第一章 距离和范数（2）

矩阵的收敛性问题

> 对于一个 $n \times n$ 的方阵 A，其所有特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$ (这些特征值构成了矩阵的“谱”) 的绝对值的最大值，称为矩阵 A 的谱半径。

记作 $\rho(A)$ (读作 “rho A”):

$$\rho(A) = \max\{|\lambda_1|, |\lambda_2|, \dots, |\lambda_n|\}$$

简单来说，谱半径就是矩阵所有特征值中，绝对值最大的那个对应特征值的绝对值。

1. 当 A 为正规矩阵时，A 的 2 范数就等于 A 的谱半径；
2. A 的谱半径小于等于 A 的**任意一个**范数（如 1 范数、无穷范数）；
3. A 为**收敛矩阵的充要条件**：A 的谱半径 $\rho(A) < 1$

结合 2 和 3 可知，**若 A 存在任何一个矩阵范数 < 1，则 A 矩阵收敛。**

做题技巧：

- 对于矩阵 A，可以先判断其 1 范数和无穷范数是不是 < 1。如果 < 1，那么可以直接判断 A 收敛；
- 否则，考虑计算矩阵 A 的谱半径 $\rho(A)$，求解特征方程 $det(A - \lambda I) = 0$ 得到特征值，找到特征值中绝对值最大的对应的特征值绝对值 $\max\{|\lambda_1|, |\lambda_2|, \dots, |\lambda_n|\}$：
  - 如果 $\rho(A) < 1$，则矩阵 A 收敛；
  - 否则，矩阵 A 不收敛。

例题 1

> 判断下列矩阵是否收敛：
>
> $$A_1 = \frac{1}{5}\begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}$$

解答：

$\|A_1\|_1 = \frac{6}{5}$，无法判断收敛；$\|A_1\|_\infty = \frac{6}{5}$，无法判断收敛。因此，考虑计算 $\rho(A_1)$.

解特征方程 $\det(A_1 - \lambda I) = 0$：

- 计算

$$A_1 - \lambda I = \begin{pmatrix} \frac{2}{5} - \lambda & \frac{1}{5} \\ \frac{4}{5} & \frac{2}{5} - \lambda \end{pmatrix}$$

- 行列式为

$$\det(A_1 - \lambda I) = (\frac{2}{5} - \lambda)^2 - \frac{1}{5} \times \frac{4}{5} = \lambda^2 - \frac{4}{5}\lambda$$

- 另行列式为 0

$$\lambda^2 - \frac{4}{5}\lambda = 0 \to \lambda(\lambda - \frac{4}{5}) = 0$$

- 得到特征值

$$\lambda_1 = 0, \quad \lambda_2 = \frac{4}{5}$$

可以发现，绝对值最大特征值为 $|\lambda_2| = |\frac{4}{5}| < 1$，可知谱半径 $\rho(A_1) < 1$，故 矩阵 A_1 收敛。

> $$A_2 = \frac{1}{3}\begin{pmatrix} 1 & -1 \\ -1 & 3 \end{pmatrix}$$

解答:

$\|A_2\|_1 = \frac{4}{3}$，无法判断收敛；$\|A_2\|_\infty = \frac{4}{3}$，无法判断收敛。因此，考虑计算 $\rho(A_2)$.

解特征方程 $\det(A_2 - \lambda I) = 0$:

- 计算

$$A_2 - \lambda I = \begin{pmatrix} \frac{1}{3} - \lambda & -\frac{1}{3} \\ -\frac{1}{3} & 1 - \lambda \end{pmatrix}$$

- 行列式为

$$\det(A_2 - \lambda I) = (\frac{1}{3} - \lambda)(1 - \lambda) - (-\frac{1}{3})(-\frac{1}{3}) = \lambda^2 - \frac{4}{3}\lambda + \frac{2}{9}$$

- 另行列式为 0

$$\lambda^2 - \frac{4}{3}\lambda + \frac{2}{9} = 0$$

- 得到特征值

$$\lambda_1 = \frac{2 + \sqrt{2}}{3}, \quad \lambda_2 = \frac{2 - \sqrt{2}}{3}$$

可以发现，绝对值最大特征值为 $|\lambda_1| = |\frac{2+\sqrt{2}}{3}| \geq 1$，可知谱半径 $\rho(A_2) \geq 1$，故矩阵 A_2 不收敛。

矩阵的条件数及应用

矩阵的条件数：线性方程组 $Ax=b$ 对输入的微小扰动的敏感程度。（系数 A 和 右端项 b，x 为方程组的解）

矩阵条件数定义

对于可逆矩阵 A，在某个范数 $\|\cdot\|$ 下的条件数定义为：

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

例如，在 2 范数下，有：

$$cond_2(A)=\frac{\sigma_{\max}(A)}{\sigma_{\min}(A)}$$

即最大奇异值和最小奇异值的比。奇异值为 A^TA 的特征值的非负平方根。
关键性质：**可逆矩阵 A 的特征值，和 A^{-1} 的特征值互为倒数。**

特别的，如果 A 为正规矩阵且可逆，$cond_2(A)$ 为矩阵 A 的最大特征值绝对值和最小特征值绝对值的比：

$$cond_2(A)=\frac{\max_i|\lambda_i(A)|}{\min_i|\lambda_i(A)|}$$

矩阵条件数在误差估计中的应用

1. 同时考虑系数 A 和右端项 b 的一般情形

$$\frac{\|\hat{x}-x\|}{\|x\|}\le\frac{cond(A)}{1-cond(A)\frac{\|\delta A\|}{\|A\|}}(\frac{\|\delta A\|}{\|A\|}+\frac{\|\delta b\|}{\|b\|})$$

   含义：**解的相对误差**被条件数 $cond(A)\times$ （系数矩阵相对误差 + 右端项相对误差）通过因子 $(1-cond(A)\frac{\|\delta A\|}{\|A\|})^{-1}$ 放大。

2. 只考虑系数矩阵 A 的误差（$\delta b=0$）

$$\frac{\|\hat{x}-x\|}{\|x\|}\le\frac{cond(A)}{1-cond(A)\frac{\|\delta A\|}{\|A\|}}(\frac{\|\delta A\|}{\|A\|})$$

例题 1

> $A=\begin{pmatrix}-1 & i & 0\\ -i & 0 & -i\\ 0 & i & -1\end{pmatrix},\quad \delta(A)\in C^{3\times3},\quad 0\ne b\in C^{3\times3}$. 为使得线性方程组 $Ax=b$ 的解 x 与 $(A+\delta(A))x=b$ 的解 $\hat{x}$ 的相对误差 $\frac{\|\hat{x}-x\|_2}{\|x\|_2}\le10^{-4}$，问 $\frac{\|\delta(A)\|_2}{\|A\|_2}$ 应该不超过什么值？

解答：

由于 $A^H=A$，可知 A 为 Hermite 矩阵，也是正规矩阵。故 $\|A\|_2=\rho(A)$.

通过求解 A 的特征值得到谱半径进而得到 $\|A\|_2$，求解特征方程 $\det(A-\lambda I)=0$.

- 计算

$$A-\lambda I=\begin{pmatrix}-1-\lambda & i & 0\\ -i & -\lambda & -i\\ 0 & i & -1-\lambda\end{pmatrix}$$

- 行列式为

$$\begin{aligned}\det(A-\lambda I)&=-(1+\lambda)\cdot\det\begin{pmatrix}-\lambda & -i\\ i & -1-\lambda\end{pmatrix}-i\cdot\det\begin{pmatrix}-i & -i\\ 0 & -1-\lambda\end{pmatrix}+0\\&=-(1+\lambda)[(-\lambda)(-1-\lambda)+i^2]-i\cdot(-i)\cdot(-1-\lambda)\\&=-\lambda(\lambda^2+\lambda-1)+(1+\lambda)\\&=(1+\lambda)[1-(\lambda^2+\lambda-1)]\end{aligned}$$

- 另行列式为 0

$$\det(A-\lambda I)=-(1+\lambda)(\lambda^2+\lambda-2)=0$$

$$(1+\lambda)(\lambda^2+\lambda-2)=0$$

得到特征值为：$\lambda_1=-2,\quad \lambda_2=-1,\quad \lambda_3=1$，故 $\|A\|_2=\rho(A)=|\lambda_1|=2.$

由可逆矩阵 A 的特征值和 A^{-1} 的特征值互为倒数，可得 $\|A^{-1}\|_2=|\frac{1}{\lambda_3}|=1.$

- 计算条件数

$$cond_2(A)=\|A\|_2\cdot\|A^{-1}\|_2=2$$

- 代入误差估计公式

$$\frac{\|\hat{x}-x\|}{\|x\|}\leq\frac{cond_2(A)}{1-cond_2(A)\frac{\|\delta A\|}{\|A\|}}(\frac{\|\delta A\|}{\|A\|})\leq 10^{-4}$$

$$\frac{2}{1-2\frac{\|\delta A\|}{\|A\|}}(\frac{\|\delta A\|}{\|A\|})\leq 10^{-4}$$

由此可得：

$$\frac{\|\delta A\|}{\|A\|}\leq\frac{1}{2.0002}\times 10^{-4}\leq\frac{1}{2}\times 10^{-4}=5\times 10^{-5}$$

例题 2

> 假设 x 和 y 分别是以下方程组的解，试估计解的相对误差 $\frac{\|x-y\|_1}{\|x\|_1}$。（保留到小数点后第 4 位）
>
> $$\begin{pmatrix}2 & 0 & 0\\ 0 & 2 & 1\\ 1 & 1 & 2\end{pmatrix}x=\begin{pmatrix}1\\2\\1\end{pmatrix},\quad \begin{pmatrix}2.05 & 0 & 0\\ 0 & 2.02 & 1.01\\ 1.02 & 1.01 & 2.02\end{pmatrix}y=\begin{pmatrix}1.001\\2.002\\1.001\end{pmatrix}$$

解答：

- 由题意可知

$$A=\begin{pmatrix}2 & 0 & 0\\ 0 & 2 & 1\\ 1 & 1 & 2\end{pmatrix},\quad b=\begin{pmatrix}1\\2\\1\end{pmatrix}$$

$$\delta A = A^{'} - A = \begin{pmatrix} 0.05 & 0 & 0 \\ 0 & 0.02 & 0.01 \\ 0.02 & 0.01 & 0.02 \end{pmatrix}, \quad \delta b = b^{'} - b = \begin{pmatrix} 0.001 \\ 0.002 \\ 0.001 \end{pmatrix}$$

- 计算 $\|A\|_1$

$$\|A\|_1 = 3$$

- 计算 $\|A^{-1}\|_1$

  首先，构造增广矩阵，通过行初等变换得到 A^{-1}：

  增广矩阵为：

$$\left[\begin{array}{ccc|ccc} 2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{array}\right]$$

  初等行变换：

  1. 将第一行除以 2：

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 2 & 1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{array}\right]$$

  2. 将第三行减去第一行：

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 2 & 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & -\frac{1}{2} & 0 & 1 \end{array}\right]$$

  3. 将第二行除以 2：

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 2 & -\frac{1}{2} & 0 & 1 \end{array}\right]$$

  4. 将第三行减去第二行：

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{3}{2} & -\frac{1}{2} & -\frac{1}{2} & 1 \end{array}\right]$$

  5. 将第三行乘以 $\frac{2}{3}$：

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{array}\right]$$

  6. 将第二行减去 $\frac{1}{2}$ 倍第三行：

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{6} & \frac{2}{3} & -\frac{1}{3} \\ 0 & 0 & 1 & -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{array}\right]$$

  左侧已化为单位矩阵，右侧即为逆矩阵 A^{-1}，故求得 $\|A^{-1}\|_1$：

$$\|A^{-1}\|_1 = 1$$

- 计算条件数

$$cond_1(A) = \|A\|_1 \cdot \|A^{-1}\|_1 = 3$$

- 计算其它条件

$$\|\delta A\|_1 = 0.07, \quad \|\delta b\|_1 = 0.004, \quad \|b\|_1 = 4$$

- 代入误差估计公式

$$\begin{aligned}\frac{\|x-y\|_1}{\|x\|_1} &\leq \frac{cond_1(A)}{1 - cond_1(A)\frac{\|\delta A\|_1}{\|A\|_1}}(\frac{\|\delta A\|_1}{\|A\|_1} + \frac{\|\delta b\|_1}{\|b\|_1}) \\ &= \frac{3}{1 - 3 \times \frac{0.07}{3}}(\frac{0.07}{3} + \frac{0.004}{4}) \\ &= 0.0785\end{aligned}$$

- 由此可得

$$\frac{\|x-y\|_1}{\|x\|_1} \leq 0.0785$$