5.4 误差分析–矩阵的条件数

朱升峰

华东师范大学 数学科学学院

2022 年 11 月

误差

求解线性方程组（A 非奇异）

$$Ax = b$$

- 测量误差
- 舍入误差

A 有扰动 δA 或 b 有扰动 δb，

实际求解

$$Ay = b + \delta b$$

或

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

估计误差 $x - y$.

例

例. 设有线性方程组

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

记为 $Ax = b$, 它的精确解为 $x = (2, 0)^{\mathrm{T}}$. 现在考虑常数项的微小变化对线性方程组解的影响, 线性方程组

$$\begin{pmatrix} 1 & 1 \\ 1 & 1.0001 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2.0001 \end{pmatrix}$$

表示为

$$A(x + \delta x) = b + \delta b,$$

其中 $\delta b = (0, 0.0001)^T$, $y = x + \delta x$, $Ax = b$。此方程组解为 $x + \delta x = (1, 1)^T$, 即 b 的小扰动造成解的变化 δx 较大。

矩阵的条件数

定义 1

如果矩阵 $\boldsymbol{A}$ 或常数项 $\boldsymbol{b}$ 的微小变化, 引起线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 解的巨大变化, 称此线性方程组为“病态”方程组, 矩阵 $\boldsymbol{A}$ 称为“病态”矩阵, 否则称方理组为“良态”方程组, $\boldsymbol{A}$ 称为“良态”矩阵.

矩阵“病态”性是其本身的特性，研究刻画矩阵“病态”性质的量。

$$Ax=b$$

研究 A 或 b 的微小扰动对解的影响。
设 b 有微小扰动 δb, 相应解为 $x+\delta x$, 则

$$A(x+\delta x)=b+\delta b,$$

则

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

从而

$$\|\delta x\|\leqslant\|A^{-1}\|\|\delta b\|,$$

b 的扰动对解 x 的影响

又由线性方程组 $Ax = b$ 得

$$\|b\| \leqslant \|A\|\|x\|$$

即

$$\frac{1}{\|x\|} \leqslant \frac{\|A\|}{\|b\|} \quad (b \neq 0)$$

综上有

定理 2

设 A 为非奇异阵，$Ax = b \neq 0$ 且

$$A(x + \delta x) = b + \delta b,$$

则

$$\frac{\|\delta x\|}{\|x\|} \leqslant \|A^{-1}\|\|A\|\frac{\|\delta b\|}{\|b\|}.$$

A 的扰动对解 x 的影响

设 $\boldsymbol{b}$ 是精确的，$\boldsymbol{A}$ 有微小误差 (扰动) $\delta\boldsymbol{A}$，解为 $\boldsymbol{x}+\delta\boldsymbol{x}$，则

$$\begin{cases} (\boldsymbol{A}+\delta\boldsymbol{A})(\boldsymbol{x}+\delta\boldsymbol{x}) = \boldsymbol{b}, \\ (\boldsymbol{A}+\delta\boldsymbol{A})\delta\boldsymbol{x} = -(\delta\boldsymbol{A})\boldsymbol{x}. \end{cases} \tag{1}$$

如果 $\delta\boldsymbol{A}$ 不受限制的话，$\boldsymbol{A}+\delta\boldsymbol{A}$ 可能奇异，而

$$(\boldsymbol{A}+\delta\boldsymbol{A}) = \boldsymbol{A}\left(\boldsymbol{I}+\boldsymbol{A}^{-1}\delta\boldsymbol{A}\right), \tag{2}$$

由定理知，当 $\left\|\boldsymbol{A}^{-1}\delta\boldsymbol{A}\right\|<1$ 时，$\left(\boldsymbol{I}+\boldsymbol{A}^{-1}\delta\boldsymbol{A}\right)^{-1}$ 存在. 由 (1)-(2)有

$$\delta\boldsymbol{x} = -\left(\boldsymbol{I}+\boldsymbol{A}^{-1}\delta\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{-1}(\delta\boldsymbol{A})\boldsymbol{x},$$

因此

$$\|\delta\boldsymbol{x}\| \leqslant \frac{\left\|\boldsymbol{A}^{-1}\right\|\|\delta\boldsymbol{A}\|\|\boldsymbol{x}\|}{1-\left\|\boldsymbol{A}^{-1}(\delta\boldsymbol{A})\right\|}.$$

设 $\left\|\boldsymbol{A}^{-1}\right\|\|\delta\boldsymbol{A}\|<1$，即得

$$\frac{\|\delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leqslant \frac{\left\|\boldsymbol{A}^{-1}\right\|\|\boldsymbol{A}\|\frac{\|\delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|}}{1-\left\|\boldsymbol{A}^{-1}\right\|\|\boldsymbol{A}\|\frac{\|\delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|}}. \tag{3}$$

从而证明了

定理 3

设 $\boldsymbol{A}$ 为非奇异阵, $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \neq \boldsymbol{0}$, 且

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

如果 $\left\|\boldsymbol{A}^{-1}\right\| \|\delta\boldsymbol{A}\| < 1$, 则

$$\frac{\|\delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leqslant \frac{\left\|\boldsymbol{A}^{-1}\right\| \|\boldsymbol{A}\| \frac{\|\delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|}}{1 - \left\|\boldsymbol{A}^{-1}\right\| \|\boldsymbol{A}\| \frac{\|\delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|}}. \tag{4}$$

综上，量 $\left\|\boldsymbol{A}^{-1}\right\| \|\boldsymbol{A}\|$ 刻画了解对原始数据变化的灵敏程度，即刻画了方程组的“病态”程度。

定义 4

设 A 为非奇异阵，称数

$$\operatorname{cond}(A)_v = \|A^{-1}\|_v \|A\|_v$$

为矩阵 A 的条件数, 其中一般 $v = 1, 2, \infty$。

条件数

当条件数相对大即 cond$(A) >> 1$, 则线性方程组 $Ax = b$ 是病态的;

当条件数相对小时, 则线性方程组 $Ax = b$ 是良态的。

条件数越大，方程组的病态程度越严重，越难以用一般的计算方法求得较准确的解。

通常使用的条件数有 (1) $\operatorname{cond}(\boldsymbol{A})_\infty = \left\|\boldsymbol{A}^{-1}\right\|_\infty \|\boldsymbol{A}\|_\infty$;
(2) $\boldsymbol{A}$ 的谱条件数

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

当 A 为对称矩阵时

$$\operatorname{cond}(\boldsymbol{A})_2 = \frac{|\lambda_1|}{|\lambda_n|},$$

其中 λ_1, λ_n 为 $\boldsymbol{A}$ 的绝对值最大和绝对值最小的特征值.

条件数的性质

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

$$\operatorname{cond}(\boldsymbol{A})_v = \left\|\boldsymbol{A}^{-1}\right\|_v \|\boldsymbol{A}\|_v \geqslant \left\|\boldsymbol{A}^{-1}\boldsymbol{A}\right\|_v = 1;$$

(2) 设 $\boldsymbol{A}$ 为非奇异阵且 $c \neq 0$ (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 $\operatorname{cond}(\boldsymbol{A})_2 = 1$; 如果 $\boldsymbol{A}$ 为非奇异矩阵, $\boldsymbol{R}$ 为正交矩阵, 则

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

例 已知 Hilbert 矩阵

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

计算 H_3 的 ∞ 范数对应的条件数。

解

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

(1) 计算 $\boldsymbol{H}_3$ 的条件数 cond $(\boldsymbol{H}_3)_\infty$:

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

所以 $\operatorname{cond}(\boldsymbol{H}_3)_\infty = 748$. 同样可计算

$$\operatorname{cond}(\boldsymbol{H}_6)_\infty = 2.9 \times 10^7, \quad \operatorname{cond}(\boldsymbol{H}_7)_\infty = 9.85 \times 10^8.$$

当 n 愈大时, $\boldsymbol{H}_n$ 矩阵病态愈严重.

(2) 考虑线性方程组

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

设 $\boldsymbol{H}_3$ 及 $\boldsymbol{b}$ 有微小误差 (取 3 位有效数字) 有

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

简记为 $(\boldsymbol{H}_3+\delta\boldsymbol{H}_3)(\boldsymbol{x}+\delta\boldsymbol{x})=\boldsymbol{b}+\delta\boldsymbol{b}$. 线性方程组 $\boldsymbol{H}_3\boldsymbol{x}=\boldsymbol{b}$ 与线性方程组 (5.8) 的精确解分别为
$\boldsymbol{x}=(1,1,1)^{\mathrm{T}},\boldsymbol{x}+\delta\boldsymbol{x}=(1.089512538,0.487967062,1.491002798)^{\mathrm{T}}$. 于是

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

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

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

文就是说 $\boldsymbol{H}_3$ 与 $\boldsymbol{b}$ 相对误差不超过 0.3%, 而引起解的相对误差超过 50%

发现及处理病态

要判别一矩阵是否病态需计算条件数 $\mathrm{cond}(\boldsymbol{A}) = \left\|\boldsymbol{A}^{-1}\right\| \|\boldsymbol{A}\|$, 而计算 $\boldsymbol{A}^{-1}$ 较困难, 在实际计算中发现病态情况:

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

(2) 系数矩阵的行列式值相对说很小, 或系数矩阵某些行近似线性相关, 这时 $\boldsymbol{A}$ 可能病态.

(3) 系数矩阵 $\boldsymbol{A}$ 元素间数量级相差很大, 并且无一定规则, $\boldsymbol{A}$ 可能病态.

用选主元素的消去法不能解决病态问题, 对于病态方程组可采用高精度 (如双倍字长) 的算术运算或采用预处理方法. 即将求解 $Ax = b$ 转化为等价线性方程组

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

选择非奇异矩阵 P, Q 使

$$\mathrm{cond}(PAQ) < \mathrm{cond}(A).$$

一般选择预处理子 P, Q 为对角阵或者三角矩阵.

当矩阵 $\boldsymbol{A}$ 的元素大小不均时, 对 $\boldsymbol{A}$ 的行 (或列) 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. 这种方法不能保证 $\boldsymbol{A}$ 的条件数一定得到改善.

例 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 $\mathrm{cond}(\boldsymbol{A})_\infty$.

解:

$$\boldsymbol{A}=\begin{pmatrix} 1 & 10^4 \\ 1 & 1 \end{pmatrix}, \quad \boldsymbol{A}^{-1}=\frac{1}{10^4-1}\begin{pmatrix} -1 & 10^4 \\ 1 & -1 \end{pmatrix}$$

$$\operatorname{cond}(\boldsymbol{A})_\infty=\frac{\left(1+10^4\right)^2}{10^4-1}\approx 10^4.$$

现在 $\boldsymbol{A}$ 的第一行引进比例因子. 如用 $s_1=\max_{1\leqslant i\leqslant 2}|a_{1i}|=10^4$ 除第一个方程, 得 $\boldsymbol{A}'\boldsymbol{x}=\boldsymbol{b}'$, 即

$$\begin{pmatrix} 10^{-4} & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

而

$$\left(\boldsymbol{A}'\right)^{-1}=\frac{1}{1-10^{-4}}\begin{pmatrix} -1 & 1 \\ 1 & -10^{-4} \end{pmatrix}$$

于是

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

当用列主元消去法解线性方程组时 (计算到三位数字),

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

于是得到坏结果: $x_2 = 1, x_1 = 0$.

现用列主元消去法解线性方程组, 得到

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

从而得到较好的计算解 $x_1 = 1, x_2 = 1$.

设 $\bar{x}$ 为线性方程组 $Ax=b$ 的近似解, 可计算 $\bar{x}$ 的剩余向量 $r=b-A\bar{x}$, 当 r 很小时, $\bar{x}$ 是否为 $Ax=b$ 较好的近似解呢?

定理 5

事后误差估计) 设 A 为非奇异矩阵, x 是线性方程组 $Ax=b\neq 0$ 的精确解. 再设 $\bar{x}$ 是此方程组的近似解, $r=b-A\bar{x}$, 则

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

证明: 由 $x-\bar{x}=A^{-1}r$, 得

$$\|x-\bar{x}\|\leqslant \left\|A^{-1}\right\|\|r\|,$$

又有

$$\|b\|=\|Ax\|\leqslant\|A\|\cdot\|x\|,\quad \frac{1}{\|x\|}\leqslant\frac{\|A\|}{\|b\|},$$

由以上两式得结论成立.