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

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

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

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

向量$\boldsymbol{x}$的Euclid范数定义为：

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

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$\boldsymbol{x}$，$\boldsymbol{y} \in \mathbf{R}^n$，则
  1. $(\boldsymbol{x}, \boldsymbol{x}) \geq 0$，当且仅当$\boldsymbol{x} = \mathbf{0}$时成立
  2. $(\alpha\boldsymbol{x}, \boldsymbol{y}) = \alpha(\boldsymbol{x}, \boldsymbol{y})$，$\alpha$为实数
  3. $(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{y}, \boldsymbol{x})$
  4. $(\boldsymbol{x}_1 + \boldsymbol{x}_2, \boldsymbol{y}) = (\boldsymbol{x}_1, \boldsymbol{y}) + (\boldsymbol{x}_2, \boldsymbol{y})$
  5. （Cauchy−Schwarz不等式）$|(\boldsymbol{x}, \boldsymbol{y})| \leq \|\boldsymbol{x}\|_2 \cdot \|\boldsymbol{y}\|_2$，等式当且仅当$\boldsymbol{x}$与$\boldsymbol{y}$线性相关时成立
  6. （三角不等式）$\|\boldsymbol{x} + \boldsymbol{y}\|_2 \leq \|\boldsymbol{x}\|_2 + \|\boldsymbol{y}\|_2$

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$\boldsymbol{x}=(x_1,x_2)^{\mathrm{T}}\in\mathbf{R}^n$，用一个$\boldsymbol{x}$的函数$N(\boldsymbol{x})=\max\limits_{i=1,2}|x_i|$来度量$\boldsymbol{x}$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$\boldsymbol{x}$的“大小”的函数$N(\boldsymbol{x})$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$\boldsymbol{x} \in \mathbf{R}^n$的某个实值函数$N(\boldsymbol{x}) = \|\boldsymbol{x}\|$，满足条件：
  1. $\|\boldsymbol{x}\| \geq 0$，$\|\boldsymbol{x}\| = 0$当且仅当$\boldsymbol{x} = \mathbf{0}$ （正定性）
  2. $\|\alpha\boldsymbol{x}\| = |\alpha|\|\boldsymbol{x}\|$，$\forall\alpha \in \mathbf{R}$ （齐次性）
  3. $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ （三角不等式）

  则称$N(\boldsymbol{x})$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

$$\big|\|\boldsymbol{x}\| - \|\boldsymbol{y}\|\big| \leq \|\boldsymbol{x} - \boldsymbol{y}\|$$

常用的向量范数

- 向量的∞-范数（最大范数）

$$\|\boldsymbol{x}\|_\infty = \max_{1\le i\le n} |x_i|$$

- 向量的1-范数

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

- 向量的2-范数

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

- 向量的p-范数，其中$p \in [1, \infty)$

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

  - 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$\{x^{(k)}\}$为$\mathbf{R}^n$中一向量序列，$\boldsymbol{x}^* \in \mathbf{R}^n$，记$\boldsymbol{x}^{(k)} = (x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)})^{\mathrm{T}}$， $\boldsymbol{x}^* = (x_1^*, x_2^*, \cdots, x_n^*)^{\mathrm{T}}$。如果$\lim\limits_{k\to\infty} \boldsymbol{x}_i^{(k)} = \boldsymbol{x}_i^*$，$i = 1,2,\cdots,n$，则称$\boldsymbol{x}^{(k)}$收敛于向量$\boldsymbol{x}^*$，记作

$$\lim_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{x}^*$$

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

范数的等价性

- **定理7.12**（向量范数的等价性）设$\|\boldsymbol{x}\|_s$，$\|\boldsymbol{x}\|_t$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

$$c_1\|\boldsymbol{x}\|_s \le \|\boldsymbol{x}\|_t \le c_2\|\boldsymbol{x}\|_s, \qquad \forall \boldsymbol{x} \in \mathbf{R}^n$$

  - 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** $\lim\limits_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{x}^*$当且仅当$k \to \infty$时，$\left\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^*\right\| \to 0$，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
  - 由向量的的2-范数可以得到$\mathbf{R}^{n\times n}$中矩阵的一种范数

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

称为$\boldsymbol{A}$的Frobenius范数

  - $\|\boldsymbol{A}\|_F$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$\boldsymbol{A} \in \mathbf{R}^{n\times n}$的某个实值函数$N(\boldsymbol{A}) = \|\boldsymbol{A}\|$，满足条件：
  1. $\|\boldsymbol{A}\| \geq 0$，　$\|\boldsymbol{A}\| = 0$当且仅当$\boldsymbol{A} = \boldsymbol{0}$（正定性）
  2. $\|c\boldsymbol{A}\| = |c|\|\boldsymbol{A}\|$，$c$为实数　（齐次性）
  3. $\|\boldsymbol{A} + \boldsymbol{B}\| \leq \|\boldsymbol{A}\| + \|\boldsymbol{B}\|$　（三角不等式）
  4. $\|\boldsymbol{AB}\| \leq \|\boldsymbol{A}\|\,\|\boldsymbol{B}\|$

  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$F(\boldsymbol{A}) = \|\boldsymbol{A}\|_F$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

$$\|\boldsymbol{Ax}\| \leq \|\boldsymbol{A}\|\,\|\boldsymbol{x}\|$$

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

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

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

矩阵的算子范数（续）

- 可验证$\|\boldsymbol{A}\|_v$满足定义7.4，所以$\|\boldsymbol{A}\|_v$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$\|\boldsymbol{x}\|_v$是$\mathbf{R}^n$上一个向量范数，则$\|\boldsymbol{A}\|_v$是$\mathbf{R}^{n\times n}$上的矩阵范数，且满足相容条件

$$\|\boldsymbol{A}\boldsymbol{x}\|_v \leq \|\boldsymbol{A}\|_v\|\boldsymbol{x}\|_v$$

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

常用的矩阵范数

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

1. $\|\boldsymbol{A}\|_\infty=\max\limits_{1\le i\le n}\sum\limits_{j=1}^{n}|a_{ij}|$ （行范数）

2. $\|\boldsymbol{A}\|_1=\max\limits_{1\le j\le n}\sum\limits_{i=1}^{n}|a_{ij}|$ （列范数）

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

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

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

定理7.18证明

- 用反证法， 若$\boldsymbol{I} \pm \boldsymbol{B}$是奇异矩阵，则行线性相关，则$(\boldsymbol{I} \pm \boldsymbol{B})\boldsymbol{x} = \boldsymbol{0}$有非零解，即存在$\boldsymbol{x}_0 \neq \boldsymbol{0}$使
$$\boldsymbol{B}\boldsymbol{x}_0 = \boldsymbol{x}_0$$
- 注意到
$$\|\boldsymbol{B}\| = \max_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{\|\boldsymbol{B}\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \geq \frac{\|\boldsymbol{B}\boldsymbol{x}_0\|}{\|\boldsymbol{x}_0\|} = 1$$
与假设矛盾，因此$\boldsymbol{I} \pm \boldsymbol{B}$为非奇异矩阵

- 由$(\boldsymbol{I} \pm \boldsymbol{B})(\boldsymbol{I} \pm \boldsymbol{B})^{-1} = \boldsymbol{I}$，有
$$(\boldsymbol{I} \pm \boldsymbol{B})^{-1} = \boldsymbol{I} \mp \boldsymbol{B}(\boldsymbol{I} \pm \boldsymbol{B})^{-1}$$
$$\Rightarrow \left\|(\boldsymbol{I} \pm \boldsymbol{B})^{-1}\right\| \leq 1 + \|\boldsymbol{B}\|\left\|(\boldsymbol{I} \pm \boldsymbol{B})^{-1}\right\| \;\; \Rightarrow \left\|(\boldsymbol{I} \pm \boldsymbol{B})^{-1}\right\| \leq \frac{1}{1 - \|\boldsymbol{B}\|}$$

目录

微小误差对解的影响

- 考虑线性方程组$\boldsymbol{Ax}=\boldsymbol{b}$，其中$\boldsymbol{A}$为非奇异矩阵，$\boldsymbol{x}$为方程组的精确解
  - 由于$\boldsymbol{A}$（或$\boldsymbol{b}$）中的元素是测量得到的，或者是计算的结果，所以$\boldsymbol{A}$（或$\boldsymbol{b}$）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$(\boldsymbol{A}+\delta\boldsymbol{A})\boldsymbol{x}=\boldsymbol{b}+\delta\boldsymbol{b}$时，通常要考虑$\boldsymbol{A}$（或$\boldsymbol{b}$）的微小误差对解的影响
  - 考虑估计$\boldsymbol{x}-\boldsymbol{y}$，其中

$$\boldsymbol{Ax}=\boldsymbol{b}$$

$$(\boldsymbol{A}+\delta\boldsymbol{A})\boldsymbol{y}=\boldsymbol{b}+\delta\boldsymbol{b}$$

例7.8

- 设有方程组$\boldsymbol{Ax}=\boldsymbol{b}$，其精确解为$\boldsymbol{x}=(2,0)^{\mathrm{T}}$

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

  - 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$\boldsymbol{A}(\boldsymbol{x}+\delta\boldsymbol{x})=\boldsymbol{b}+\delta\boldsymbol{b}$，$\delta\boldsymbol{b}=(0,0.0001)^{\mathrm{T}}$，易得 $$\boldsymbol{x}+\delta\boldsymbol{x}=(1,1)^{\mathrm{T}}$$
  - 可以看到，常数项$\boldsymbol{b}$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$\boldsymbol{A}$或常数项$\boldsymbol{b}$的微小变化，引起方程组$\boldsymbol{Ax}=\boldsymbol{b}$解的巨大变化，则称此方程组为病态方程组，矩阵$\boldsymbol{A}$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，$\boldsymbol{A}$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$\boldsymbol{Ax}=\boldsymbol{b}$，其中$\boldsymbol{A}$为非奇异阵，$\boldsymbol{x}$为方程组的精确解

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

- 首先

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

$$\Rightarrow \|\delta\boldsymbol{x}\|=\left\|\boldsymbol{A}^{-1}\delta\boldsymbol{b}\right\|\le\left\|\boldsymbol{A}^{-1}\right\|\|\delta\boldsymbol{b}\|$$

- 其次

$$\boldsymbol{Ax}=\boldsymbol{b} \quad \Rightarrow \|\boldsymbol{b}\|=\|\boldsymbol{Ax}\|\le\|\boldsymbol{A}\|\|\boldsymbol{x}\| \quad \Rightarrow \frac{1}{\|\boldsymbol{x}\|}\le\frac{\|\boldsymbol{A}\|}{\|\boldsymbol{b}\|}$$

其中假设$\boldsymbol{b}\neq\boldsymbol{0}$

微小误差影响的定量分析（续）

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

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

$$\|\delta\boldsymbol{x}\| \le \|\boldsymbol{A}^{-1}\|\|\delta\boldsymbol{b}\| \qquad \frac{1}{\|\boldsymbol{x}\|} \le \frac{\|\boldsymbol{A}\|}{\|\boldsymbol{b}\|}$$

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

$$\frac{\|\delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \le \|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\| \frac{\|\delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|}$$

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

微小误差影响的定量分析（续）

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

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

- 首先

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

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

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

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

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

微小误差影响的定量分析（续）

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

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

$$\le \frac{\left\|A^{-1}\right\|\|\delta A\|\|x\|}{1 - \left\|A^{-1}\delta A\right\|} \le \frac{\left\|A^{-1}\right\|\|\delta A\|\|x\|}{1 - \left\|A^{-1}\right\|\|\delta A\|}$$

其中假设$\left\|A^{-1}\right\|\|\delta A\| < 1$

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

$$\frac{\|\delta x\|}{\|x\|} \le \frac{\left\|A^{-1}\right\|\|\delta A\|}{1 - \left\|A^{-1}\right\|\|\delta A\|} = \frac{\left\|A^{-1}\right\| \|A\| \dfrac{\|\delta A\|}{\|A\|}}{1 - \left\|A^{-1}\right\| \|A\| \dfrac{\|\delta A\|}{\|A\|}}$$

微小误差影响的定量分析（续）

- 若$\delta\boldsymbol{A}$充分小，且$\|\boldsymbol{A}^{-1}\|\|\delta\boldsymbol{A}\| < 1$，则矩阵$\boldsymbol{A}$的相对误差$\frac{\|\delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|}$在解中可能放大$\|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|$倍

- 从上面的讨论可知，量$\|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|$愈小，由$\boldsymbol{A}$（或$\boldsymbol{b}$）的相对误差引起的解的相对误差就愈小；量$\|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|$愈大，该解的相对误差就可能愈大
  - 量$\|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|$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$\boldsymbol{A}$为非奇异矩阵，称$\mathrm{cond}(\boldsymbol{A})_v = \left\|\boldsymbol{A}^{-1}\right\|_v \|\boldsymbol{A}\|_v$（$v = 1,2$或$\infty$）为矩阵$\boldsymbol{A}$的条件数
  - 矩阵的条件数与范数有关
  - 当$\boldsymbol{A}$的条件数相对地大时，即$\mathrm{cond}(\boldsymbol{A}) \gg 1$时，式$\boldsymbol{Ax} = \boldsymbol{b}$是病态的（即$\boldsymbol{A}$是病态矩阵，或者说$\boldsymbol{A}$是坏条件的）
  - 当$\boldsymbol{A}$的条件数相对地小时，式$\boldsymbol{Ax} = \boldsymbol{b}$是良态的（或者说$\boldsymbol{A}$是好条件的）
  - $\boldsymbol{A}$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. $\mathrm{cond}(\boldsymbol{A})_\infty = \|\boldsymbol{A}^{-1}\|_\infty \|\boldsymbol{A}\|_\infty$

2. $\boldsymbol{A}$的谱条件数

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

- 当$\boldsymbol{A}$为对称阵时，$\mathrm{cond}(\boldsymbol{A})_2 = \frac{|\lambda_1|}{|\lambda_n|}$，其中$\lambda_1$和$\lambda_n$为$\boldsymbol{A}$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$\boldsymbol{A}$，都有$\mathrm{cond}(\boldsymbol{A})_v \geq 1$

$$\mathrm{cond}(\boldsymbol{A})_v = \left\|\boldsymbol{A}^{-1}\right\|_v \|\boldsymbol{A}\|_v \geq \left\|\boldsymbol{A}^{-1}\boldsymbol{A}\right\|_v = \|\boldsymbol{I}\|_v = 1$$

2. $\boldsymbol{A}$为非奇异矩阵，c为不等于零的常数，则

$$\mathrm{cond}(c\boldsymbol{A})_v = \mathrm{cond}(\boldsymbol{A})_v$$

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

$$\mathrm{cond}(\boldsymbol{R}\boldsymbol{A})_2 = \mathrm{cond}(\boldsymbol{A}\boldsymbol{R})_2 = \mathrm{cond}(\boldsymbol{A})_2$$

例7.9

- 已知Hilbert矩阵$\boldsymbol{H}_n$，计算$\boldsymbol{H}_3$的条件数

$$\boldsymbol{H}_n=\begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\ \vdots & \vdots & & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \cdots & \frac{1}{2n-1} \end{bmatrix}$$

- 解

$$\boldsymbol{H}_3=\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}, \qquad \boldsymbol{H}_3^{-1}=\begin{bmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{bmatrix}$$

例7.9（续）

- 计算$\boldsymbol{H}_3$条件数$\mathrm{cond}(\boldsymbol{H}_3)_\infty$

$$\|\boldsymbol{H}_3\|_\infty = \frac{11}{6}, \qquad \left\|\boldsymbol{H}_3^{-1}\right\|_\infty = 408$$

所以$\mathrm{cond}(\boldsymbol{H}_3)_\infty = 748$

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

考虑方程组

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

- 精确解为$\boldsymbol{x} = (1,1,1)^{\mathrm{T}}$

例7.9（续）

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

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

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

- 式(7.6.8)的解为

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

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

例7.9（续）

- 于是

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

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

- $\boldsymbol{H}_3$与$\boldsymbol{b}$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$\mathrm{cond}(\boldsymbol{A}) = \|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|$，而计算$\|\boldsymbol{A}^{-1}\|$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$\boldsymbol{A}$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，$\boldsymbol{A}$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

$$I_{23}(H_3+\delta H_3)=\begin{bmatrix}1 & & \\ 0.333 & 1 & \\ 0.500 & 0.994 & 1\end{bmatrix}\begin{bmatrix}1 & 0.5000 & 0.3330 \\ & 0.0835 & 0.0891 \\ & & -0.00507\end{bmatrix}$$

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

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

定理**7.16** $\Rightarrow |\lambda_1|\leq\|\boldsymbol{A}\|,\frac{1}{|\lambda_n|}\leq\left\|\boldsymbol{A}^{-1}\right\|\Rightarrow \mathrm{cond}(\boldsymbol{A})\geq\frac{|\lambda_1|}{|\lambda_n|}\gg 1$

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$\boldsymbol{A}$可能是病态的

4. 如果系数矩阵$\boldsymbol{A}$元素间数量级相差很大，并且无一定规则，则$\boldsymbol{A}$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$\boldsymbol{Ax}=\boldsymbol{b}$的问题转化为求解一等价方程组

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

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

病态问题的求解（续）

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

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

$$\begin{bmatrix} 1 & 10^4 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 10^4 \\ 2 \end{bmatrix},$$

计算$\mathrm{cond}(\boldsymbol{A})_\infty$和解

例7.10（续）

- 由
$$A=\begin{bmatrix}1 & 10^4\\ 1 & 1\end{bmatrix}$$
可得
$$A^{-1}=\frac{1}{10^4-1}\begin{bmatrix}-1 & 10^4\\ 1 & -1\end{bmatrix}$$
$$\mathrm{cond}(A)_\infty=\frac{(1+10^4)^2}{10^4-1}\approx 10^4$$

- 用列主元素消去法解$\boldsymbol{Ax}=\boldsymbol{b}$时（计算到三位数字），得
$$(\boldsymbol{A},\boldsymbol{b})\to\left[\begin{array}{cc:c}1 & 10^4 & 10^4\\ 0 & -10^4 & -10^4\end{array}\right]$$
于是得到很坏的结果：$x_1=0$，$x_2=1$

例7.10（续）

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

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

$$\operatorname{cond}(\boldsymbol{A}')_\infty = \frac{1}{1-10^{-4}} \approx 1$$

大幅下降

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

$$(\boldsymbol{A}', \boldsymbol{b}') \to \left[\begin{array}{cc:c} 1 & 1 & 2 \\ 10^{-4} & 1 & 1 \end{array}\right] \to \left[\begin{array}{cc:c} 1 & 1 & 2 \\ 0 & 1 & 1 \end{array}\right]$$

从而得到较好的结果：$x_1 = 1$，$x_2 = 1$

事后误差估计

- 设$\bar{\boldsymbol{x}}$为方程组$\boldsymbol{Ax}=\boldsymbol{b}$的近似解，于是可计算$\bar{\boldsymbol{x}}$的剩余向量$\boldsymbol{r}=\boldsymbol{b}-\boldsymbol{A}\bar{\boldsymbol{x}}$，当$\boldsymbol{r}$很小时，$\bar{\boldsymbol{x}}$是否为$\boldsymbol{Ax}=\boldsymbol{b}$一个较好的近似解呢？

- **定理7.21**（事后误差估计） 设$\boldsymbol{A}$为非奇异阵，$\boldsymbol{Ax}=\boldsymbol{b}\neq\boldsymbol{0}$，$\bar{\boldsymbol{x}}$是方程组的近似解，$\boldsymbol{r}=\boldsymbol{b}-\boldsymbol{A}\bar{\boldsymbol{x}}$，则

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

定理7.21（证明）

- 由
$$\boldsymbol{A}(\boldsymbol{x}-\overline{\boldsymbol{x}})=\boldsymbol{b}-\boldsymbol{A}\overline{\boldsymbol{x}}=\boldsymbol{r}$$
得
$$\boldsymbol{x}-\overline{\boldsymbol{x}}=\boldsymbol{A}^{-1}\boldsymbol{r} \quad \Rightarrow \|\boldsymbol{x}-\overline{\boldsymbol{x}}\| \le \left\|\boldsymbol{A}^{-1}\right\|\|\boldsymbol{r}\|$$
- 又有
$$\|\boldsymbol{b}\|=\|\boldsymbol{A}\boldsymbol{x}\| \le \|\boldsymbol{A}\|\,\|\boldsymbol{x}\|$$
得
$$\frac{\mathbf{1}}{\|\boldsymbol{x}\|} \le \frac{\|\boldsymbol{A}\|}{\|\boldsymbol{b}\|}$$
- 因此
$$\frac{\|\boldsymbol{x}-\overline{\boldsymbol{x}}\|}{\|\boldsymbol{x}\|} \le \frac{\|\boldsymbol{A}\|\left\|\boldsymbol{A}^{-1}\right\|\|\boldsymbol{r}\|}{\|\boldsymbol{b}\|}=\operatorname{cond}(\boldsymbol{A})\frac{\|\boldsymbol{r}\|}{\|\boldsymbol{b}\|}$$

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$\bar{\boldsymbol{x}}$为选主元素Gauss消去法解$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$的计算解，$\boldsymbol{x}$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$\|\boldsymbol{x}-\bar{\boldsymbol{x}}\|$，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$\bar{\boldsymbol{x}}$是扰动方程组$(\boldsymbol{A}+\delta\boldsymbol{A})\boldsymbol{x}=\boldsymbol{b}$的精确解
  - 具体定理可以查看教材

- 计算解$\bar{\boldsymbol{x}}$的相对误差限依赖于$\mathrm{cond}(\boldsymbol{A})_\infty$，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

- Guass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