

矩阵理论

王磊

自动化科学与电气工程学院

第四章 矩阵分析

- □ 向量范数
- □ 矩阵范数
- □ 相容范数
- □ 矩阵扰动分析
- □ 特征值估计

- □ 矩阵级数
- □ 矩阵函数
- □ 函数矩阵
- □ 应用: 主元分析法

第四章 矩阵分析

向量范数

定义4.1.1(向量范数)设V是数域F上的线性空间,若对任意向量 $x \in V, ||x||$ 是以x为自变量的实值函数,且满足以下三条性质:

- (1) 正定性(或非负性): $||x|| \ge 0$, 当且仅当 $x = \theta$ 时有||x|| = 0;
 - (2) 齐次性: $\forall k \in F, \in V, ||kx|| = |k|||x||;$
 - (3) 三角不等式: $\forall x, y \in F$,

$$||x + y|| \le ||x|| + ||y||$$
;

则称 $\|x\|$ 是向量x的范数, V是数域F上的赋范线性_<mark>空间, 记为(V, $\|\cdot\|$).</mark>



注1: 定义4.1.1中的三条性质称为范数的三条公理,即只要满足范数公理的实值函数均可定义为向量的范数.

注2: 三角不等式的等价形式:

$$\forall x, y \in V, |||x|| + ||y||| \le ||x - y||$$

思考: 内积空间是赋范线性空间吗?



思考: 内积空间是赋范线性空间吗?

分析: 判断内积空间是否为赋范线性空间只需判断 内积空间是否定义了符合范数三条公理的实值函 数. 显然. 内积空间定义的向量长度符合这一要求. 因此, 内积空间是赋范线性空间, 实际上, 内积空间 的长度定义不仅满足范数的三条公理, 还满足平行 四边形法则:而赋范线性空间并不一定满足这一法 则.



例4.1.1
$$\forall x = [x_1, x_2, \cdots, x_n]^T \in \mathbb{C}^n$$
, 定义

$$||x||_{1} = \sum_{i=1}^{n} |x_{i}|$$

$$||x||_{\infty} = \max_{1 \le i \le n} |x_{i}|$$

$$||x||_{2} = (\sum_{i=1}^{n} |x_{i}|^{2})^{\frac{1}{2}}$$

则 $\|x\|_1$, $\|x\|_\infty$, $\|x\|_2$ 均是向量x的范数, 分别称为1-范数、 ∞ –范数和2-范数(或欧几里得范数).

同一线性空间可定义不同的范数.



例4.1.2 设向量组 $\alpha_1, \alpha_2, \dots, \alpha_n$ 是n维线性空间V的一组基,V中任意向量 α 在这组基下的坐标为 $x = [x_1, x_2, \dots, x_n]^T$. 由此,可定义 α 的范数为

$$\|\alpha\| \triangleq \sqrt{\sum_{i=1}^{n} |x_i|^2}$$

例4.1.2 设向量组 $\alpha_1, \alpha_2, \dots, \alpha_n$ 是n维线性空间V的一组基,V中任意向量 α 在这组基下的坐标为 $x = [x_1, x_2, \dots, x_n]^T$. 由此,可定义 α 的范数为

$$\|\alpha\| \triangleq \sqrt{\sum_{i=1}^{n} |x_i|^2}$$

思考: 例4.1.2定义的范数是2-范数吗?



例4.1.3 设 $1 \le p \le \infty$, 对任意向量

$$\mathbf{x} = [x_1, x_2, \cdots, x_n]^T \in \mathbb{C}^n$$

定义

$$||x||_p = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$$

则 $||x||_p$ 是向量x的范数,称为p-范数.

注3:当p = 1,2时, $||x||_p$ 分别是例4.1.1中的1-范数和2-范数.



注3: 当 $p = \infty$ 时, $\lim_{p \to \infty} ||x||_p = ||x||_\infty$. 当 $0 时, 此时<math>||x||_p$ 不满足三角不等式.

注4: 范数证明常涉及两个重要的不等式, 分别为

(1) Holder不等式: 设p, q > 1且对任意 $x = [x_1, x_2, \dots, x_n]^T \in \mathbb{C}^n$ 和 $\mathbf{y} = [y_1, y_2, \dots, y_n]^T \in \mathbb{C}^n$,

$$\sum_{i=1}^{n} |x_i y_i| \le \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} |y_i|^q\right)^{\frac{1}{q}}$$



(2) Minkowski不等式:

$$\forall \mathbf{x} = [x_1, x_2, \cdots, x_n]^T \in \mathbb{C}^n, \ \mathbf{y} = [y_1, y_2, \cdots, y_n]^T \in \mathbb{C}^n$$
以及 $p \ge 1$

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{\frac{1}{p}} \le \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{\frac{1}{p}}$$

例4.1.4 设 $A \in \mathbb{C}^{n \times n}$ 是正定Hermite矩阵, 对任意向量 $x \in \mathbb{C}^n$, 定义 $\|x\|_A = \sqrt{x^H A x}$, 则 $\|x\|_A$ 是向量范数, 常称为加权范数或椭圆范数.

注5: 例4.1.4提供了一种借助已有向量范数构造新的向量范数方法. 实际上, 我们可利用定义式(4.1.6)构造更一般的向量范数(见习题2).

【应用】:系统稳定性分析. 我们常采用二次型 Lyapunov函数

$$V(\mathbf{x}) = \mathbf{x}^H P \mathbf{x} = \|\mathbf{x}\|_P^2$$

研究线性系统乃至非线性系统的稳定性, 其中, P 是正定实对称矩阵.



【应用】:模式分类.模式分类常通过已知类型属 性的观测样本判断未知样本的类型属性, 其基本思 想是根据未知样本的模式向量x与已知样本的模 式向量 s_1, \dots, s_M 相似度来判断未知样本的类型属 性. 最简单的方法就是用向量间的距离来度量模式 向量的相似度, 比如经典的欧几里得距离, 就是采 用向量的2-范数:

$$D(x, s_i) = ||x - s_i||_2 = \sqrt{(x - s_i)^H (x - s_i)}$$



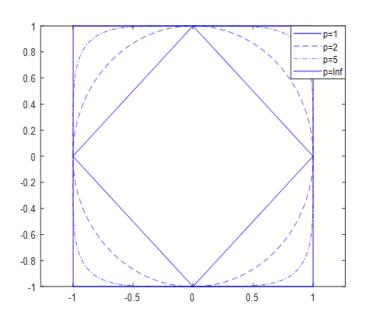
思考: 范数的本质是什么?

思考: 范数的本质是什么?

定理4.1.1 线性空间V中任一范数 $\|x\|$ 都是其坐标的连续函数.



例4.1.5 画出 \mathbb{R}^n 空间不同范数下的"单位圆".





M4.1.5 画出 \mathbb{R}^n 空间不同范数下的"单位圆".

注6: 赋范线性空间中的单位圆具有重要意义. 在几何上, 它们相当于实数轴上的单位闭区间或其端点, 或三维空间中的单位球或球面. 因此, 它们都是有界闭集. 连续的函数在有界闭集上一定有最大值和最小值, 这是高等数学课程中的一个重要结论.

研究赋范线性空间上的连续函数或变换的一个 重要技巧就是设法将函数的定义域限制或转移到 单位圆上.



则称 $\|x\|_{\alpha}$ 与 $\|x\|_{\beta}$ 是等价的.

定义4.1.2(范数等价)设V是数域F上的有限维线性空间, $\|x\|_{\alpha}$ 和 $\|x\|_{\beta}$ 是V中任意两个向量范数. 若存在正数 k_1 和 k_2 使得 $\forall x \in V$,都有 $k_1\|x\|_{\beta} \leq \|x\|_{\alpha} \leq k_2\|x\|_{\beta}$

定义4.1.2(范数等价)设V是数域F上的有限维线性空间, $\|x\|_{\alpha}$ 和 $\|x\|_{\beta}$ 是V中任意两个向量范数. 若存在正数 k_1 和 k_2 使得 $\forall x \in V$,都有

 $k_1 \|\boldsymbol{x}\|_{\beta} \le \|\boldsymbol{x}\|_{\alpha} \le k_2 \|\boldsymbol{x}\|_{\beta}$

则称 $\|x\|_{\alpha}$ 与 $\|x\|_{\beta}$ 是等价的.

定理4.1.2 有限维线性空间中的任意向量范数都是等价的.



<mark>命题4.1.1(范数等价的性质</mark>)设V是数域F上的有限维线性空间,向量范数 $\|x\|_{\alpha}$ 与 $\|x\|_{\beta}$ 等价,则

- (1) $\mathbf{p} = \mathbf{p} \cdot \mathbf{p$
- (2) 对称性: $\frac{1}{k_1} \|x\|_{\alpha} \le \|x\|_{\beta} \le \frac{1}{k_2} \|x\|_{\alpha}$;
- (3) 传递性: 若 $\|x\|_{\beta}$ 与 $\|x\|_{\gamma}$ 等价,则向量范数 $\|x\|_{\alpha}$ 与 $\|x\|_{\gamma}$ 等价.

赋范线性空间中范数等价性结论是后续向量序列和矩阵序 列的收敛性证明的重要基础理论.



第四章 矩阵分析

矩阵范数

定义4.2.1(矩阵的向量范数)对任意矩阵 $A \in \mathbb{C}^{m \times n}$,定义 $\|A\|$ 是对应以A为自变量的实值函数,且满足以下三条性质:

- (1) 正定性(或非负性): $||A|| \ge 0$, 当且仅当A = 0时有||A|| = 0;
 - (2) 齐次性: $\forall k \in \mathbb{C}, ||kA|| = |k|||A||$;
 - (3) 三角不等式: $\forall A, B \in \mathbb{C}^{m \times n}$, $\|A + B\| \le \|A\| + \|B\|$;

则称||A||是矩阵A的向量范数.



例4.2.1 设 $A = (a_{ij}) \in \mathbb{C}^{m \times n}$,则

$$||A||_{v1} = \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|$$

$$||A||_{v\infty} = \max_{\forall i,j} |a_{ij}|$$

$$||A||_{v2} = (\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2)^{\frac{1}{2}}$$

$$||A||_{vp} = (\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^p)^{\frac{1}{p}}, p \ge 1$$

均为矩阵A的向量范数.

$$||A||_{vp} = (\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^p)^{\frac{1}{p}}, p \ge 1$$

均为矩阵A的向量范数.

类似于上一节,对于矩阵的向量范数,我们同样有以下结论.

定理4.2.1 矩阵 $A = (a_{ij}) \in \mathbb{C}^{m \times n}$ 的任一向量范数均是A元素的连续函数;



定理4.2.2 线性空间 $\mathbb{C}^{m \times n}$ 的任意两个向量范数是等价的,即对 $\|A\|_{v\alpha}$ 和 $\|A\|_{v\beta}$ 存在正整数 k_1 和 k_2 使得对任意矩阵 $A \in \mathbb{C}^{m \times n}$,有 $k_1 \|A\|_{v\beta} \leq \|A\|_{v\alpha} \leq k_2 \|A\|_{v\beta}$

定理4.2.2 线性空间 $\mathbb{C}^{m\times n}$ 的任意两个向量范数是等价的,即对 $\|A\|_{v\alpha}$ 和 $\|A\|_{v\beta}$ 存在正整数 k_1 和 k_2 使得对任意矩阵 $A\in\mathbb{C}^{m\times n}$,有

 $k_1 ||A||_{v\beta} \le ||A||_{v\alpha} \le k_2 ||A||_{v\beta}$

尽管矩阵可视为拉直的向量,但矩阵和向量还是有所不同,典型的是矩阵有乘法运算.因此,在考虑范数时须兼顾矩阵的乘法运算.



定义4.2.2(矩阵范数)对任意矩阵 $A \in \mathbb{C}^{m \times n}$,定义 ||A|| 是对应以A为自变量的实值函数,且满足以下四条性质:

- (1) 正定性: $||A|| \ge 0$, 当且仅当A = 0时有||A|| = 0;
- (2) 齐次性: $\forall k \in \mathbb{C}, ||kA|| = |k|||A||$;
- (3) 三角不等式: $||A + B|| \le ||A|| + ||B||$;
- (4) 矩阵乘法的相容性: ||AB|| ≤ ||A||||B|| (须矩阵乘积有意义).

则称||A||是矩阵A的矩阵范数.



注1: 若反向性质(4)中矩阵乘法相容性的不等式,即 $||AB|| \ge ||A||||B||$,此时幂零矩阵(对于n阶复方阵A,存在正整数k使得 $A^k = 0$,则称A为幂零矩阵)的矩阵范数将是0,这与矩阵范数的正定性要求矛盾.

矩阵乘法相容性的不等式实际上保证了矩阵幂级数的收敛性(设 $||A|| \le 1$,则根据矩阵乘法相容性得当 $k \to \infty$, $||A^k|| \to 0$).

思考:设 $A = (a_{ij}) \in \mathbb{C}^{m \times n}, ||A||_{v \infty} = \max_{\forall i,j} |a_{ij}|$ 是矩

阵范数吗?

思考: 设
$$A = (a_{ij}) \in \mathbb{C}^{m \times n}, ||A||_{v\infty} = \max_{\forall i,j} |a_{ij}|$$
是矩

阵范数吗?

考察
$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$
, $B = \begin{bmatrix} -1 & 0 \\ -1 & 1 \end{bmatrix}$,

则

$$AB = \begin{bmatrix} -1 & 0 \\ -2 & 1 \end{bmatrix}.$$

例4.2.2 设 $A = (a_{ij}) \in \mathbb{C}^{m \times n}$, 证明:

$$||A||_{v2} = (\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2)^{\frac{1}{2}} = (\operatorname{tr}(A^H A))^{\frac{1}{2}}$$

是A的矩阵范数.

例4.2.2 设 $A = (a_{ij}) \in \mathbb{C}^{m \times n}$, 证明:

$$||A||_{v2} = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2\right)^{\frac{1}{2}} = \left(\operatorname{tr}(A^H A)\right)^{\frac{1}{2}}$$

是A的矩阵范数.

注2: $||A||_{v_2}$ 范数称为Frobenious范数, 简称为F — 范数, 并常记为 $||A||_F$.



定理4.2.3 设 $A = (a_{ij}) \in \mathbb{C}^{m \times n}$ 和 $x \in \mathbb{C}^n$,则

- (1) $||UA||_F = ||AV||_F = ||UAV||_F = ||A||_F$, 其中 U和V是酉矩阵;
- (2) $\|A\|_F^2 = \sum_{i=1}^n \|\boldsymbol{\beta}_i\|_2^2$, 其中矩阵A按列分块, 记为 $A = [\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_n]$; 或 $\|A\|_F^2 = \sum_{i=1}^n \|\boldsymbol{\alpha}_i\|_2^2$, 其中矩阵A按行分块;
 - (3) $||Ax|| \le ||A||_F ||x||_2$
- 注3: 定理4.2.2性质(1)称为F -范数的酉不变性.



第四章 矩阵分析——矩阵范数

思考:如何根据已有的矩阵范数构造新的矩阵范数?

第四章 矩阵分析——矩阵范数

思考:如何根据已有的矩阵范数构造新的矩阵范数?

例4.2.3 设 $A \in \mathbb{C}^{n \times n}$, $\|\cdot\|$ 是某一给定矩阵范数. 定义

$$||A||_m = ||P^{-1}AP||$$

则 $\|\cdot\|_m$ 是矩阵范数,其中P是n阶可逆矩阵.



第四章 矩阵分析

相容范数



定义4.3.1(向量范数与矩阵范数相容)若对 $A \in \mathbb{C}^{m \times n}$ 和 $x \in \mathbb{C}^n$,向量范数 $\|x\|_v$ 与矩阵范数 $\|A\|_m$ 满足

 $||Ax||_v \le ||A||_m ||x||_v$

则称向量范数 $\|x\|_v$ 与矩阵范数 $\|A\|_m$ 相容.



定义4.3.1(向量范数与矩阵范数相容)若对 $A \in \mathbb{C}^{m \times n}$ 和 $x \in \mathbb{C}^n$,向量范数 $\|x\|_v$ 与矩阵范数 $\|A\|_m$ 满足

 $||Ax||_v \le ||A||_m ||x||_v$

则称向量范数 $\|x\|_v$ 与矩阵范数 $\|A\|_m$ 相容.

思考: 给定矩阵范数, 是否存在与之相容的向量范数? 反之, 给定向量范数, 是否存在与之相容的矩阵范数?



例4.3.1 由定理4.2.3可知, 对任意矩阵 $A \in \mathbb{C}^{m \times n}$ 和 $x \in \mathbb{C}^n$, $||Ax||_2 \le ||A||_F ||x||_2$. 由此, 向量2-范数和矩阵F -范数是相容的.

定理4.3.1 设 $\|A\|_m$ 是 $\mathbb{C}^{n\times n}$ 的一个矩阵范数,则必存在 \mathbb{C}^n 上与之相容的向量范数.

定理4.3.2 设 $\|x\|_v$ 是 \mathbb{C}^n 上的一个向量范数, 对任意 $A \in \mathbb{C}^{m \times n}$, 定义

$$||A|| = \max_{\|x\|_v=1} ||Ax||_v$$

则||A||是一个与 $||x||_v$ 相容的矩阵范数, 称其是从属于向量范数 $||\cdot||_v$ 的算子范数或由向量范数 $||\cdot||_v$ 诱导的矩阵范数(或简称诱导范数).



注1: 我们常用到算子范数有如下等价定义式:

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

常见的向量范数有 $\|x\|_1$, $\|x\|_\infty$, $\|x\|_2$, 则从属于它们的算子范数分别记为 $\|A\|_1$, $\|A\|_\infty$, $\|A\|_2$, 并分别称为列和范数、行和范数和谱范数.

定理4.2.3 设 $A = (a_{ij}) \in \mathbb{C}^{m \times n}$ 和 $x \in \mathbb{C}^n$, $\sigma_{\max}(A)$ 是矩阵A的最大奇异值, 则从属于向量范数 $\|x\|_1$, $\|x\|_{\infty}$, $\|x\|_2$ 的算子范数分别为

$$||A||_1 = \max_{1 \le j \le n} \sum_{i=1}^m |a_{ij}|, (列和范数)$$

$$||A||_{\infty} = \max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}|, (行和范数)$$



定理4.2.3

$$||A||_2 = \sqrt{\lambda_{\max}(A^H A)} = \sigma_{\max}(A), ($$
im 范数)

注2: 由定理4.2.3知, $||A||_F = ||x||_2$ 是相容的, 而 $||A||_2$ 作为从属于 $||x||_2$ 的算子范数, 自然也是相容的, 所以相容于同一向量范数的矩阵范数有多个. 事实上, 我们总有 $||A||_2 \le ||A||_F$.

注3: 若存在常数M使得对任一向量 $x \in \mathbb{C}^n$,有 $\|Ax\|_v \leq M\|x\|_v$,则 $\|A\|_v \leq M$,即从属于范数 $\|x\|_v$ 的算子范数 $\|A\|_v$ 是使不等式 $\|Ax\|_v \leq M\|x\|_v$ 成立的最小常数.



例4.3.2 设
$$A = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 2 & 3 \\ 1 & 2 & 0 \end{bmatrix}$$
,分别计算 $\|A\|_1$, $\|A\|_\infty$, $\|A\|_2$ 和 $\|A\|_F$.

例4.3.2 设
$$A = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 2 & 3 \\ 1 & 2 & 0 \end{bmatrix}$$
,分别计算 $\|A\|_1$, $\|A\|_\infty$, $\|A\|_2$ 和 $\|A\|_F$.

解:
$$||A||_1 = 5$$
, $||A||_{\infty} = 5$, $||A||_F = \sqrt{23}$, $||A||_2 = \sqrt{15}$.

注: $||A||_1$, $||A||_\infty$ 容易计算; $||A||_2$ 计算复杂,对矩阵元素变化比较敏感.



注4: 设 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ 是Hermite矩阵A的n个特征值,则由谱范数证明过程知

$$\lambda_n \le R(\mathbf{x}) \le \lambda_1$$

其中R(x)的定义为

$$R(\mathbf{x}) = \frac{\mathbf{x}^H A \mathbf{x}}{\mathbf{x}^H \mathbf{x}}, \forall \mathbf{x} \neq 0$$

称R(x)为Hermite矩阵A的Rayleigh商.



推论4.3.1 设 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ 是Hermite矩阵 A的n个特征值, 则

$$\lambda_1 = \max_{orall x
eq 0} R(x)$$
 , $\lambda_n = \min_{orall x
eq 0} R(x)$

基于Rayleigh商,有如下著名的极大极小定理或极小极大定理。

定理4.3.4 设 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ 是Hermite矩阵 A的n个特征值, V_i 是 \mathbb{C}^n 中i维子空间,则 $\forall i=1,\cdots,n$

$$\lambda_{i} = \max_{\forall V_{i}} \min_{\mathbf{x} \in V_{i}, \mathbf{x} \neq 0} R(\mathbf{x})$$

$$\lambda_{i} = \min_{\forall V_{n-i+1}} \max_{\mathbf{x} \in V_{n-i+1}, \atop \mathbf{x} \neq 0} R(\mathbf{x})$$



【应用】: 利用极大极小定理可研究Hermite矩阵特征值的扰动问题(讨论Hermite矩阵元素发生变化时相应矩阵特征值的变化范围).

定理4.3.5 设A和E均为n阶Hermite复方阵,则对 $i=1,\cdots,n$,有 $\lambda_i(A)+\lambda_n(E)\leq \lambda_i(A+E)\leq \lambda_i(A)+\lambda_1(E)$ 式中, $\lambda_i(A)$ 是矩阵A的第i大特征值.



定理4.3.6 设 $A \in \mathbb{C}^{n \times n}$, ||A||是某一矩阵范数. 若 ||A|| < 1, 证明I - A非奇异, 且

$$||(I - A)^{-1}|| \le \frac{||I||}{1 - ||A||}$$