第三章 线性方程

第（3）部分：误差分析

目录

问题背景

解 $\boldsymbol{Ax}=\boldsymbol{b}$ 时，实际在计算机中存储的 $\boldsymbol{A}$ 和 $\boldsymbol{b}$ 未必是真实值。

误差的可能来源：

- $\boldsymbol{A}$ 和 $\boldsymbol{b}$ 中可能带有实验测量造成的误差
- 计算机存储时，有效数字造成的浮点精度误差
- $\cdots$

我们这一部分探究的核心问题是：误差如何影响真解。

假设真实的方程组是

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

而我们实际在计算机里解的方程是

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

我们假设误差 $\delta\boldsymbol{A}$（或 $\delta\boldsymbol{b}$）相对于 $\boldsymbol{A}$（或 $\boldsymbol{b}$）很小，我们希望量化

$$\frac{\text{解的相对误差}}{\text{输入来源的相对误差}} = \frac{\|\tilde{\boldsymbol{x}} - \boldsymbol{x}\|/\|\boldsymbol{x}\|}{\|\delta\boldsymbol{A}\|/\|\boldsymbol{A}\| \quad \text{或} \quad \|\delta\boldsymbol{b}\|/\|\boldsymbol{b}\|}$$

为实现这个目标，需要引入一个概念叫做向量和矩阵的“ 大小 ”；这一概念类似于实数的绝对值。

例子 $\rightarrow$ 现象

我们考虑如下的方程

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

其精确解为 $\boldsymbol{x} = \begin{bmatrix} 2 & 0 \end{bmatrix}^T$

而若我们考虑

$$\begin{bmatrix} 1 & 1 \\ 1 & 1.0001 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2.0001 \end{bmatrix}$$

得到的解为 $\boldsymbol{x} = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$

问题：此处我们的计算没有任何的误差，那么为什么会差那么多呢?

现象 $\to$ 定义

定义 (定义 7.7)

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。

注：病态与否是问题本身的问题，与算法无关。

目标：理解病态是什么意思，以及如何使用一个量化工具来刻画它。

目标

我们首先需要回顾和了解三个概念:

- 特征值和特征向量
- 矩阵范数
- 谱半径

然后我们会引入条件数这一概念和量化误差

目录

特征值和特征向量

定义 (特征值和特征向量)

(1) 已知 $\boldsymbol{A} = (a_{ij})_{n\times n}$, 要求代数方程 $\varphi(\lambda) = \det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0$ 的根。$\varphi(\lambda)$ 称为 $\boldsymbol{A}$ 的特征多项式。该式子展开即有

$$\varphi(\lambda) = \lambda^n + c_1\lambda^{n-1} + \cdots + c_n = 0$$

一般 $\varphi(\lambda)$ 有 n 个零点, 称为 $\boldsymbol{A}$ 的特征值。

(2) 设 λ 为 $\boldsymbol{A}$ 的特征值，要求 $\boldsymbol{Ax} = \lambda\boldsymbol{x}$ 的非零解。对应的 $\boldsymbol{x}$ 被称为特征向量。

回顾练习： 请计算特征值和特征向量 $\boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

定理

如果 $\lambda_i(i=1,2,\cdots,n)$ 是矩阵 $\boldsymbol{A}$ 的特征值, 则有

- $\sum_{i=1}^{n}\lambda_i=\sum_{i=1}^{n}a_{ii}=\operatorname{tr}\boldsymbol{A}$;
- $\det\boldsymbol{A}=\lambda_1\lambda_2\cdots\lambda_n$

课外练习: 请验证

$$\det(\boldsymbol{A}-\lambda\boldsymbol{I})=(-1)^n\big(\lambda^n-\operatorname{tr}(\boldsymbol{A})\lambda^{n-1}+\cdots+(-1)^n\det(\boldsymbol{A})\big)$$

定理 (定理 9.2；考虑相似矩阵)

设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 为相似矩阵 (即存在非奇异阵 $\boldsymbol{T}$ 使 $\boldsymbol{B} = \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$), 则

- $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的特征值;
- 若 $\boldsymbol{x}$ 是 $\boldsymbol{B}$ 的一个特征向量, 则 $\boldsymbol{T}\boldsymbol{x}$ 是 $\boldsymbol{A}$ 的特征向量。

定理 (考虑对称矩阵)

- 若 $\boldsymbol{A}$ 为实对称矩阵，存在互相垂直的向量 $u_1, u_2, \ldots, u_n$ 使得 $\boldsymbol{A}u_i = \lambda_i u_i$, 其中特征向量 $\lambda_i \in \mathbb{R}$ 是实数。
- 并且若 $\boldsymbol{A}^{-1}$ 存在（即若任何特征值皆非零 $\lambda_i \neq 0$)，则 $\boxed{\boldsymbol{A}^{-1}u_i = \dfrac{1}{\lambda_i}u_i}$,

  即 $\boldsymbol{A}^{-1}$ 具有特征值 $1/\lambda_i$ 和特征向量 u_i。

对于向量 $\boldsymbol{x}$， $\boxed{R(\boldsymbol{x}) = \frac{(\boldsymbol{A}\boldsymbol{x},\boldsymbol{x})}{(\boldsymbol{x},\boldsymbol{x})}}$ 被称为**Rayleigh 商**。

问题：这个量有什么几何含义？假设我们考虑 $\boldsymbol{x}$ 是单位向量。

练习： 我们考虑 $\boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$，$\boldsymbol{x} = \begin{bmatrix} r\cos(\theta) \\ r\sin(\theta) \end{bmatrix}$

- 模长 r 对于计算 Rayleigh 商重要吗?
- 请直接计算 $\min\limits_{\substack{\boldsymbol{x}\in\mathbb{R}^n \\ \boldsymbol{x}\neq 0}} R(\boldsymbol{x})$ 和 $\max\limits_{\substack{\boldsymbol{x}\in\mathbb{R}^n \\ \boldsymbol{x}\neq 0}} R(\boldsymbol{x})$，这里 $n=2$。
- $\boldsymbol{x}$ 什么时候使得 $R(\boldsymbol{x})$ 能取到最小值? $\boldsymbol{x}$ 什么时候使得 $R(\boldsymbol{x})$ 能取到最大值? 这些极值和特征值有什么关系?

定理 (定理 9.4)

设 $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ 为对称矩阵（其特征值次序记作 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ ，对应的特征向量 $u_1, u_2, \cdots, u_n$ 组成规范化正交组, 即 $(u_i, u_j) = \delta_{ij}$), 则

- 对于任何非零 $\boldsymbol{x} \in \mathbb{R}^n$:

$$\boxed{\lambda_n \leqslant \frac{(\boldsymbol{Ax}, \boldsymbol{x})}{(\boldsymbol{x}, \boldsymbol{x})} \leqslant \lambda_1} \tag{1}$$

- *Rayleigh* 商的极值分别是最大和最小的特征值:

$$\lambda_1 = \max_{\substack{\boldsymbol{x}\in\mathbb{R}^n \\ \boldsymbol{x}\neq 0}} \frac{(\boldsymbol{Ax}, \boldsymbol{x})}{(\boldsymbol{x}, \boldsymbol{x})} \qquad \lambda_n = \min_{\substack{\boldsymbol{x}\in\mathbb{R}^n \\ \boldsymbol{x}\neq 0}} \frac{(\boldsymbol{Ax}, \boldsymbol{x})}{(\boldsymbol{x}, \boldsymbol{x})}.$$

目录

向量的范数

我们先考虑一下向量的范数：

对于一个向量 $\boldsymbol{x} = [x_1, x_2, \cdots, x_n]$

- 数量积 $(\boldsymbol{x}, \boldsymbol{y}) = \sum_i x_i y_i$（是使用尖括号还是圆括号都可以）
- $\|\boldsymbol{x}\|_2 = (\boldsymbol{x}, \boldsymbol{x})^{1/2} = (\sum_i x_i^2)^{1/2}$ 为范数（模）

性质：

- $(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{y}, \boldsymbol{x})$
- Cauchy-Schwarz 不等式：$|(\boldsymbol{x}, \boldsymbol{y})| \leq \|\boldsymbol{x}\|_2 \|\boldsymbol{y}\|_2$
- 三角不等式：$\|\boldsymbol{x} + \boldsymbol{y}\|_2 \leq \|\boldsymbol{x}\|_2 + \|\boldsymbol{y}\|_2$

把范数中下标 2 改为任何的 $p \in [1, \infty)$:

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$$ p-范数

例子: $\boldsymbol{x} = [1, 3]$，请计算 $\|\boldsymbol{x}\|_5$, $\|\boldsymbol{x}\|_{10}$, $\|\boldsymbol{x}\|_{20}$，你有什么发现?

把范数中下标 2 改为任何的 $p \in [1, \infty)$:

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$$ p-范数

例子: $\boldsymbol{x} = [1, 3]$,请计算 $\|\boldsymbol{x}\|_5$, $\|\boldsymbol{x}\|_{10}$, $\|\boldsymbol{x}\|_{20}$,你有什么发现?

答案: 对于这个例子,当 $p \to \infty$, $\|\boldsymbol{x}\|_p \to 3$

把范数中下标 2 改为任何的 $p \in [1, \infty)$：

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} \quad \text{p-范数}$$

例子： $\boldsymbol{x} = [1, 3]$，请计算 $\|\boldsymbol{x}\|_5$, $\|\boldsymbol{x}\|_{10}$, $\|\boldsymbol{x}\|_{20}$，你有什么发现?

<u>答案</u>：对于这个例子，当 $p \to \infty$, $\|\boldsymbol{x}\|_p \to 3$

更一般的，当 $p \to \infty$,

$$\lim_{p\to\infty} \|\boldsymbol{x}\|_p \overset{\text{现象/结论}}{=} \max_i |x_i| \overset{\text{定义为}}{=} \|\boldsymbol{x}\|_\infty$$

$\|\cdot\|$ 是一个函数：输入是一个向量 $\boldsymbol{x}$，输出 $\|\boldsymbol{x}\|$ 是一个非负实数。

定义 (向量的范数)

我们说它是一个范数如果它满足如下条件:

- $\|\boldsymbol{x}\| \geq 0$ ($\|\boldsymbol{x}\| = 0$ 当且仅当 $\boldsymbol{x} = 0$)
- $\|\alpha\boldsymbol{x}\| = |\alpha|\|\boldsymbol{x}\|$ 对于任何一个实数 α
- $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$

我们使用下标来注明这个范数是什么意义下的。

该量本质是用来描述向量的大小。

对于有限维空间，向量范数具有连续性（定理 7.11）和等价性（定理 7.12）。

矩阵的范数

对于这样的矩阵

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

目的：是找一个概念来定义矩阵的大小。

我们希望该概念满足一些好的性质：

- $\|\boldsymbol{A}\| \geq 0$（其中 $\|\boldsymbol{A}\| = 0$ 当且仅当 $\boldsymbol{A} = 0$）（正定条件）
- $\|c\boldsymbol{A}\| = |c|\|\boldsymbol{A}\|$ （齐次条件）
- $\|\boldsymbol{A} + \boldsymbol{B}\| \leq \|\boldsymbol{A}\| + \|\boldsymbol{B}\|$（三角不等式）
- $\|\boldsymbol{A}\boldsymbol{B}\| \leq \|\boldsymbol{A}\|\|\boldsymbol{B}\|$

例 (Frobenius 范数)

视角 1：把矩阵看成是一个 n^2 维的向量： $F(\boldsymbol{A}) = \|\boldsymbol{A}\|_F = \left(\sum_{i,j=1}^{n} a_{ij}^2\right)^{1/2}$

例 (算子范数)

视角 2：把矩阵看成是作用在向量上的一个变换：

$$\|\boldsymbol{A}\|_v := \max_{\boldsymbol{x}\neq 0} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_v}{\|\boldsymbol{x}\|_v} \tag{2}$$

由于该定义，我们可知 $\|\boldsymbol{A}\boldsymbol{x}\|_v \leqslant \|\boldsymbol{A}\|_v \|\boldsymbol{x}\|_v$；且我们可以证明 $\|\cdot\|_v$ 是一个合适的矩阵范数。

练习：请根据定义来证明

$$\|\boldsymbol{A}\boldsymbol{B}\|_v \leq \|\boldsymbol{A}\|_v \|\boldsymbol{B}\|_v \tag{3}$$

矩阵

假设我们考虑 $\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$, $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

$$\|\boldsymbol{A}\boldsymbol{x}\|_2 = \sqrt{x_1^2 + (2x_2)^2}, \qquad \|\boldsymbol{x}\|_2 = \sqrt{x_1^2 + x_2^2}$$

因此

$$\|\boldsymbol{A}\|_2 = \max_{\boldsymbol{x}\neq 0} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2} = \max_{\boldsymbol{x}\neq 0} \frac{\sqrt{x_1^2 + 4x_2^2}}{\sqrt{x_1^2 + x_2^2}} \leq \max_{\boldsymbol{x}\neq 0} \frac{\sqrt{4x_1^2 + 4x_2^2}}{\sqrt{x_1^2 + x_2^2}} = 2.$$

练习题：计算 $\|\boldsymbol{A}\|_\infty$

定理 (定理 7.15)

我们考虑任意的矩阵 $\boldsymbol{A} \in \mathbb{R}^{n\times n}$，我们有如下的结论：

$$
\begin{aligned}
&1^\circ \quad \|\boldsymbol{A}\|_1 = \max_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}| \quad (\boldsymbol{A}\text{ 的列范数}) \\
&2^\circ \quad \|\boldsymbol{A}\|_2 = \sqrt{\lambda_{\max}\left(\boldsymbol{A}^T\boldsymbol{A}\right)} \quad (\boldsymbol{A}\text{ 的2范数}) \\
&3^\circ \quad \|\boldsymbol{A}\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}| \quad (\boldsymbol{A}\text{ 的行范数})
\end{aligned}
\tag{4}
$$

（具体的证明请见课本或课内。）

不同范数的优势是什么?

$\|\cdot\|_\infty$ 和 $\|\cdot\|_1$ 在实际中比较容易算

$\|\cdot\|_2$ 具有很高的理论价值

定理

- $\boldsymbol{A}\boldsymbol{A}^\top$ 和 $\boldsymbol{A}^\top\boldsymbol{A}$ 有同样的特征值（课后作业第 *26* 题）
- $\|\boldsymbol{A}\|_2 = \left\|\boldsymbol{A}^\top\right\|_2$
- 若 $\boldsymbol{A}^{-1}$ 存在，则

$$\left\|\boldsymbol{A}^{-1}\right\|_2 = \frac{1}{\sqrt{\lambda_{\min}\left(\boldsymbol{A}\boldsymbol{A}^\top\right)}} = \frac{1}{\sqrt{\lambda_{\min}\left(\boldsymbol{A}^\top\boldsymbol{A}\right)}} \tag{5}$$

证明.

结论一：$\boldsymbol{A}\boldsymbol{A}^\top u = \lambda u$ 令 $v = \boldsymbol{A}^\top u$ 则 $\boldsymbol{A}^\top\boldsymbol{A}v = \boldsymbol{A}^\top\boldsymbol{A}\boldsymbol{A}^\top u = \boldsymbol{A}^\top\lambda u = \lambda v$.

结论二可由结论一和 2 范数的公式(4)得到。

对于结论三，

$$\left\|\boldsymbol{A}^{-1}\right\|_2 \overset{\text{公式(4)}}{=} \sqrt{\lambda_{\max}\left(\left(\boldsymbol{A}^{-1}\right)^\top\boldsymbol{A}^{-1}\right)} = \sqrt{\lambda_{\max}\left(\left(\boldsymbol{A}\boldsymbol{A}^\top\right)^{-1}\right)} = 1/\sqrt{\lambda_{\min}\left(\boldsymbol{A}\boldsymbol{A}^T\right)}$$

序列的收敛

实数的收敛：实数序列 $\{x_n\} \subset \mathbb{R}$ 收敛于 x^* 的意思是 $\lim_{n\to\infty} |x_n - x^*| = 0$

向量的收敛：向量序列 $\{\boldsymbol{x}_n\} \subset \mathbb{R}^n$ 收敛于 $\boldsymbol{x}^*$ 的意思是

$$\lim_{n\to\infty} \|\boldsymbol{x}_n - \boldsymbol{x}^*\| = 0$$

（此处选择不同的范数，对于收敛现象无区别，但是范数值还是有区别的）

矩阵的收敛：矩阵序列 $\{\boldsymbol{A}_n\} \subset \mathbb{R}^{n\times n}$ 收敛于 $\boldsymbol{A}^*$ 的意思是

$$\lim_{n\to\infty} \|\boldsymbol{A}_n - \boldsymbol{A}^*\| = 0$$

（此处选择不同的范数，对于收敛现象无区别，但是范数值还是有区别的）

目录

谱半径

定义 (谱半径)

设 $\boldsymbol{A}$ 的特征值为 λ_i $(i=1,2,\ldots,n)$，我们称

$$\rho(\boldsymbol{A}) \overset{\text{定义}}{=} \max_{1\le i\le n} |\lambda_i| \text{ 为}\boldsymbol{A}\text{ 的谱半径}$$

定理 (特征值上界)

假设 $\boldsymbol{A}$ 是 n 阶实矩阵，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。

定理

若 $\boldsymbol{A}$ 为对称矩阵，则 $\|\boldsymbol{A}\|_2 = \rho(\boldsymbol{A})$。*(证明：作业题)*

目录

情景 1：$\boldsymbol{A}$ 是精确的，$\boldsymbol{b}$ 有误差

定理 (定理 7. 19)

设 $\boldsymbol{A}$ 是非奇异阵, $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \neq \mathbf{0}$, 且 $\boldsymbol{A}(\boldsymbol{x} + \delta\boldsymbol{x}) = \boldsymbol{b} + \delta\boldsymbol{b}$, 则

$$\underbrace{\frac{\|\delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|}}_{\text{输出误差的相对误差}} \leq \underbrace{\|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|}_{\substack{\text{误差放大的}\\ \text{倍数上界}}} \underbrace{\frac{\|\delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|}}_{\text{输入误差的相对误差}}$$

（证明：参见课本或课内）

定理

若 $\|\boldsymbol{B}\| < 1$，则 $\boldsymbol{I} \pm \boldsymbol{B}$ 为非奇异矩阵，且 $\|(\boldsymbol{I} \pm \boldsymbol{B})^{-1}\| \leq \frac{1}{1-\|\boldsymbol{B}\|}$。其中 $\|\cdot\|$ 可以是任何一种算子范数。

情景 2：$\boldsymbol{b}$ 是精确的，$\boldsymbol{A}$ 有误差

定理 (定理 7.20)

设 $\boldsymbol{A}$ 是非奇异阵, $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \neq \mathbf{0}$, 且 $(\boldsymbol{A} + \delta\boldsymbol{A})(\boldsymbol{x} + \delta\boldsymbol{x}) = \boldsymbol{b}$。若 $\|\boldsymbol{A}^{-1}\|\|\delta\boldsymbol{A}\| < 1$，则

$$\underbrace{\frac{\|\delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|}}_{\text{输出误差的相对误差}} \leq \frac{\|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|\frac{\|\delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|}}{1 - \|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|\frac{\|\delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|}}$$

定义 (定义 7.8)

设 $\boldsymbol{A}$ 为非奇异矩阵，称数 $\operatorname{cond}(\boldsymbol{A})_v = \left\|\boldsymbol{A}^{-1}\right\|_v \|\boldsymbol{A}\|_v \quad (v \in [1, \infty])$ 为矩阵 $\boldsymbol{A}$ 的**条件数**。

由前面的两个结论可知，条件数可以作为衡量病态程度的指标。

条件数小（良态）$\longrightarrow$ 条件数大（病态）

常用的两种:

(1) $\operatorname{cond}(\boldsymbol{A})_\infty := \left\|\boldsymbol{A}^{-1}\right\|_\infty \|\boldsymbol{A}\|_\infty$

(2) 谱条件数

$$\operatorname{cond}(\boldsymbol{A})_2 := \|\boldsymbol{A}\|_2 \left\|\boldsymbol{A}^{-1}\right\|_2 = \sqrt{\frac{\lambda_{\max}\left(\boldsymbol{A}^\top \boldsymbol{A}\right)}{\lambda_{\min}\left(\boldsymbol{A}^\top \boldsymbol{A}\right)}} \tag{6}$$

证明：请参考 PPT 内的公式(4)和(5)

条件数的性质

定理

条件数有如下性质:

(1) 对任何非奇异矩阵 $\boldsymbol{A}$, 都有 $\operatorname{cond}(\boldsymbol{A})_v \geqslant 1$。

(2) 设 $\boldsymbol{A}$ 为非奇异矩阵, c 为不等于零的常数, 则 $\operatorname{cond}(c\boldsymbol{A})_v = \operatorname{cond}(\boldsymbol{A})_v$.

(3) 如果 $\boldsymbol{A}$ 为正交矩阵 *, 则 $\operatorname{cond}(\boldsymbol{A})_2 = 1$。†

(4) 如果 $\boldsymbol{A}$ 为非奇异矩阵, $\boldsymbol{R}$ 为正交矩阵, 则

$$\operatorname{cond}(\boldsymbol{RA})_2 = \operatorname{cond}(\boldsymbol{AR})_2 = \operatorname{cond}(\boldsymbol{A})_2$$

---

* 若 $\boldsymbol{A}$ 满足 $\boldsymbol{AA}^T = \boldsymbol{I}$, 则称 $\boldsymbol{A}$ 为正交矩阵

† 证明: 使用公式(6)

应用：Hilbert 矩阵是病态的

预告：

后面章节会介绍高维 Hilbert 矩阵的条件数过大（不稳定）$\Longrightarrow$ 正交多项式在计算时的重要性！

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

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

$$
\mathrm{cond}(\boldsymbol{H}_6)_\infty \approx 3 \times 10^6
$$

实际中如何判断问题是否病态

- 三角约化时出现小主元
- 最大特征值和最小特征值的比值（按照绝对值）过大
- 系数矩阵的行列式过小，或者系数矩阵的某些行线性相关性比较大
- 元素间的数量级相差很大，且无规律

则 $\boldsymbol{A}$ 有很大的概率是病态的矩阵。

定理 (定理 7.21，事后误差估计)

假设

- $\boldsymbol{A}$ 为非奇异矩阵，$\boldsymbol{x}$ 是精确解，$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \neq \mathbf{0}$;
- 设 $\bar{\boldsymbol{x}}$ 是方程组的近似解，定义残差 $\boldsymbol{r} = \boldsymbol{b} - \boldsymbol{A}\bar{\boldsymbol{x}}$。

则

$$\frac{\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|}{\|\boldsymbol{x}\|} \leqslant \mathrm{cond}(\boldsymbol{A}) \cdot \frac{\|\boldsymbol{r}\|}{\|\boldsymbol{b}\|}$$

提示：

- 条件数如果比较大，则上述估计很难有用；
- 若条件数比较小，则残差的大小可以控制方程组解的误差。

目录

重点知识总结

- 病态方程组和病态矩阵的含义
- Rayleigh 商的定义、计算、以及它和特征值的关系
- 向量和矩阵范数的定义、性质等；重点需要掌握算子范数
- 谱半径的定义和性质
- 对于矩阵误差分析的三个定理

  (对应课本里的定理 7.19、定理 7.20、定理 7.21)
- 条件数的定义、计算和性质