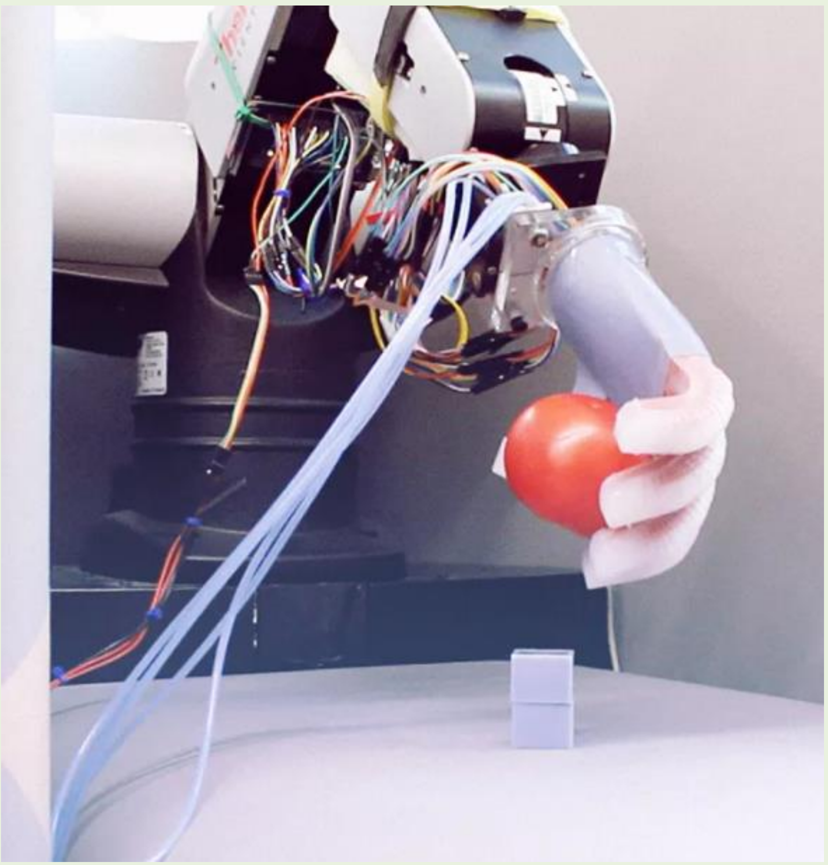
第五章

矩阵范数

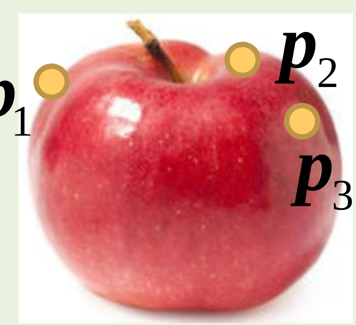
引导问题

$$\boldsymbol{h}_i = (x_i, y_i, z_i)$$



A

寻找旋转*Q*，使得*Q*和*A*最接近。



$$\boldsymbol{p}_i = (u_i, v_i, w_i)$$

提 纲

- **1. 向量范数**
- 2. 方阵范数
- 3. 一般矩阵范数
- 4. 矩阵的谱半径
- 5. 矩阵的条件数

向量(矩阵)的数字特征，某种意义上相当于实数和复数的绝对值。是研究序列、级数极限的基础。

定义(向量范数) 对任一向量 $x\in C^n$, 按照一定规则确定一个实数与它对应, 该实数记为$\|x\|$, 若$\|x\|$满足下面三个性质:

1) $\|x\|\geq 0$；$\|x\|=0$当且仅当$x=0$(零向量)　**正定性**

2) 对任意复数a, $\|ax\|=|a|\,\|x\|$　**齐次性**

3) 对任意向量 $x, y\in C^n$, $\|x+y\| \leq \|x\|+\|y\|$　**三角不等式**

则称该实数$\|x\|$为向量x的范数.

线性空间 ⟶ 赋范线性空间

- C^n中常用的三种范数

曼哈顿范数、出租车范数

- 1-范数

$$\|x\|_1=|x_1|+|x_2|+\ldots+|x_n|=\sum_{i=1}^{n}|x_i|$$

- 2-范数

$$\|x\|_2=\sqrt{|x_1|^2+|x_2|^2+\ldots+|x_n|^2}=\left(\sum_{i=1}^{n}|x_i|^2\right)^{\frac{1}{2}}$$

欧几里得范数、向量的模

- ∞-范数

$$\|x\|_\infty=\max\{|x_1|,|x_2|,\ldots,|x_n|\}=\max_{1\le i\le n}\{|x_i|\}$$

其中 $x_1, x_2, \ldots, x_n$ 分别是 x 的n个分量.

最大值范数

酉变换不改变向量的2范数。

p范数 $p \geq 1$

$$\boxed{\|x\|_p = \sqrt[p]{|x_1|^p + |x_2|^p + \ldots + |x_n|^p} = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}}$$

$$\lim_{p \to \infty} \|x\|_p = \|x\|_\infty$$

意义

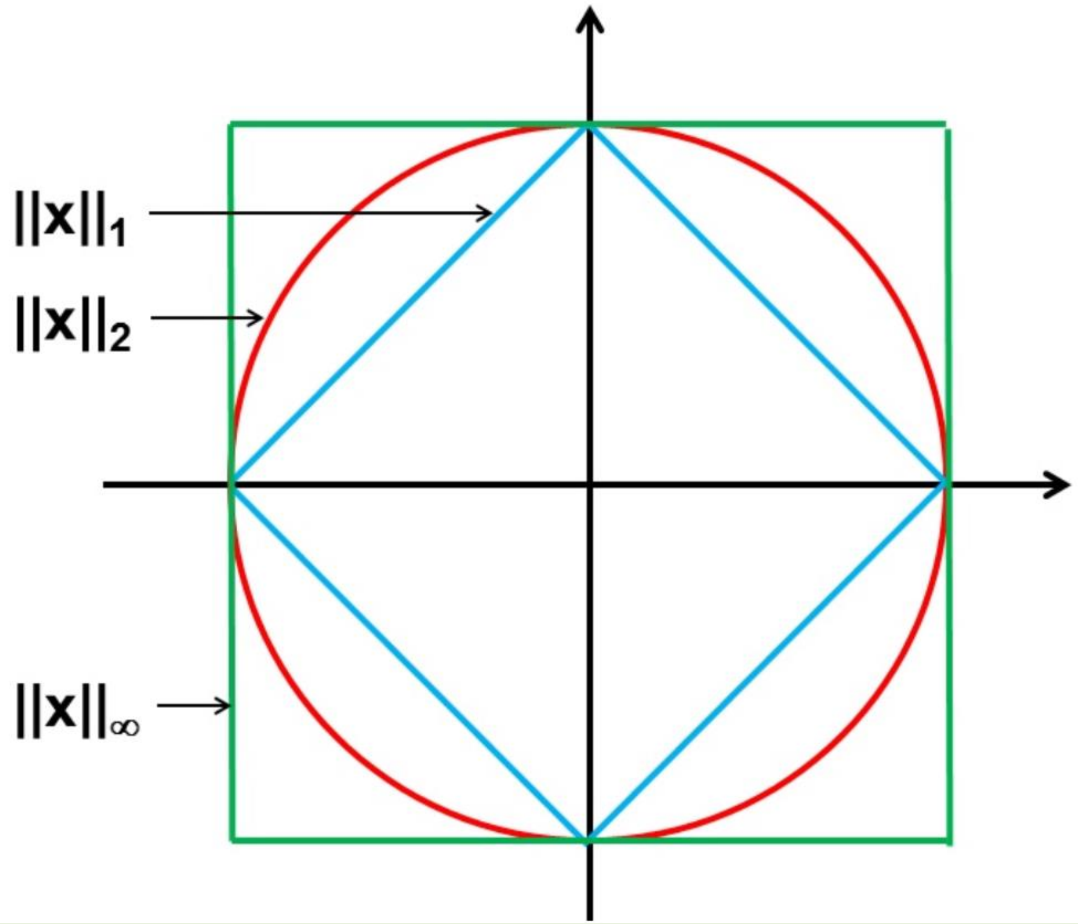
反映大小或距离，比较不同的向量

$(0,6)$ $\qquad \|(0,6)\|_2=6$

$(0,6)>(3,4)$

$(3,4)$ $\qquad \|(3,4)\|_2=5$

向量1范数、2范数、∞范数的对比



性质

【定理】对赋范空间的任意向量 x,y ，有

（1）$||-x||=||x||$

（2）$||x-y||\geq| \, ||x|| - ||y|| \, |$

【Hölder不等式】

对任意的向量$x=(x_1,x_2,\ldots,x_n)$和 $y=(y_1,y_2,\ldots,y_n)$有

$$\sum_{i=1}^{n}|x_i||y_i| \leq ||x||_p \, ||y||_q$$

其中$p>1$，$q>1$，且$\frac{1}{p}+\frac{1}{q}=1$

$p=q=2$
Cauchy-Schwarz
不等式

定理(范数的连续性):

C^n中的任何范数$\|x\|$均为x的连续函数。

定义(范数等价) 在C^n中有两个范数$\|\cdot\|$和$\|\cdot\|'$，

若存在实数$M, m>0$，使得对任意的 n 维向量 x，都有

$$m\|x\| \le \|x\|' \le M\|x\|$$

则称这两个范数**等价**.

- C^n中范数的重要性质: 范数等价定理

范数等价定理: C^n中任意两个范数等价.

例 1-范数, 2-范数和 ∞-范数是两两等价的.

$$\| x \|_\infty \le \| x \|_2 \le \sqrt{n} \| x \|_\infty$$

$$\frac{1}{n} \| x \|_1 \le \| x \|_2 \le \| x \|_1$$

$$\| x \|_\infty \le \| x \|_1 \le n \| x \|_\infty$$

提 纲

- 1. 向量范数
- 2. 方阵范数
- 3. 一般矩阵范数
- 4. 矩阵的谱半径
- 5. 矩阵的条件数

● 矩阵范数是用于定义矩阵“大小”的量，类似向量范数，定义 n 阶方阵A的范数.

定义(矩阵范数) 设A为n 阶方阵，按照一定规则有一实数与之对应，记为$\|A\|$，若$\|A\|$满足：

1) $\|A\|\geq 0$，$\|A\|=0$当且仅当$A=0$时 **正定性**

2) 对任意复数a，$\|aA\|=|a|\ \|A\|$ **齐次性**

3) 对任意两个n阶方阵A, B，都有 $\|A+B\|\leq\|A\|+\|B\|$ **三角不等式**

4) $\|AB\|\leq\|A\|\ \|B\|$ **相容性**

则称$\|A\|$为矩阵A的范数.

- 设矩阵 $A=[a_{ij}]_{n\times n}\in M_n$

矩阵的$\mathbf{m_1}$范数

$$\|A\|_{\mathrm{m1}}=\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|$$

矩阵的$\mathbf{m_2}$范数(Frobenius范数，F范数)

$$\|A\|_{\mathrm{m2}}=\|A\|_{\mathrm{F}}=\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2}$$

矩阵的$\mathbf{m_\infty}$范数

$$\|A\|_{\mathrm{m}\infty}=n\max_{i,j}|a_{ij}|$$

【Frobenius范数性质】:

设A为n阶矩阵, 则

$$\|A\|_F=\sqrt{\sigma_1^2+\sigma_2^2+\ldots+\sigma_r^2}$$

其中$\sigma_1, \sigma_2,\ldots, \sigma_r$为$A$的所有非零奇异值。

$$\|A\|_F^2=\mathrm{tr}(AA^H)$$

相似矩阵有相同的迹

$$\mathrm{tr}(AA^H)=\sigma_1^2+\sigma_2^2+\ldots+\sigma_r^2$$

$$\|A\|_F=\sqrt{\sigma_1^2+\sigma_2^2+\ldots+\sigma_r^2}$$

【Frobenius范数性质】：　F范数-酉不变性

对于任意的n阶酉矩阵U, V

$$\|A\|_F = \|UA\|_F = \|AV\|_F = \|UAV\|_F$$

$$UA(UA)^H = UAA^HU^H$$

$$AV(AV)^H = AVV^HA^H = AA^H$$

$$UAV(UAV)^H = UAVV^HA^HU^H = UAA^HU^H$$

相似矩阵具有相同的迹

【意义】比较不同矩阵的差异大小

低秩逼近 在秩为$k(<r=\mathrm{rank}(A))$的所有$m\times n$矩阵中与A的距离最近的矩阵B?

$$\|A-B\|_F^2=\|U\Sigma V^T-B\|_F^2=\|\Sigma-U^TBV\|_F^2$$

上式取到最小值时，U^TBV必须为对角矩阵,又$\mathrm{rank}(B)=k$,

当 $U^TBV=\begin{bmatrix}\Sigma_k & 0\\ 0 & 0\end{bmatrix}_{m\times n}$， $\Sigma_k=\mathrm{diag}(\sigma_1,\sigma_2,\cdots,\sigma_k)$

即当 $\boxed{B=\sigma_1u_1v_1^T+\sigma_2u_2v_2^T+\cdots+\sigma_ku_kv_k^T}$ 时取到最小值

$$\min_{\mathrm{rank}(B)=k}\|A-B\|_F^2=\sigma_{k+1}^2+\sigma_{k+2}^2+\cdots+\sigma_r^2$$

图像压缩的理论依据之一.

【意义】 · 比较不同矩阵的差异大小

给定矩阵*A*，求矩阵*X*，使得

$$\min_X \|X - A\|_F^2,\ XX^{\mathrm{T}} = I$$

$$\|X-A\|_F = \|X-QP\|_F = \|Q^TX-P\|_F = \|Q^TX-Q_1^T\Sigma Q_1\|_F$$

$$= \|Q_1Q^TXQ_1^T-\Sigma\|_F$$

上式取到最小值时，$Q_1Q^TXQ_1^T$必须为对角半正定矩阵，

正交矩阵的乘积是正交矩阵，所以$Q_1Q^TXQ_1^T$还是正交矩阵。

对角的正交半正定矩阵为单位矩阵*E*。

所以$X=Q$

性质

- 单位矩阵的范数值

$$\|I_n\|_{\mathrm{m1}} = n$$

$$\|I_n\|_{\mathrm{F}} = \sqrt{n}$$

$$\|I_n\|_{\mathrm{m}\infty} = n$$

提纲

- 1. 向量范数
- 2. 方阵范数
- **3. 一般矩阵范数**
- 4. 矩阵的谱半径
- 5. 矩阵的条件数

方阵范数与向量范数相容性

- 向量范数

$$\|\boldsymbol{a}\|_1 \quad \|\boldsymbol{a}\|_2 \quad \|\boldsymbol{a}\|_\infty \quad \cdots$$

- 方阵范数

$$\|A\|_{\mathrm{m}1} \quad \|A\|_{\mathrm{m}2} \quad \|A\|_{\mathrm{m}\infty} \quad \cdots$$

- 【相容性】

设$\|\cdot\|_{\mathrm{m}}$是$\boldsymbol{M_n}$上的矩阵范数，$\|\cdot\|_{\mathrm{v}}$是$\boldsymbol{F^n}$上的向量范数，如果对任意$A \in M_n$和$\boldsymbol{a} \in F^n$都有

$$\|A\boldsymbol{a}\|_{\mathrm{v}} \le \|A\|_{\mathrm{m}} \|\boldsymbol{a}\|_{\mathrm{v}}$$

则称矩阵范数$\|\cdot\|_{\mathrm{m}}$和向量范数$\|\cdot\|_{\mathrm{v}}$是相容的。

方阵范数与向量范数相容性

【例】：M_n上的矩阵m_1范数$\|\cdot\|_{m1}$和F范数$\|\cdot\|_F$分别与向量1范数和向量2范数相容。

【例】 M_n上的矩阵 m_∞ 范数 $\|\cdot\|_{m\infty}$ 分别与向量1范数、向量2范数和向量∞范数相容。

方阵范数与向量范数相容性

【定理】设$\|\cdot\|_v$是F^n上的一种向量范数，则在$M_n(F)$上必存在与它相容的矩阵范数。

从属范数(导出范数)

【定义】已知向量范数 $\|\cdot\|_v$，对任意矩阵 $A\in M_{m\times n}$，规定

$$\|A\|=\max_{x\neq 0}\frac{\|A\boldsymbol{x}\|_v}{\|\boldsymbol{x}\|_v}=\max_{\|\boldsymbol{x}\|_v=1}\|A\boldsymbol{x}\|_v$$

则上述 $\|\cdot\|$ 称为由向量范数 $\|\cdot\|_v$ 导出的矩阵范数，或者从属于向量范数 $\|\cdot\|_v$ 的矩阵范数，简称导出范数或从属范数。

从属范数

【定理】从属范数 $\|\cdot\|$ 具有如下性质：

（1）$\|I_n\|=1$

（2）对任意方阵 $A\in M_n$ 以及任意向量 $\boldsymbol{x}\in F^n$ 都有

$$\|A\boldsymbol{x}\|_{\mathrm{v}}\le\|A\|\,\|\boldsymbol{x}\|_{\mathrm{v}}$$

（3）$\|\cdot\|$ 是 M_n 上的一个矩阵范数。

常用矩阵范数

- 对于任意$A\in M_{m\times n}$，常用的以下7种矩阵范数

（1）$\mathbf{m_1}$范数：$\|A\|_{\mathrm{m1}}=\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|$

（2）$\mathbf{m_2}$范数(F范数)：

$$\|A\|_{\mathrm{F}}=\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2}=\sqrt{\mathrm{tr}(A^*A)}$$

（3）$\mathbf{M_\infty}$范数或最大范数：

$$\|A\|_{\mathrm{M}}=\max\{m,n\}\max_{i,j}|a_{ij}|$$

常用矩阵范数

- 对于任意$A \in M_{m\times n}$，常用的以下7种矩阵范数

（4）G范数或几何平均范数：$\|A\|_G = \sqrt{mn}\max_{i,j}|a_{ij}|$

（5）1范数或列范数：$\|A\|_1 = \max_j \sum_{i=1}^{m}|a_{ij}|$

（6）2范数或谱范数：$\|\boldsymbol{A}\|_2 = \sqrt{\boldsymbol{\rho}(\boldsymbol{A}^*\boldsymbol{A})}$

（7）∞范数或行范数：

$$\|A\|_\infty = \max_i \sum_{j=1}^{n}|a_{ij}|$$

提 纲

- 1. 向量范数
- 2. 方阵范数
- 3. 一般矩阵范数
- **4. 矩阵的谱半径**
- 5. 矩阵的条件数

谱半径

【定义】设方阵$A \in M_n$的n个特征值为λ_1、λ_2、…、λ_n，称

$$\rho(A) = \max_j |\lambda_j|$$

为A的谱半径。

性质

【定理】设方阵 $A \in M_n$，则

（1）$\rho(A^k) = \rho(A)^k$

（2）$\rho(A^* A) = \rho(AA^*) = \|A\|_2^2$

（3）当**A**是正规矩阵时，$\rho(A) = \|A\|_2$

性质

- 【定理】设方阵$A \in M_n$，则对$\boldsymbol{M_n}$上的任一相容的矩阵范数 $\|\cdot\|$，都有

$$\rho(A) \le \| A \|$$

性质

- 意义：谱半径是矩阵的任意一种范数的下界

例：试估计如下矩阵的谱半径

$$A=\begin{bmatrix} 0 & 0.2 & 0.1 \\ -0.2 & 0 & 0.2 \\ -0.1 & -0.2 & 0 \end{bmatrix}$$

$$\|A\|_1=\|A\|_\infty=0.4 \qquad \|A\|_{m1}=1 \qquad \|A\|_{m\infty}=0.6$$

$$\|A\|_F=\sqrt{0.18}$$

$$\Rightarrow \quad \rho(A)\le 0.4$$

提 纲

- 1. 向量范数
- 2. 方阵范数
- 3. 一般矩阵范数
- 4. 矩阵的谱半径
- **5. 矩阵的条件数**

条件数

条件数同时描述了对向量的拉伸能力和压缩能力，即令向量发生形变的能力。
条件数越大，向量在变换后越可能变化得越多。

【定义】

设矩阵$A\in M_n$ 是可逆矩阵，$\|.\|$是矩阵范数，称

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

为矩阵A的条件数。

$$\|A\| = \max_{x} \frac{\| A\boldsymbol{x} \|}{\| \boldsymbol{x} \|}$$

A对向量的拉伸能力

$$\|A^{-1}\| = \max_{y} \frac{\| A^{-1}\boldsymbol{y} \|}{\| \boldsymbol{y} \|} = 1 / \min_{x} \frac{\| A\boldsymbol{x} \|}{\| \boldsymbol{x} \|}$$

A对向量的压缩能力

条件数

- 常用条件数
  - **例1：∞-条件数**

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

  - **例2：1-条件数**

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

  - **例3：2-条件数**

$$\operatorname{cond}_2(A)=\|A\|_2\,\|A^{-1}\|_2=\sqrt{\frac{\lambda_1}{\lambda_n}}$$

λ_1是A^*A的最大特征值， λ_n是A^*A的最小特征值

性质

- 矩阵**A**的条件数满足以下性质：

（**1**）$\mathrm{cond}(A) \geq 1$

（**2**）$\mathrm{cond}(A) = \mathrm{cond}(A^{-1})$

（**3**）$\mathrm{cond}(kA) = \mathrm{cond}(A) \quad k \neq 0$

用途

- 线性方程的稳定性

理想线性方程组$Ax=b$

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

噪声带来的扰动

$$A\Delta x=\Delta b \Rightarrow \|A\Delta x\|=\|\Delta b\| \qquad \|A\|\,\|\Delta x\|\geq\|\Delta b\|$$

$$A^{-1}b=x \qquad \|A^{-1}\|\,\|b\|\geq\|x\|$$

$$\|A\|\,\|A^{-1}\|\,\|\Delta x\|\,\|b\|\geq\|\Delta b\|\,\|x\|$$

$$\frac{\|\Delta x\|}{\|x\|}\geq\frac{\|\Delta b\|}{\|b\|}\frac{1}{cond(A)}$$

下界

用途

- 线性方程的稳定性

理想线性方程组$Ax=b$

噪声带来的扰动

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

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

$Ax=b \qquad ||b||\le||A||\ ||x||$

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

$$\frac{||\Delta b||}{||b||}\frac{1}{cond(A)}\le\frac{||\Delta x||}{||x||}\le\frac{||\Delta b||}{||b||}cond(A)$$

上界

用途

【例】：方程组I

$$\begin{cases} x_1 + x_2 = 2, \\ x_1 + 1.00001x_2 = 2, \end{cases}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

方程组II

$$\begin{cases} x_1 + x_2 = 2, \\ x_1 + 1.00001x_2 = 2.00001, \end{cases}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

- 右端项有0.00001的差别, 最大相对误差为 0.5×10^{-5}, 但解分量的相对误差为50%.

- 平面上两条接近于平行的直线的交点, 当其中一条直线稍有变化时，新的交点可与原交点相差甚远.

计算$A=\begin{bmatrix}1 & 1\\ 1 & 1.00001\end{bmatrix}$的条件数cond($A$)

$\|A\|_1=2.00001$

$A^{-1}=\begin{bmatrix}100001 & -100000\\ -100000 & 100000\end{bmatrix}$

$\|A^{-1}\|_1=200001$

cond(A)=400004.00001

用途

- 线性方程的稳定性

如果线性方程组系数矩阵A的条件数大就称方程是病态的或坏条件的；

否则，接近于1，则称为良态的或好条件的。