \mathbf{A} 为正交矩阵，则 $\text{cond}(\mathbf{A})_2 = 1$ ；如果 \mathbf{A} 为非奇异矩阵， \mathbf{R} 为正交矩阵，则

$$\text{cond}(\mathbf{R}\mathbf{A})_2 = \text{cond}(\mathbf{A}\mathbf{R})_2 = \text{cond}(\mathbf{A})_2$$



例7.9

□ 已知Hilbert矩阵 H_n ，计算 H_3 的条件数

$$H_n = \begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \cdots & \frac{1}{2n-1} \end{bmatrix}$$

■ 解

$$H_3 = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}, \quad H_3^{-1} = \begin{bmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{bmatrix}$$



例7.9 (续)

- 计算 \mathbf{H}_3 条件数 $\text{cond}(\mathbf{H}_3)_\infty$

$$\|\mathbf{H}_3\|_\infty = \frac{11}{6}, \quad \|\mathbf{H}_3^{-1}\|_\infty = 408$$

所以 $\text{cond}(\mathbf{H}_3)_\infty = 748$

- 同样可计算 $\text{cond}(\mathbf{H}_6)_\infty = 2.9 \times 10^6$
- 一般当 n 愈大时, \mathbf{H}_n 的病态愈严重

□ 考虑方程组

$$\mathbf{H}_3 \mathbf{x} = \left(\frac{11}{6}, \frac{13}{12}, \frac{47}{60} \right)^T = \mathbf{b}$$

- 精确解为 $\mathbf{x} = (1, 1, 1)^T$



例7.9 (续)

- 设 \mathbf{H}_3 及 \mathbf{b} 有微小误差（取三位有效数字），

$$\begin{bmatrix} 1.00 & 0.500 & 0.333 \\ 0.500 & 0.333 & 0.250 \\ 0.333 & 0.250 & 0.200 \end{bmatrix} \begin{bmatrix} x_1 + \delta x_1 \\ x_2 + \delta x_2 \\ x_3 + \delta x_3 \end{bmatrix} = \begin{bmatrix} 1.83 \\ 1.08 \\ 0.783 \end{bmatrix} \quad (7.6.8)$$

简记作 $(\mathbf{H}_3 + \delta\mathbf{H}_3)(\mathbf{x} + \delta\mathbf{x}) = \mathbf{b} + \delta\mathbf{b}$

- 式(7.6.8)的解为

$$\mathbf{x} + \delta\mathbf{x} = (1.089512538, 0.487967062, 1.491002798)^T$$

- 可得 $\delta\mathbf{x} = (0.0895, -0.5120, 0.4910)^T$



例7.9 (续)

■ 于是

$$\frac{\|\delta \mathbf{H}_3\|_\infty}{\|\mathbf{H}_3\|_\infty} \approx 0.18 \times 10^{-3} < 0.02\%,$$

$$\frac{\|\delta \mathbf{b}\|_\infty}{\|\mathbf{b}\|_\infty} \approx 0.182\%, \quad \frac{\|\delta \mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} \approx 51.2\%$$

■ \mathbf{H}_3 与 \mathbf{b} 相对误差不超过0.2%，而引起解的误差超过50%



判断矩阵是否病态

□ 由上面讨论可知，要判别一个矩阵是否病态，需要计算条件数 $\text{cond}(\mathbf{A}) = \|\mathbf{A}^{-1}\| \|\mathbf{A}\|$ ，而计算 $\|\mathbf{A}^{-1}\|$ 是比较困难的，那么在实际计算中如何发现病态情况呢？

1. 如果在 \mathbf{A} 的三角约化时（尤其是用主元素消去法解方程组时）出现小主元，那么对大多数矩阵来说， \mathbf{A} 是病态矩阵

■ 例如用选主元的直接三角分解法解(7.6.8)

$$\begin{bmatrix} 1.00 & 0.500 & 0.333 \\ 0.500 & 0.333 & 0.250 \\ 0.333 & 0.250 & 0.200 \end{bmatrix} \begin{bmatrix} x_1 + \delta x_1 \\ x_2 + \delta x_2 \\ x_3 + \delta x_3 \end{bmatrix} = \begin{bmatrix} 1.83 \\ 1.08 \\ 0.783 \end{bmatrix} \quad (7.6.8)$$



判断矩阵是否病态（续）

■ 得到

$$I_{23}(\mathbf{H}_3 + \delta\mathbf{H}_3) = \begin{bmatrix} 1 & & \\ 0.333 & 1 & \\ 0.500 & 0.994 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0.5000 & 0.3330 \\ & 0.0835 & 0.0891 \\ & & -0.00507 \end{bmatrix}$$

2. 如果 \mathbf{A} 的最大特征值和最小特征值之比（按绝对值）是大的，则 \mathbf{A} 是病态的，即 $\frac{|\lambda_1|}{|\lambda_n|}$ 是大的，其中 λ_1 和 λ_n 为 \mathbf{A} 的绝对值最大和绝对值最小的特征值

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n| > 0$$

定理**7.16** $\Rightarrow |\lambda_1| \leq \|\mathbf{A}\|, \frac{1}{|\lambda_n|} \leq \|\mathbf{A}^{-1}\| \Rightarrow \text{cond}(\mathbf{A}) \geq \frac{|\lambda_1|}{|\lambda_n|} \gg 1$



判断矩阵是否病态（续）

3. 如果系数矩阵的行列式值相对来说很小，或系数矩阵某些行近似线性相关，则 \mathbf{A} 可能是病态的
 4. 如果系数矩阵 \mathbf{A} 元素间数量级相差很大，并且无一定规则，则 \mathbf{A} 可能是病态的
- 病态问题通常不能用选主元素的消去法来解决



病态问题的求解

1. 一般采用高精度的算术运算（采用双倍字长进行运算）

2. 预处理方法，即将求解 $\mathbf{Ax} = \mathbf{b}$ 的问题转化为求解一等价方程组

$$\begin{cases} \mathbf{PAQy} = \mathbf{Pb} \\ \mathbf{y} = \mathbf{Q}^{-1}\mathbf{x} \end{cases}$$

- 通过选择非奇异矩阵 \mathbf{P} 与 \mathbf{Q} ，使 $\text{cond}(\mathbf{PAQ}) < \text{cond}(\mathbf{A})$
- 一般选择 \mathbf{P} 与 \mathbf{Q} 为对角阵或者三角阵



病态问题的求解（续）

3. 当矩阵 \mathbf{A} 的元素大小不均时，在 \mathbf{A} 的行（或列）中引进适当的比例因子（使矩阵 \mathbf{A} 的所有行或列按 ∞ -范数大体上有相同的长度，使 \mathbf{A} 的系数均衡）
- 对 \mathbf{A} 的条件数是有影响的
 - 但这种方法不能保证 \mathbf{A} 的条件数一定得到改善

□ 例7.10 设 $\mathbf{A} = \begin{bmatrix} 1 & 10^4 \\ 1 & 1 \end{bmatrix}$ ，方程组 $\mathbf{Ax} = \mathbf{b}$ 为

$$\begin{bmatrix} 1 & 10^4 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 10^4 \\ 2 \end{bmatrix}, \text{ 计算 } \text{cond}(\mathbf{A})_{\infty} \text{ 和解}$$



例7.10 (续)

■ 由
$$\mathbf{A} = \begin{bmatrix} 1 & 10^4 \\ 1 & 1 \end{bmatrix}$$

可得
$$\mathbf{A}^{-1} = \frac{1}{10^4 - 1} \begin{bmatrix} -1 & 10^4 \\ 1 & -1 \end{bmatrix}$$

$$\text{cond}(\mathbf{A})_{\infty} = \frac{(1 + 10^4)^2}{10^4 - 1} \approx 10^4$$

- 用列主元素消去法解 $\mathbf{Ax} = \mathbf{b}$ 时 (计算到三位数字), 得

$$(\mathbf{A}, \mathbf{b}) \rightarrow \begin{bmatrix} 1 & 10^4 & \vdots & 10^4 \\ 0 & -10^4 & \vdots & -10^4 \end{bmatrix}$$

于是得到很坏的结果: $x_1 = 0, x_2 = 1$



例7.10 (续)

- 在**A**的第一行引进比例因子, 如用 $s_1 = \max_{1 \leq i \leq 2} |a_{1i}| = 10^4$ 除第一个方程式, 得 $\mathbf{A}'\mathbf{x} = \mathbf{b}'$, 即

$$\begin{bmatrix} 10^{-4} & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad (\mathbf{A}')^{-1} = \frac{1}{1 - 10^{-4}} \begin{bmatrix} -1 & 1 \\ 1 & -10^{-4} \end{bmatrix}$$

$$\text{cond}(\mathbf{A}')_{\infty} = \frac{1}{1 - 10^{-4}} \approx 1 \quad \text{大幅下降}$$

- 用列主元素消去法解 $\mathbf{A}'\mathbf{x} = \mathbf{b}'$ 时, 得

$$(\mathbf{A}', \mathbf{b}') \rightarrow \begin{bmatrix} 1 & 1 & \vdots & 2 \\ 10^{-4} & 1 & \vdots & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & \vdots & 2 \\ 0 & 1 & \vdots & 1 \end{bmatrix}$$

从而得到较好的结果: $x_1 = 1, x_2 = 1$



事后误差估计

- 设 $\bar{\mathbf{x}}$ 为方程组 $\mathbf{Ax} = \mathbf{b}$ 的近似解，于是可计算 $\bar{\mathbf{x}}$ 的剩余向量 $\mathbf{r} = \mathbf{b} - \mathbf{A}\bar{\mathbf{x}}$ ，当 \mathbf{r} 很小时， $\bar{\mathbf{x}}$ 是否为 $\mathbf{Ax} = \mathbf{b}$ 一个较好的近似解呢？
- 定理**7.21**（事后误差估计） 设 \mathbf{A} 为非奇异阵， $\mathbf{Ax} = \mathbf{b} \neq \mathbf{0}$ ， $\bar{\mathbf{x}}$ 是方程组的近似解， $\mathbf{r} = \mathbf{b} - \mathbf{A}\bar{\mathbf{x}}$ ，则

$$\frac{\|\mathbf{x} - \bar{\mathbf{x}}\|}{\|\mathbf{x}\|} \leq \text{cond}(\mathbf{A}) \frac{\|\mathbf{r}\|}{\|\mathbf{b}\|}$$



定理7.21 (证明)

■ 由
$$A(\mathbf{x} - \bar{\mathbf{x}}) = \mathbf{b} - A\bar{\mathbf{x}} = \mathbf{r}$$

得
$$\mathbf{x} - \bar{\mathbf{x}} = A^{-1}\mathbf{r} \Rightarrow \|\mathbf{x} - \bar{\mathbf{x}}\| \leq \|A^{-1}\| \|\mathbf{r}\|$$

■ 又有
$$\|\mathbf{b}\| = \|A\mathbf{x}\| \leq \|A\| \|\mathbf{x}\|$$

得
$$\frac{1}{\|\mathbf{x}\|} \leq \frac{\|A\|}{\|\mathbf{b}\|}$$

■ 因此
$$\frac{\|\mathbf{x} - \bar{\mathbf{x}}\|}{\|\mathbf{x}\|} \leq \frac{\|A\| \|A^{-1}\| \|\mathbf{r}\|}{\|\mathbf{b}\|} = \text{cond}(A) \frac{\|\mathbf{r}\|}{\|\mathbf{b}\|}$$



舍入误差

- 在复杂的计算中，由浮点运算而引进的舍入误差可能积累而影响答案，因此对任何算法都需要进行舍入误差分析，看其是否过度影响所得的结果

- 设 \bar{x} 为选主元素Gauss消去法解 $Ax = b$ 的计算解， x 为精确解
 - 若要直接计算每一步舍入误差对解的影响来获得界的估计 $\|x - \bar{x}\|$ ，是非常困难的



舍入误差（续）

- Wilkinson等人提出了“向后误差分析方法”，基本思想是把舍入误差对解的影响归为原始数据变化对解的影响，即 $\bar{\mathbf{x}}$ 是扰动方程组 $(\mathbf{A} + \delta\mathbf{A})\mathbf{x} = \mathbf{b}$ 的精确解
 - 具体定理可以查看教材

- 计算解 $\bar{\mathbf{x}}$ 的相对误差限依赖于 $\text{cond}(\mathbf{A})_\infty$ ，元素的增长因子、方程组阶数、计算机字长等



总结

□ Gauss消去法

- 消元过程、回代过程、能够成功的充分条件
- Gauss消去法的矩阵描述、矩阵的LU分解
- Gauss消去法的计算量

□ Gauss主元素消去法

- 完全主元素消去法、列主元素消去法
- 列主元素的三角分解定理
- Gauss-Jordan消去法、矩阵求逆



总结（续）

□ Gauss消去法的变形

- 不选主元的三角分解法、选主元的三角分解法
- 平方根法、对称（正定）矩阵的三角分解
- 数值稳定性、避免开方的分解
- 追赶法、对角占优的矩阵三角分解、分解的性质

□ 向量和矩阵的范数

- 数量积与Euclid范数、范数、连续性和等价性
- 矩阵范数、算子范数、特征值、谱半径

□ 误差分析

- 病态方程组、误差分析、条件数、舍入误差