第二讲 线性方程组的直接解法

扰动分析 • 解的改进

目录

2-3 扰动分析

2.3 扰动分析

https://math.ecnu.edu.cn/~jypan/Teaching/SC

2-3-1 矩阵条件数

为什么要考虑扰动分析

除了数据误差和截断误差外, 在用计算机进行数值计算时, 每次运算 (加减乘除等) 还会产生舍入误差. 对于大规模问题, 实际运算次数往往非常巨大, 因此直接分析舍入误差比较复杂, 而且得到的结果往往不一定能反映实际计算情况. 当前一种比较实用的误差分析方法是将舍入误差看作是对原始数据的扰动, 即将其归结到数据误差中, 这样就可以在很大程度上简化误差分析过程. 这种误差分析方法称为**向后误差分析**.

病态线性方程组/病态矩阵

考虑线性方程组 $Ax=b$, 如果 A 或 b 的微小变化 (也称为 **扰动**) 会导致解的巨大变化, 则称此线性方程组是 **病态** 的, 反之则是 **良态** 的.

病态线性方程组举例

例 考虑线性方程组 $Ax=b$, 其中 $A=\begin{bmatrix}1 & 1\\ 1 & 1.0001\end{bmatrix}$, $b=\begin{bmatrix}2\\2\end{bmatrix}$. $\Longrightarrow$ 解为 $x=\begin{bmatrix}2\\0\end{bmatrix}$.

病态线性方程组举例

例 考虑线性方程组 $Ax=b$, 其中 $A=\begin{bmatrix}1 & 1\\ 1 & 1.0001\end{bmatrix}$, $b=\begin{bmatrix}2\\ 2\end{bmatrix}$. $\Longrightarrow$ 解为 $x=\begin{bmatrix}2\\ 0\end{bmatrix}$.

☞ 若 b 的第二个元素出现小扰动, 变为 $b=\begin{bmatrix}2\\ 2.0001\end{bmatrix}$, 则解变为 $x=\begin{bmatrix}1\\ 1\end{bmatrix}$.

当右端项出现细微变化时, 解会出现很大的变化, 因此该线性方程组是病态的.

病态线性方程组举例

例 考虑线性方程组 $Ax=b$, 其中 $A=\begin{bmatrix}1 & 1\\ 1 & 1.0001\end{bmatrix}$, $b=\begin{bmatrix}2\\2\end{bmatrix}$. $\Longrightarrow$ 解为 $x=\begin{bmatrix}2\\0\end{bmatrix}$.

☞ 若 b 的第二个元素出现小扰动, 变为 $b=\begin{bmatrix}2\\2.0001\end{bmatrix}$, 则解变为 $x=\begin{bmatrix}1\\1\end{bmatrix}$.

当右端项出现细微变化时, 解会出现很大的变化, 因此该线性方程组是病态的.

怎样判断一个线性方程组是否病态?

对于线性方程组而言, 问题是否变态主要取决于 系数矩阵是否病态.

判断一个矩阵是否病态的一个重要指标是 矩阵条件数.

矩阵条件数

定义 设 A 非奇异, $\|\cdot\|$ 是任一算子范数, 则称

$$\kappa(A) \triangleq \|A^{-1}\| \cdot \|A\|$$

为 A 的 **条件数**.

矩阵条件数

定义 设 A 非奇异, $\|\cdot\|$ 是任一算子范数, 则称

$$\kappa(A) \triangleq \|A^{-1}\| \cdot \|A\|$$

为 A 的 **条件数**.

✍ 常用的矩阵条件数有

$$\kappa_2(A) \triangleq \|A^{-1}\|_2 \cdot \|A\|_2, \quad \kappa_1(A) \triangleq \|A^{-1}\|_1 \cdot \|A\|_1, \quad \kappa_\infty(A) \triangleq \|A^{-1}\|_\infty \cdot \|A\|_\infty.$$

✍ $\kappa_2(A)$ 也称为 **谱条件数**, 当 A 对称时, 有

$$\boxed{\kappa_2(A) = \frac{\max\limits_{1\leq i\leq n} |\lambda_i|}{\min\limits_{1\leq i\leq n} |\lambda_i|}}$$

例 计算前例中的系数矩阵 A 的条件数 $\kappa_\infty(A)$ 和 $\kappa_2(A)$. $\left(A=\begin{bmatrix}1 & 1\\ 1 & 1.0001\end{bmatrix}\right)$

解. 通过直接计算可得 $A^{-1}=\begin{bmatrix}10001 & -10000\\ -10000 & 10000\end{bmatrix}$. 所以

$$\kappa_\infty(A)=\|A\|_\infty\cdot\|A^{-1}\|_\infty\approx 4\times 10^4.$$

由于 A 是对称的, 且 A 的特征值为

$$\lambda(A)=\frac{2.0001\pm\sqrt{2.0001^2-0.0004}}{2}>0,$$

所以

$$\kappa_2(A)=\frac{\lambda_{\max}}{\lambda_{\min}}\approx 4\times 10^4.$$

□

例 (条件数与特征值) 已知矩阵 $A=\begin{bmatrix}1 & -0.5 & \cdots & -0.5\\ & 1 & \ddots & \vdots\\ & & \ddots & -0.5\\ & & & 1\end{bmatrix}\in\mathbb{R}^{n\times n}$.

直接计算可知 $A^{-1}=\begin{bmatrix}1 & \frac{1}{2} & \frac{1.5}{2} & \frac{1.5^2}{2} & \cdots & \frac{1.5^{n-2}}{2}\\ & 1 & \frac{1}{2} & \frac{1.5}{2} & \ddots & \vdots\\ & & 1 & \ddots & \ddots & \frac{1.5^2}{2}\\ & & & \ddots & \frac{1}{2} & \frac{1.5}{2}\\ & & & & 1 & \frac{1}{2}\\ & & & & & 1\end{bmatrix}$.

直接观察可知, $\kappa_1(A)$ 和 $\kappa_\infty(A)$ 会随着 n 的增大而快速增长.

下表中是 $n=10,20,30,40,50$ 时 $\kappa_1(A)$ 和 $\kappa_2(A)$ 的值.

n	10	20	30	40	50
$\kappa_1(A)$	2.1×10^2	2.3×10^4	2.0×10^6	1.5×10^8	1.1×10^{10}
$\kappa_2(A)$	6.3×10	7.6×10^3	6.8×10^5	5.5×10^7	3.9×10^9

下表中是 $n = 10, 20, 30, 40, 50$ 时 $\kappa_1(A)$ 和 $\kappa_2(A)$ 的值.

n	10	20	30	40	50
$\kappa_1(A)$	2.1×10^2	2.3×10^4	2.0×10^6	1.5×10^8	1.1×10^{10}
$\kappa_2(A)$	6.3×10	7.6×10^3	6.8×10^5	5.5×10^7	3.9×10^9

注记

- 对称矩阵的 (谱) 条件数取决于特征值, 但该结论不适用于一般矩阵.
- 条件数是矩阵是否病态的一个重要指标, 但两者并不完全等价.

矩阵条件数的基本性质

引理 条件数具有以下性质:

- $\kappa(A) \geq 1, \quad \forall\ A \in \mathbb{R}^{n\times n}$.
- 对任意非零常数 $\alpha \in \mathbb{R}$, 都有

$$\kappa(\alpha A) = \kappa(A).$$

- 对任意正交矩阵 $Q \in \mathbb{R}^{n\times n}$, 有 $\kappa_2(Q) = 1$.
- 设 Q 是正交矩阵, 则对任意矩阵 A 有

$$\kappa_2(QA) = \kappa_2(A) = \kappa_2(AQ).$$

(留作练习)

2-3-2 条件数与扰动分析

考虑线性方程组 $Ax = b$, 真解 $x_* = A^{-1}b$

- 假定系数矩阵 A 是精确的, 而右端项 b 有个微小扰动 δb.

2-3-2 条件数与扰动分析

考虑线性方程组 $Ax = b$, 真解 $x_* = A^{-1}b$

- 假定系数矩阵 A 是精确的, 而右端项 b 有个微小扰动 δb.
- 因此我们实际求解的是线性方程组

$$\boxed{Ax = b + \delta b}$$

我们称这个方程组为 **扰动方程组**, 其解记为 $\tilde{x}$, 并记

$$\boxed{\delta x \triangleq \tilde{x} - x_*}$$

δx 与 x_* 之间的关系

定理 设 $\|\cdot\|$ 是任一向量范数（当该范数作用在矩阵上时就是相应的算子范数），若 A 是精确的, b 有个小扰动 δb, 则有

$$\boxed{\frac{\|\delta x\|}{\|x_*\|} \leq \|A^{-1}\| \cdot \|A\| \cdot \frac{\|\delta b\|}{\|b\|}}$$

(板书)

δx 与 x_* 之间的关系

定理 设 $\|\cdot\|$ 是任一向量范数（当该范数作用在矩阵上时就是相应的算子范数），若 A 是精确的, b 有个小扰动 δb, 则有

$$\boxed{\frac{\|\delta x\|}{\|x_*\|} \leq \|A^{-1}\| \cdot \|A\| \cdot \frac{\|\delta b\|}{\|b\|}}$$

(板书)

✍ 由于右端项的扰动而产生的解的相对误差, 大约被放大 $\|A^{-1}\| \cdot \|A\|$ 倍.
这个倍数正好是系数矩阵的 **条件数**.

✍ 需要指出的是, 这是最坏的情况.

如果 A 也存在扰动

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

首先需要确保 $A+\delta A$ 的非奇异性.

引理 设 $\|\cdot\|$ 是 $\mathbb{R}^{n\times n}$ 上的任一算子范数, 若 $\|B\|<1$, 则 $I+B$ 非奇异, 且

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

(板书)

如果 A 也存在扰动

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

首先需要确保 $A+\delta A$ 的非奇异性.

引理 设 $\|\cdot\|$ 是 $\mathbb{R}^{n\times n}$ 上的任一算子范数, 若 $\|B\|<1$, 则 $I+B$ 非奇异, 且

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

(板书)

假定 $\|\delta A\|$ 很小, 满足 $\|A^{-1}\|\cdot\|\delta A\|<1$, 则 $\|A^{-1}\delta A\| \leq \|A^{-1}\|\cdot\|\delta A\|<1$.

因此 $A+\delta A = A(I+A^{-1}\delta A)$ 非奇异且 $\|(A+\delta A)^{-1}\| \leq \frac{\|A\|}{1-\|A^{-1}\|\,\|\delta A\|}$.

如果 A 也存在扰动 (续)

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

定理 设 $\|A^{-1}\| \cdot \|\delta A\| < 1$, 则

$$\frac{\|\delta x\|}{\|x_*\|} \le \frac{\kappa(A)}{1-\kappa(A)\cdot\frac{\|\delta A\|}{\|A\|}}\left(\frac{\|\delta A\|}{\|A\|}+\frac{\|\delta b\|}{\|b\|}\right)$$

当 $\delta b = 0$ 时, 有

$$\frac{\|\delta x\|}{\|x_*\|} \le \frac{\kappa(A)}{1-\kappa(A)\cdot\frac{\|\delta A\|}{\|A\|}}\cdot\frac{\|\delta A\|}{\|A\|}$$

(板书)

δx 与 $\tilde{x}$ 之间的关系

由于 x_* 通常是未知的, 因此一个更加实际的情况是考虑 δx 与 $\tilde{x}$ 之间的关系.

定理 设 $\|\cdot\|$ 是任一向量范数（当该范数作用在矩阵上时就是相应的算子范数），则 δx 与 $\tilde{x}$ 满足下面的关系式

$$\frac{\|\delta x\|}{\|\tilde{x}\|} \leq \|A^{-1}\| \cdot \|A\| \left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|A\| \cdot \|\tilde{x}\|} \right).$$

当 $\delta b = 0$ 时, 有

$$\boxed{\frac{\|\delta x\|}{\|\tilde{x}\|} \leq \kappa(A) \frac{\|\delta A\|}{\|A\|}}$$

(板书)

δx 与残量之间的关系

前面的结论具有理论指导作用, 但通常无法获得 δA 和 δb 的大小, 因此很难用来估计 δx.

δx 与残量之间的关系

前面的结论具有理论指导作用, 但通常无法获得 δA 和 δb 的大小, 因此很难用来估计 δx.

记 **残量** (**残差**, residual) 为 $r \triangleq b - A\tilde{x}$, 则有

$$\delta x = \tilde{x} - x_* = \tilde{x} - A^{-1}b = A^{-1}(A\tilde{x} - b) = -A^{-1}r,$$

所以可得

$$\boxed{\|\delta x\| \leq \|A^{-1}\| \cdot \|r\|}$$

✍ 这个估计式的优点是不需要知道 δA 和 δb 的大小. 而且在实际计算中, r 通常是可计算的, 因此该估计式比较实用.

例 Hilbert 矩阵是一个典型的病态矩阵, 其定义如下

$$H_n = \begin{bmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\ \vdots & \vdots & & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \cdots & \frac{1}{2n-1} \end{bmatrix}, \quad \text{即} \quad H_n = [h_{ij}],\ h_{ij} = \frac{1}{i+j-1},\ i,j = 1,2,\ldots,n.$$

可以验证 H_n 是对称正定的. 通过计算可知

$$\kappa_\infty(H_2) = 27, \quad \kappa_\infty(H_3) = 748, \quad \kappa_\infty(H_4) = 28375, \quad \kappa_\infty(H_{10}) \approx 3.5 \times 10^{13}.$$

例 考虑线性方程组

$$H_n x = b,$$

设精确解为 $x_* = [1, 1, \ldots, 1]^{\mathsf{T}}$, 计算出右端项 b, 然后用 Cholesky 分解求解该线性方程组, 求得的近似解记为 $\tilde{x}$.

例 考虑线性方程组

$$H_n x = b,$$

设精确解为 $x_* = [1, 1, \ldots, 1]^{\mathsf{T}}$, 计算出右端项 b, 然后用 Cholesky 分解求解该线性方程组, 求得的近似解记为 $\tilde{x}$.

对于不同的 n, 下表中列出了近似解的误差.

n	5	10	20	30
$\|\tilde{x} - x_*\|_\infty$	2.3×10^{-11}	6.2×10^{-4}	> 7.8	> 42.8
$\kappa_\infty(H_n)$	9.4×10^5	3.5×10^{13}	$> 10^{19}$	$> 10^{19}$

当 $n \geq 20$ 时, Cholesky 分解计算出来的近似解已经没有任何意义了.

2-4 扰动分析

2.4 解的改进

高精度运算

当矩阵 A 是病态时, 即使残量 $r = b - A\tilde{x}$ 很小, 所求得的数值解 $\tilde{x}$ 仍可能带有较大的误差. 此时需要通过一些方法来提高解的精度.

高精度运算

- 在计算中, 采用更高精度的运算.
  比如, 原始数据是单精度的, 但在计算时都采用双精度运算, 或者更高精度的运算.
- 但更高精度的运算会带来更大的开销.

矩阵元素缩放 (Scaling)

- 如果 A 的元素在数量级上相差很大, 则在计算过程中很可能会出现大数与小数的加减运算, 这样就可能会引入更多的舍入误差.
- 为了避免由于这种情况而导致的舍入误差, 我们可以在求解之前先对矩阵元素进行缩放 (Scaling), 即在矩阵两边同时乘以两个 **适当的对角矩阵**.

$$\boxed{Ax = b} \quad \Longleftrightarrow \quad \boxed{D_r^{-1} A D_c^{-1} y = D_r^{-1} b, \quad x = D_c^{-1} y}$$

迭代改进法

设近似解 $\tilde{x}$, 残量 $r = b - A\tilde{x}$.

基本思想与实施过程

(1) 构造修正向量 z 使得 $\tilde{x} + z$ 是原问题的精确解, 即

$$\boxed{A(\tilde{x} + z) = b}$$

也就是说, 向量 z 满足方程组

$$Az = b - A\tilde{x} = r.$$

这就是 **残量方程** (右端项是残量).

(2) 求得近似解 $\tilde{z}$, 得到新的近似解 $\tilde{x} + \tilde{z}$.

(3) 如果还不满足精度要求, 则可重复以上过程.

算法 迭代改进法

设 $PA = LU$, $\tilde{x}$ 是 $Ax = b$ 的近似解
while 近似解 $\tilde{x}$ 不满足精度要求, **do**
    计算 $r = b - A\tilde{x}$
    求解 $Ly = Pr$, 即 $y = L^{-1}Pr$
    求解 $Uz = y$, 即 $z = U^{-1}y$
    令 $\tilde{x} = \tilde{x} + z$
end while

每改进一次额外增加的运算量约为 $\mathcal{O}(n^2)$, 所以相对来讲还是比较经济的.

算法 迭代改进法

1: 设 $PA = LU$, $\tilde{x}$ 是 $Ax = b$ 的近似解
2: **while** 近似解 $\tilde{x}$ 不满足精度要求, **do**
3: 计算 $r = b - A\tilde{x}$
4: 求解 $Ly = Pr$, 即 $y = L^{-1}Pr$
5: 求解 $Uz = y$, 即 $z = U^{-1}y$
6: 令 $\tilde{x} = \tilde{x} + z$
7: **end while**

每改进一次额外增加的运算量约为 $\mathcal{O}(n^2)$, 所以相对来讲还是比较经济的.

- 为了提高计算精度, 在计算残量 r 时最好使用原始数据 A, 而不是 $P^{\mathsf{T}}LU$, 因此对 A 做 PLU 分解时需要保留原矩阵 A, 不能被 L 和 U 覆盖.
- 实际计算经验表明, 当 A 病态不是很严重时, 即 $\varepsilon_u \kappa_\infty(A) < 1$, 迭代法可以有效改进解的精度, 最后达到机器精度. 但 $\varepsilon_u \kappa_\infty(A) \geq 1$ 时, 一般没什么效果.

谢谢

THANK YOU