北京航空航天大學
BEIHANG UNIVERSITY

数值分析

主讲教师：贺慧霞

北京航空航天大学数学科学学院

第二章 线性方程组的解法

2.3 矩阵的条件数与病态线性方程组

问题：求解Ax＝b时，A和b的误差对x有何影响?

数值算例

$$\begin{bmatrix} 2 & 3 \\ 2 & 3.000\,01 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 8 \\ 8.000\,02 \end{bmatrix} \qquad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

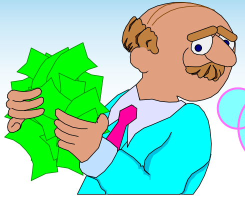
$$\begin{bmatrix} 2 & 3 \\ 2 & 2.999\,99 \end{bmatrix}\begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix} = \begin{bmatrix} 8 \\ 8.000\,03 \end{bmatrix} \qquad \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix} = \begin{bmatrix} 8.5 \\ -3 \end{bmatrix}$$

perturbation

perturbation

病态线性方程组

It's funny that such small perturbations in the coefficients lead to so big change in the solution!



由实际问题建立起来的线性方程组$Ax=b$本身存在**模型误差和观测误差**，或者是由计算得到的，存在**舍入误差**等。总之，A,b都会有一定扰动ΔA, Δb, 因此实际处理的是A+ ΔA或b+ Δb ，我们需要分析A或b的扰动对解的影响.

$$Ax=b \longrightarrow \underline{(A+\Delta A)x = b+\Delta b}$$

一、矩阵的条件数与病态线性方程组

$\Delta x = ?$

1：A非奇异,设精确,b有误差Δb,导致解x有多大误差?

$$A(x+\Delta x) = b + \Delta b$$

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

$$\Delta x = A^{-1}\Delta b$$

$$\Downarrow$$

$$\|\Delta x\| \le \|A^{-1}\| \cdot \|\Delta b\|$$

$$Ax = b \qquad \|\boldsymbol{A}\| \cdot \|\boldsymbol{x}\| \ge \|\boldsymbol{b}\|$$

$$\Downarrow$$

$$\frac{1}{\|x\|} \le \frac{\|A\|}{\|b\|}$$

$$\Downarrow$$

b的相对误差被放大

$$\frac{\|\Delta x\|}{\|x\|} \le \|A\| \cdot \|A^{-1}\| \frac{\|\Delta b\|}{\|b\|}$$

2：设b精确，A有误差ΔA，导致解x有多大误差？ 设解为 $x+\Delta x$

$$(A+\Delta A)(x+\Delta x)=b$$

$$Ax=b$$

$$(A+\Delta A)\Delta x=-(\Delta A)\cdot x$$

$A\Delta x=-(\Delta A)x-(\Delta A)\Delta x$

$\Delta x=-A^{-1}(\Delta A)x-A^{-1}(\Delta A)\Delta x$

$\|\Delta x\|\leq\|A^{-1}\|\cdot\|\Delta A\|\,\|x\|+\|A^{-1}\|\|\Delta A\|\cdot\|\Delta x\|$

$(1-\|A^{-1}\|\cdot\|\Delta A\|)\dfrac{\|\Delta x\|}{\|x\|}\leq\|A^{-1}\|\cdot\|\Delta A\|$

设A非奇异，$\|\Delta A\|\,\|A^{-1}\|<1$

$\dfrac{\|\Delta x\|}{\|x\|}\leq\dfrac{\|A^{-1}\|\cdot\|\Delta A\|}{1-\|A^{-1}\|\cdot\|\Delta A\|}$

$$\frac{\|\Delta x\|}{\|x\|}\leq\frac{\|A\|\cdot\|A^{-1}\|}{1-\|A\|\cdot\|A^{-1}\|\dfrac{\|\Delta A\|}{\|A\|}}\frac{\|\Delta A\|}{\|A\|}$$

3：设b,A分别有误差Δb和ΔA,导致解x有多大误差?

$$(A+\Delta A)(x+\Delta x)=b+\Delta b$$

A非奇异，$\|\Delta A\|\,\|A^{-1}\|<1$

解的相对误差

b的相对误差

A的相对误差

$$\frac{\|\Delta x\|}{\|x\|}\le\frac{\|A\|\cdot\|A^{-1}\|}{1-\|A\|\cdot\|A^{-1}\|\dfrac{\|\Delta A\|}{\|A\|}}\left(\frac{\|\Delta b\|}{\|b\|}+\frac{\|\Delta A\|}{\|A\|}\right)$$

当方程组的系数矩阵A或右端项b受到扰动ΔA, Δb时，引起的解的相对误差完全由$\|A\|\cdot\|A^{-1}\|$来决定，它刻画了方程组的解对原始数据的敏感程度。

4. 矩阵的条件数（Condition number ）

定义：对非奇异矩阵A，称乘积$\|A\|\,\|A^{-1}\|$为矩阵A的**条件数**，记为 $\mathrm{cond}(A) = \|A\|\,\|A^{-1}\|$

- $\mathrm{cond}(A)$的具体大小与$\|\cdot\|$有关，但相对大小一致；
- $\mathrm{cond}(A)$的大小本质取决于A，与解题的方法无关；
- 如果A是奇异的，$\mathrm{cond}(A) = \infty$。

5. 常用的矩阵条件数

1. 矩阵A的行和范数： $\|A\|_\infty = \max\limits_{1\le i\le n}\sum\limits_{j=1}^{n}|a_{ij}|$

$$cond(A)_\infty = \|A\|_\infty \cdot \|A^{-1}\|_\infty$$

2. 矩阵A的列和范数： $\|A\|_1 = \max\limits_{1\le j\le n}\sum\limits_{i=1}^{n}|a_{ij}|$

$$cond(A)_1 = \|A\|_1 \cdot \|A^{-1}\|_1$$

3. 矩阵A的谱范数： $\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)}$

$$cond(A)_2 = \|A\|_2 \cdot \|A^{-1}\|_2$$

例1：Hilbert 阵 $H_n=\begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{1}{n} & \frac{1}{n+1} & \cdots & \frac{1}{2n-1} \end{bmatrix}$ 计算H_3的条件数.

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

$$\|H_3\|_\infty=\frac{11}{6},\quad \|H_3^{-1}\|_\infty=408,\text{所以}cond(H_3)=748.$$

注：现在用Matlab数学软件可以很方便求矩阵的条件数!

$$cond(H_6)_\infty = 2.9\times 10^7,\quad cond(H_7)_\infty = 9.85\times 10^8.$$

$\therefore$ 当 n 越大时，H_n 矩阵病态越严重．

(2)考虑方程组 $H_3 x = (11/6,\quad 13/12,\quad 47/60)^T = b,$ $\boxed{(1)}$

设 H_3 及 有微小误差(取3位有效数字)有

$$\begin{pmatrix} 1.00 & 0.500 & 0.333 \\ 0.500 & 0.333 & 0.250 \\ 0.333 & 0.250 & 0.200 \end{pmatrix}\begin{pmatrix} x_1+\delta x_1 \\ x_2+\delta x_2 \\ x_3+\delta x_3 \end{pmatrix} = \begin{pmatrix} 1.83 \\ 1.08 \\ 0.783 \end{pmatrix},$$

简记作 $(H_3+\delta H_3)(x+\delta x) = b+\delta b.$ $\boxed{(2)}$

方程组（1）与 (2) 的精确解分别为：$x = (1,\quad 1,\quad 1,)^T,$

$$x+\delta x = (1.089512538,\quad 0.487967062,\quad 1.491002798)^T.$$

于是 $\delta x = (0.0895,\quad -0.5120,\quad 0.4910)^T$,

$$\frac{\|\delta H_3\|_\infty}{\|H_3\|_\infty} \approx 0.18 \times 10^{-3} < 0.02\%, \quad \frac{\|\delta b\|_\infty}{\|b\|_\infty} \approx 0.182\%,$$

$$\frac{\|\delta x\|_\infty}{\|x\|_\infty} \approx 51.2\%.$$

即H_3与b的相对误差很小，但引起解的相对误差超过50%.

6. 矩阵条件数的一些性质

1．A非奇异，则cond（A）⩾1；

2．A非奇异，k≠0，则cond（kA）＝cond（A）；

3．A是正交矩阵，则$\mathrm{cond}(A)_2=1$；

4．A可逆，R正交，则$\mathrm{cond}(RA)_2=\mathrm{cond}(AR)_2=\mathrm{cond}(A)_2$

5.设A是非奇异实对称矩阵，则有 $\mathrm{cond}(A)_2=\left|\dfrac{\lambda_1}{\lambda_n}\right|$

$\lambda_1 \to A$的模最大的特征值．$\lambda_n \to A$的模最小的特征值．

二、病态、良态线性方程组

1. 病态线性方程组

定义：对线性方程组$Ax=b$，若$cond\left(A\right)$相对很大,则称$Ax=b$是病态的线性方程组(也称A是病态矩阵)；

若$cond\left(A\right)$相对很小，则称$Ax=b$是良态的线性方程组(也称A是良态矩阵).

例2　设线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 为

$$\begin{bmatrix} 0.2161 & 0.1441 \\ 1.2969 & 0.8648 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 0.144 \\ 0.8642 \end{bmatrix}$$

(1) 求 $\mathrm{cond}(\boldsymbol{A})_\infty$；　　$\|A\|_\infty = 1.2969+0.8648$

(2) 在八位十进制的限制下，用列主元素 Gauss 消去法求解。

解　(1) $\boldsymbol{A}^{-1}=\begin{bmatrix} -0.8648\times 10^8 & 0.1441\times 10^8 \\ 1.2969\times 10^8 & -0.2161\times 10^8 \end{bmatrix}$　　$\|A^{-1}\|_\infty = (1.2969+0.2161)\times 10^{-8}$

$$\mathrm{cond}(\boldsymbol{A})_\infty = \|\boldsymbol{A}\|_\infty \|\boldsymbol{A}^{-1}\|_\infty \approx \underline{3.27\times 10^8}$$ 非常大

(2)

$$\begin{bmatrix} 0.2161 & 0.1441 & 0.1440 \\ \boxed{1.2969} & 0.8648 & 0.8642 \end{bmatrix} \xrightarrow[\text{选主元}]{r_1 \leftrightarrow r_2} \begin{bmatrix} 1.2969 & 0.8648 & 0.8642 \\ 0.2161 & 0.1441 & 0.1440 \end{bmatrix} \xrightarrow{r_2 - \frac{0.2161}{1.296} r_1}$$

$$\begin{bmatrix} 1.2969 & 0.8648 & 0.8642 \\ 0 & 1\times 10^{-8} & -1\times 10^{-8} \end{bmatrix}$$

$0.1441 - \frac{0.2161}{1.296} \times 0.8648$

$= 0.1441 - 0.14409999$

$= 1 \times 10^{-8}$

得到解 $x_2=-1, x_1=1.3331791$ 原方程组的精确解 $x_1=2, x_2=-2$

其相对误差为 $\frac{\|\Delta x\|_\infty}{\|x\|_\infty} = 50\%$ 解的精确性却很低

该方程组系数矩阵的条件数很大，方程组病态很严重

2. 判断线性方程组是否病态

A^{-1} 工作量大

注： 一般判断矩阵是否病态，并不计算A^{-1}，而由经验得出.

- ☞ 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.

$A=\begin{pmatrix} 0.1 & 0.1 \\ 0.1 & 10^{10} \end{pmatrix}$, $\mathrm{cond}(A)_\infty \approx 10^{11}$, $Ax=b$ 病态

$B=\begin{pmatrix} 10^{10} & 0.1 \\ 0.1 & 10^{10} \end{pmatrix}$ $\mathrm{cond}(B)_\infty \approx 1$ $Bx=b$ 良态

3. 缓和甚至解决线性方程组病态的手段

1. 采用高精度的算法（不是总有效）；

2. 通过行或列的平衡来降低矩阵的条件数（平衡法）；

平衡法: 取矩阵 A 的行绝对值最大元 $s_i = \max\limits_{1 \le j \le n} |a_{ij}|,$

构造对角阵 $D = \begin{pmatrix} \frac{1}{s_1} & & & \\ & \frac{1}{s_2} & & \\ & & \ddots & \\ & & & \frac{1}{s_n} \end{pmatrix},$

解方程组 $DAx = Db,$

可能良态

3. 残差校正法(迭代改善法)

例1 设 $\begin{pmatrix} 1 & 10^4 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 10^4 \\ 2 \end{pmatrix}$，计算 $\mathrm{cond}(A)_\infty$.

解: $A = \begin{pmatrix} 1 & 10^4 \\ 1 & 1 \end{pmatrix}$, $A^{-1} = \dfrac{1}{10^4 - 1}\begin{pmatrix} -1 & 10^4 \\ 1 & -1 \end{pmatrix}$ $cond(A)_\infty = \dfrac{(1+10^4)^2}{10^4 - 1} \approx 10^4$.

这个矩阵的元素大小不均，在 A 的第一行引进比例因子.

用平衡法计算 .则$\begin{cases} s_1 = \max\limits_{1\le i\le 2} |a_{1i}| = 10^4 \\ s_2 = 1, \end{cases}$，得 $DAx = Db$，记为 $A'x = b'$，

$\begin{pmatrix} 10^{-4} & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$，$(A')^{-1} = \dfrac{1}{1-10^{-4}}\begin{pmatrix} -1 & 1 \\ 1 & -10^{-4} \end{pmatrix}$，

于是：$cond(A')_\infty = \dfrac{1}{1-10^{-4}} \approx 4.$ $\Rightarrow A'x = b'$为良态的 .

用Gauss消元法求解原方程组：$(A|b) \to \left(\begin{array}{cc:c} 1 & 10^4 & 10^4 \\ 0 & -10^4 & -10^4 \end{array}\right)$,

得到一个很坏的解： $x_2 = 1, \quad x_1 = 0.$

用列主元 Gauss 消元法：$(A'|b') \to \left(\begin{array}{cc:c} 1 & 1 & 2 \\ 10^{-4} & 1 & 1 \end{array}\right) \to \left(\begin{array}{cc:c} 1 & 1 & 2 \\ 0 & 1 & 1 \end{array}\right)$,

得到一个较好的解： $x_2 = 1, \quad x_1 = 1.$

设 $Ax=b$, $A\in \mathrm{R}^{n\times n}$ 为非奇异矩阵,且为病态方程组(但不过分病态).

下面研究的问题是,当求得方程组的近似解 后,如何对其精度进行改善

单精度，也就是 float ，在 32 位机器上用 4 个字节来存储的；

双精度double是用 8 个字节来存储的。

3. 残差校正法步骤:

1. 用选主元三角分解法实现分解计算 $PA \doteq LU$,

2. 求解$PAx=LUx=Pb$，得到x的近似解x_1；(用单精度)

3. 校正x_1，计算残差$r_1=b-Ax_1$, 解$A\Delta x_1=r_1$, 得到Δx_1；

双精度

4. $x_2=x_1+\Delta x_1$；

如果第二步计算没有误差，则 $x_2=x_1+\Delta x_1$ 就为Ax=b的精确解；

$$Ax_2 = A(x_1+\Delta x_1) = Ax_1 + A\Delta x_1 = Ax_1 + r_1 = b$$

4. $x_2=x_1+\Delta x_1$；

事实上，总有舍入误差．

5.校正x_2，计算残差$r_2=b-Ax_2$，解$A\Delta x_2 = r_2$；得到Δx_2；

6. $x_3=x_2+\Delta x_2$；

7. …… ；

8. $r_k=b-Ax_k$，解 $A\Delta x_k=r_k$；得到Δx_k；

9. $x_{k+1}=x_k+\Delta x_k$，……

$\xrightarrow{n\to\infty} x^*$

重复该过程，就产生一近似解序列$\{x_k\}$，有时可能得到比较好的近似.

例2 用迭代改善法解 $\begin{pmatrix} 1.0303 & 0.99030 \\ 0.99030 & 0.95285 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2.4944 \\ 2.3988 \end{pmatrix}$

(记为 $Ax = b$) 矩阵元素保留5位有效数字.

解 精确解 $x^* = (1.2240, 1.2454)^T$(舍入到小数后第4位)

$$cond(A)_\infty = \|A\|_\infty \|A^{-1}\|_\infty \doteq 2 \times 2000 = 4000.$$

1. 首先实现分解计算 $PA \doteq LU$,且求 x_1

$$A \doteq \begin{pmatrix} 1 & 0 \\ 0.96118 & 1 \end{pmatrix}\begin{pmatrix} 1.0303 & 0.99030 \\ 0 & 0.00099 \end{pmatrix} = LU,$$

$$\frac{0.99030}{1.0303}$$

$$Ly = \begin{pmatrix} 1 & 0 \\ 0.96118 & 1 \end{pmatrix} y = \begin{pmatrix} 2.4944 \\ 2.3988 \end{pmatrix}, \Rightarrow y = \begin{pmatrix} 2.4944 \\ 0.0012326 \end{pmatrix},$$

$$Ux_1=\begin{pmatrix}1.0303 & 0.99030\\ 0 & 0.00099\end{pmatrix}x_1=\begin{pmatrix}2.4944\\ 0.0012326\end{pmatrix},\Rightarrow x_1=\begin{pmatrix}1.21549\\ 1.25424\end{pmatrix}$$

2. 校正x_1，令$r_1=b-Ax_1$，解$A\Delta x_1=r_1$

计算解$\Delta x_1=(0.00853501,-0.00887291)^T$.

3. $x_2=x_1+\Delta x_1$ $=(1.2250\ ,1.2387\ )^T$.

4. 令$r_2=b-Ax_2$，解$A\Delta x_2=r_2$；

$$\begin{pmatrix}1.0303 & 0.99030\\ 0.99030 & 0.95285\end{pmatrix}\begin{pmatrix}1.2250\\ 1.2387\end{pmatrix}=\begin{pmatrix}2.4888\\ 2.39341\end{pmatrix}$$

$$r_2=b-A\begin{pmatrix}x_{21}\\ x_{22}\end{pmatrix}=\begin{pmatrix}0.0056\\ 0.00539\end{pmatrix}\Rightarrow \Delta x_2=(-\ 0.00171037\ ,\ 0.0074343\ )^T.$$

5. $x_3 = x_2 + \Delta x_2 = (1.2233\ ,\ 1.2461\ )^T.$

6. 令$r_3 = b - Ax_3$，解$A\Delta x_3 = r_3$；

$$\begin{pmatrix} 1.0303 & 0.99030 \\ 0.99030 & 0.95285 \end{pmatrix}\begin{pmatrix} 1.2233 \\ 1.2461 \end{pmatrix} = \begin{pmatrix} 2.49438 \\ 2.39878 \end{pmatrix}$$

$$r_3 = b - A\begin{pmatrix} x_{31} \\ x_{32} \end{pmatrix} = \begin{pmatrix} 0.00002 \\ 0.00002 \end{pmatrix} \qquad \Rightarrow \Delta x_3 = (0,\quad 0.00002019\ 59)^T.$$

7. $x_4 = x_3 + \Delta x_3 = (1.2233\ ,\ 1.24612\ )^T.$

如果x_k 需要更多的数位，迭代可以继续.

小 结

定义：对非奇异矩阵A，称乘积$\|A\|\,\|A^{-1}\|$为矩阵A的**条件数**，记为 $\mathrm{cond}(A)=\|A\|\,\|A^{-1}\|$

定义：对线性方程组$Ax=b$，若$cond(A)$相对很大,则称$Ax=b$是病态的线性方程组(也称A是病态矩阵)；

若$cond(A)$相对很小，则称$Ax=b$是良态的线性方程组(也称A是良态矩阵).

作　业

❖ 教材第46页习题11、12.