第一讲　引论与预备知识

— 线性代数基础

1 线性空间基本概念

- 数域, 如: $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$
- 线性空间, 如: $\mathbb{R}^n$, $\mathbb{C}^n$, $\mathbb{R}^{m\times n}$, $\mathbb{C}^{m\times n}$, $\mathbb{H}_n$, $C[a,b]$, $C^p[a,b]$
- 线性相关与线性无关, 线性表示, 秩, 基, 维数
- 线性子空间 (简称子空间), 零空间 (核) $\mathrm{Ker}(A)$

张成的线性空间

例 设 $\mathbb{S}$ 是数域 $\mathbb{F}$ 上的一个线性空间, $x_1, x_2, \ldots, x_k \in \mathbb{S}$, 记

$$\operatorname{span}\{x_1, x_2, \ldots, x_k\} \triangleq \left\{ \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_k x_k \ : \ \alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{F} \right\},$$

即由 $x_1, x_2, \ldots, x_k$ 的所有线性组合构成的集合, 则 $\operatorname{span}\{x_1, x_2, \ldots, x_k\}$ 是 $\mathbb{S}$ 的一个线性子空间, 称为由 $x_1, x_2, \ldots, x_k$ **张成的线性空间**.

✍ $\operatorname{span}(A) \to$ 由 A 的列向量所张成的线性空间, 即 A 的像空间、值域、列空间

2 范数与赋范线性空间

定义 (范数与赋范线性空间) 设 $\mathbb{S}$ 为数域 $\mathbb{F}$ ($\mathbb{F}$ 可以是 $\mathbb{R}$ 或 $\mathbb{C}$) 上的线性空间, 若对任意 $x \in \mathbb{S}$, 存在唯一实数与之对应, 记为 $\|x\|$, 它满足条件:

(1) $\|x\| \geq 0$, 等号当且仅当 $x = 0$ 时成立; (正定性)

(2) $\|\alpha x\| = |\alpha| \cdot \|x\|, \forall\, \alpha \in \mathbb{F}$; (正齐次性)

(3) $\|x + y\| \leq \|x\| + \|y\|, \forall\, x, y \in S$; (三角不等式)

则称 $\|\cdot\|$ 为线性空间 $\mathbb{S}$ 上的**范数**, 定义了范数的线性空间称为**赋范线性空间**.

✍ 范数是从 $\mathbb{S}$ 到 $\mathbb{R}_+ \cup \{0\}$ 的**一元函数**.

例 $C[a,b]$ 上的常用范数: 设 $f(x)\in C[a,b]$

- 1-范数: $\|f\|_1 = \displaystyle\int_a^b |f(x)|\,\mathrm{d}x$;
- 2-范数: $\|f\|_2 = \left(\displaystyle\int_a^b |f(x)|^2\,\mathrm{d}x\right)^{\frac{1}{2}}$;
- p-范数: $\|f\|_p = \left(\displaystyle\int_a^b |f(x)|^p\,\mathrm{d}x\right)^{\frac{1}{p}}$; ($p$ 为正整数)
- ∞-范数: $\|f\|_\infty = \max\limits_{a\le x\le b} |f(x)|$.

3 内积与内积空间

定义 (内积与内积空间) 设 $\mathbb{S}$ 是数域 $\mathbb{F}$ ($\mathbb{R}$ 或 $\mathbb{C}$) 上的一个线性空间, 定义一个从 $\mathbb{S}\times\mathbb{S}$ 到 $\mathbb{F}$ 的代数运算, 记为 "$(\cdot\,,\,\cdot)$", 即对任意 $x,y\in\mathbb{S}$, 都存在唯一的 $f\in\mathbb{F}$, 使得 $f=(x,y)$. 如果该运算满足

(1) $(x,y)=\overline{(y,x)},\quad \forall\, x,y\in\mathbb{S}$;

(2) $(\alpha x,y)=\alpha(x,y),\quad \forall\,\alpha\in\mathbb{F},\ x,y\in\mathbb{S}$;

(3) $(x+y,z)=(x,z)+(y,z),\quad \forall\, x,y,z\in\mathbb{S}$;

(4) $(x,x)\geq 0$, 等号当且仅当 $x=0$ 时成立;

则称 $(\cdot,\cdot)$ 为 $\mathbb{S}$ 上的一个 **内积**, 有时也称为 **标量积**. 定义了内积的线性空间称为 **内积空间**.

- ✍ 定义在实数域 $\mathbb{R}$ 上的内积空间称为**欧氏空间**, 定义在复数域 $\mathbb{C}$ 上的内积空间称为**酉空间**.

- ✍ $\overline{(u,v)}$ 表示 (u,v) 的共轭. 当 $\mathbb{F}=\mathbb{R}$ 时, 该条件即为 $(v,u)=(u,v)$.

- ✍ 内积可以看作是从线性空间 $\mathbb{S}$ 到数域 $\mathbb{F}$ 的**二元函数**.

内积基本性质

- $(u,\beta v)=\bar{\beta}(u,v), \forall\, \beta\in\mathbb{F}$
- $(u,v+w)=(u,v)+(u,w)$
- $(u_1+u_2+\cdots+u_n,v)=(u_1,v)+(u_2,v)+\cdots+(u_n,v)$

例 考虑线性空间 $\mathbb{R}^n$, 对任意 $x=[x_1,x_2,\ldots,x_n]^\mathsf{T}\in\mathbb{R}^n, y=[y_1,y_2,\ldots,y_n]^\mathsf{T}\in\mathbb{R}^n$, 定义

$$(x,y)\triangleq y^\mathsf{T}x=\sum_{i=1}^n x_iy_i,$$

则 (x,y) 是 $\mathbb{R}^n$ 上的内积, 因此 $\mathbb{R}^n$ 是一个欧氏空间. 这种方式定义的内积称为 **欧拉内积**.

☞ 如果 $x,y\in\mathbb{C}^n$, 则内积定义为 $(x,y)\triangleq y^*x=\sum_{i=1}^n x_i\bar{y}_i$

✍ 以上定义的内积是 $\mathbb{R}^n$ 和 $\mathbb{C}^n$ 上的标准内积, 若不特别声明, $\mathbb{R}^n$ 和 $\mathbb{C}^n$ 上的内积均是指标准内积.

例 考虑实数域 $\mathbb{R}$ 上的线性空间 $C[a,b]$, 对任意 $f(x), g(x) \in C[a,b]$, 定义

$$(f,g) \triangleq \int_a^b f(x)g(x)\,\mathrm{d}x,$$

则 (f,g) 是 $C[a,b]$ 上的内积, 此时 $C[a,b]$ 是一个欧氏空间.

正交

有了内积以后，我们就可以定义正交.

定义 设 $\mathbb{S}$ 是内积空间, $u, v \in \mathbb{S}$, 若 $(u, v) = 0$, 则称 u 与 v 是**正交**的.

✍ 正交可以看作是垂直概念的推广.

定理 设 $\mathbb{S}$ 是内积空间, $u_1, u_2, \ldots, u_n \in \mathbb{S}$, 则 $u_1, u_2, \ldots, u_n$ 线性无关的充要条件是矩阵 G 非奇异, 其中

$$G = \begin{bmatrix} (u_1, u_1) & (u_2, u_1) & \cdots & (u_n, u_1) \\ (u_1, u_2) & (u_2, u_2) & \cdots & (u_n, u_2) \\ \vdots & \vdots & & \vdots \\ (u_1, u_n) & (u_2, u_n) & \cdots & (u_n, u_n) \end{bmatrix}.$$

这个矩阵就称为 $u_1, u_2, \ldots, u_n$ 的 Gram **矩阵**.

(留作自习)

内积与范数

设 $\mathbb{S}$ 是内积空间, 对任意 $u \in \mathbb{S}$, 定义

$$\|u\| \triangleq (u,u)^{\frac{1}{2}},$$

则可以验证, $\|u\|$ 是 $\mathbb{S}$ 上的范数. 这就是**由内积导出的范数**. 即任意一个内积都可以定义一个范数. (证明留作练习)

例 $\mathbb{R}^n$ 上由标准内积导出的范数为

$$\|x\| = (x,x)^{\frac{1}{2}} = \left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}}.$$

这就是 2-范数.

带权内积

设 $\omega_1, \omega_2, \ldots, \omega_n$ 为任意给定的正实数, 对任意 $x, y \in \mathbb{R}^n$, 定义

$$\boxed{(x, y)_\omega \triangleq \sum_{i=1}^{n} \omega_i x_i y_i}.$$

可以验证, $(x, y)_\omega$ 是 $\mathbb{R}^n$ 上的内积, 我们称其为**带权内积**, 其中 ω_i 称为**权系数**.

☞ 同样, 可以在 $\mathbb{C}^n$ 中定义带权内积.

权函数

为了在 $C[a,b]$ 上定义带权内积, 我们首先定义权函数.

定义 设 $\omega(x)$ 是 $[a,b]$ 上的非负函数, 如果 $\omega(x)$ 满足

(1) 对任意非负整数 k, $\int_a^b x^k\omega(x)\,\mathrm{d}x$ 存在且为有限值;

(2) 对 $[a,b]$ 上的任意非负连续函数 $g(x)$, 如果 $\int_a^b g(x)\omega(x)\,\mathrm{d}x=0$, 则 $g(x)\equiv 0$;

则称 $\omega(x)$ 是 $[a,b]$ 上的一个**权函数**.

✍ $[a,b]$ 可以是有限区间, 也可以是无限区间, 即 a 可以是 $-\infty$, b 可以是 ∞;

✍ 权函数与定义区间相关.

例 常见的权函数有

- $\omega(x) = 1$ 是 $[-1, 1]$ 上的权函数;
- $\omega(x) = \dfrac{1}{\sqrt{1-x^2}}$ 是 $[-1, 1]$ 上的权函数;
- $\omega(x) = e^{-x}$ 是 $[0, \infty)$ 上的权函数;
- $\omega(x) = e^{-x^2}$ 是 $(-\infty, \infty)$ 上的权函数.

例 设 $\omega(x)$ 是 $[a,b]$ 上的权函数, 对任意 $f(x),g(x)\in C[a,b]$, 定义

$$(f,g)_\omega=\int_a^b \omega(x)f(x)g(x)\,\mathrm{d}x,$$

则 $(f,g)_\omega$ 是 $C[a,b]$ 上的带权内积. 对应的带权内积导出范数为

$$\|f\|_\omega=(f,f)_\omega^{\frac{1}{2}}=\left(\int_a^b \omega(x)f^2(x)\,\mathrm{d}x\right)^{\frac{1}{2}}.$$

特别地, 当 $\omega(x)\equiv 1$ 时, 内积导出范数为

$$\|f\|=(f,f)^{\frac{1}{2}}=\left(\int_a^b f^2(x)\,\mathrm{d}x\right)^{\frac{1}{2}},$$

即为 $f(x)$ 的 2-范数.

推论 设 $\varphi_1(x), \varphi_2(x), \ldots, \varphi_n(x) \in C[a, b]$, 则 $\varphi_1(x), \varphi_2(x), \ldots, \varphi_n(x)$ 线性无关的充要条件是矩阵 G 非奇异, 其中 G 是 Gram 矩阵:

$$G = \begin{bmatrix} (\varphi_1, \varphi_1) & (\varphi_2, \varphi_1) & \cdots & (\varphi_n, \varphi_1) \\ (\varphi_1, \varphi_2) & (\varphi_2, \varphi_2) & \cdots & (\varphi_n, \varphi_2) \\ \vdots & \vdots & & \vdots \\ (\varphi_1, \varphi_n) & (\varphi_2, \varphi_n) & \cdots & (\varphi_n, \varphi_n) \end{bmatrix}.$$

4 矩阵特征值与谱半径

定义 设 $A \in \mathbb{C}^{n\times n}$, 若存在 $\lambda \in \mathbb{C}$ 和非零向量 $x, y \in \mathbb{C}^n$, 使得

$$Ax = \lambda x, \qquad y^* A = \lambda y^*,$$

则 λ 是 A 的**特征值**, x 称为 A 对应于 λ 的 (右) **特征向量**, y 称为 A 对应于 λ 的**左特征向量**, 并称 (λ, x) 为 A 的一个**特征对** (eigenpair).

关于特征值的几点说明

- 只有当 A 是方阵时, 才具有特征值与特征向量;
- 实矩阵的特征值与特征向量有可能是复的;
- n 阶矩阵总是存在 n 个特征值 (其中可能有相等的);
- 特征值有代数重数和几何重数, 几何重数不超过代数重数;
- 相似变换不改变矩阵的特征值, 合同变换 (congruence transformation) 不改变矩阵的惯性指数.

思考

$A^{\intercal}$ 和 A, 它们的特征值和特征向量有什么关系?

A^{-1} 和 A, 它们的特征值和特征向量有什么关系?

设 $p(t)$ 是一个多项式, 则 $p(A)$ 与 A 的特征值与特征向量有什么关系?

谱半径

定义 (谱半径) 设 $A \in \mathbb{C}^{n\times n}$, 则称

$$\rho(A) \triangleq \max_{\lambda\in\sigma(A)} |\lambda|$$

为 A 的**谱半径**, 其中 $\sigma(A)$ 为 A **谱**, 即所有特征值组成的集合:

$$\sigma(A) = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}.$$

设 $A \in \mathbb{C}^{n\times n}$, 称 $\mathrm{tr}(A) \triangleq a_{11} + a_{22} + \cdots + a_{nn}$ 称为矩阵 A 的迹 (trace), 则

$$\det(A) = \lambda_1\lambda_2\cdots\lambda_n, \qquad \mathrm{tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n.$$

特征值分解 (谱分解)

定义 设 $A \in \mathbb{C}^{n\times n}$. 若存在一个非奇异矩阵 $X \in \mathbb{C}^{n\times n}$, 使得

$$X^{-1}AX = \Lambda, \tag{1.1}$$

其中 $\Lambda \in \mathbb{C}^{n\times n}$ 是对角矩阵. 则称 A 是**可对角化**的, 矩阵 Λ 的对角线元素即为 A 的特征值, 分解 (1.1) 称为矩阵 A 的**特征值分解**或**谱分解**.

其他基本概念

- 特征多项式, 最小多项式
- 代数重数和几何重数, 特征空间
- 可对角化, 可对角化的充要条件/充分条件

特征值估计：Bendixson 定理

对任意矩阵 $A \in \mathbb{C}^{n\times n}$, 我们记

$$\mathcal{H}_A \triangleq \frac{1}{2}(A + A^*), \quad \mathcal{S}_A \triangleq \frac{1}{2}(A - A^*).$$

则 $\boxed{(\mathcal{H}_A)^* = \mathcal{H}_A, \quad (\mathcal{S}_A)^* = -\mathcal{S}_A}$, 分别称为 A 的 Hermite 部分和 Skew-Hermite (斜 Hermite) 部分.

定理 (Bendixson) 设 $A \in \mathbb{C}^{n\times n}$, 则

$$\lambda_{\min}(\mathcal{H}_A) \leq \mathrm{Re}(\lambda(A)) \leq \lambda_{\max}(\mathcal{H}_A),$$

$$\lambda_{\min}(-i\mathcal{S}_A) \leq \mathrm{Im}(\lambda(A)) \leq \lambda_{\max}(-i\mathcal{S}_A),$$

其中 Re(·) 和 Im(·) 分别表示实部和虚部.

特征值估计：圆盘定理

设 $A=[a_{ij}]\in\mathbb{C}^{n\times n}$, 定义集合

$$\mathcal{D}_i \triangleq \left\{ z\in\mathbb{C} \,:\, |z-a_{ii}| \le \sum_{j=1,j\neq i}^{n} |a_{ij}| \right\}, \quad i=1,2,\ldots,n.$$

这就是 A 的 n 个 **Gerschgorin 圆盘**.

定理 (Gerschgorin 圆盘定理) 设 $A=[a_{ij}]\in\mathbb{C}^{n\times n}$. 则 A 的所有特征值都包含在 A 的 Gerschgorin 圆盘的并集中, 即 $\sigma(A)\subset\bigcup_{i=1}^{n}\mathcal{D}_i$.

将 A 的非对角线元素换成 τa_{ij}, 其中 $0 \le \tau \le 1$, 并利用特征值关于矩阵元素的连续性, 我们就可以得到下面的结论.

定理 设 $A=[a_{ij}] \in \mathbb{C}^{n\times n}$. 如果 $\bigcup\limits_{i=1}^{n} \mathcal{D}_i$ 可分解成两个不相交的子集 S 和 T, 即

$$\bigcup_{i=1}^{n} \mathcal{D}_i = \mathsf{S} \bigcup \mathsf{T} \quad 且 \quad \mathsf{S} \bigcap \mathsf{T} = \emptyset.y$$

假定 S 由 k 的圆盘组成, 而 T 由其它 $n-k$ 个圆盘组成. 则 S 中恰好包含 A 的 k 个特征值 (重特征值按重数计算), 而 T 中则包含 A 的其它 $n-k$ 个特征值.

5 向量范数与矩阵范数

向量范数

定义 (向量范数) 若函数 $f:\mathbb{C}^n\to\mathbb{R}$ 满足

(1) $f(x)\geq 0, \forall\, x\in\mathbb{C}^n$, 等号当且仅当 $x=0$ 时成立;

(2) $f(\alpha x)=|\alpha|\cdot f(x), \forall\, x\in\mathbb{C}^n, \alpha\in\mathbb{C}$;

(3) $f(x+y)\leq f(x)+f(y), \forall x, y\in\mathbb{C}^n$;

则称 $f(x)$ 为 $\mathbb{C}^n$ 上的范数, 通常记作 $\|\cdot\|$

✍ 相类似地, 我们可以定义实数空间 $\mathbb{R}^n$ 上的向量范数.

$\mathbb{C}^n$ 和 $\mathbb{R}^n$ 中常见的向量范数

- 1-范数: $\|x\|_1 = |x_1| + |x_2| + \cdots + |x_n|$
- 2-范数: $\|x\|_2 = \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2}$
- ∞-范数: $\|x\|_\infty = \max_{1\leq i\leq n} |x_i|$
- p-范数: $\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}, \quad 1 \leq p < \infty$

向量范数的等价性

定义 (范数的等价性)

$\mathbb{C}^n$ 上的向量范数 $\|\cdot\|_\alpha$ 与 $\|\cdot\|_\beta$ 等价: 存在正常数 c_1, c_2, 使得

$$c_1\|x\|_\alpha \leq \|x\|_\beta \leq c_2\|x\|_\alpha, \quad \forall\, x \in \mathbb{C}^n$$

✍ 该定义可推广到任意赋范线性空间

定理 $\mathbb{C}^n$ 空间上的所有向量范数都是等价的, 特别地, 有

$$\|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\,\|x\|_2,$$

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\,\|x\|_\infty,$$

$$\|x\|_\infty \leq \|x\|_1 \leq n\,\|x\|_\infty.$$

(证明留作练习)

✍ 事实上, 任意有限维赋范线性空间上的所有范数都是等价的

Cauchy-Schwartz 不等式

定理 设 $(\cdot,\cdot)$ 是 $\mathbb{R}^n$ 上的内积, 则对任意 $x, y \in \mathbb{R}^n$, 有

$$\boxed{|(x,y)|^2 \le (x,x)\cdot(y,y)}$$

(板书)

☞ 等号成立的条件是什么?

☞ 该结论对 $\mathbb{C}^n$ 也成立.

范数的连续性

定理 设 $\|\cdot\|$ 是 $\mathbb{C}^n$ 上的向量范数, 则 $f(x) \triangleq \|x\|$ 是 $\mathbb{C}^n$ 上的连续函数.

(证明留作练习)

矩阵范数

定义 (矩阵范数) 若函数 $f : \mathbb{C}^{n\times n} \to \mathbb{R}$ 满足

(1) $f(A) \geq 0, \forall\, A \in \mathbb{C}^{n\times n}$, 等号当且仅当 $A = 0$ 时成立; **(非负性)**

(2) $f(\alpha A) = |\alpha| \cdot f(A), \forall\, A \in \mathbb{C}^{n\times n}, \alpha \in \mathbb{C}$; **(正齐次性)**

(3) $f(A + B) \leq f(A) + f(B), \forall A, B \in \mathbb{C}^{n\times n}$; **(三角不等式)**

则称 $f(x)$ 为 $\mathbb{C}^{n\times n}$ 上的范数, 通常记作 $\|\cdot\|$.

相容的矩阵范数: (4) $f(AB) \leq f(A)f(B), \quad \forall\, A, B \in \mathbb{C}^{n\times n}$

若未明确指出, 讲本课程所涉及矩阵范数都是指 **相容矩阵范数**

算子范数

引理 设 $\|\cdot\|$ 是 $\mathbb{C}^n$ 上的向量范数, 则

$$\|A\| \triangleq \sup_{x\in\mathbb{C}^n,\, x\neq 0} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \|Ax\|$$

是 $\mathbb{C}^{n\times n}$ 上的范数, 称为算子范数, 或诱导范数, 导出范数. (板书)

✍ 算子范数都是相容的, 且

$$\|Ax\| \leq \|A\|\cdot\|x\|, \quad A\in\mathbb{C}^{n\times n}, x\in\mathbb{C}^n$$

✍ 类似地, 我们可以定义 $\mathbb{R}^{n\times n}$, $\mathbb{C}^{m\times n}$, $\mathbb{R}^{m\times n}$ 上的矩阵范数.

常见算子范数及计算公式

引理 可以证明:

(1) 1-范数 (列范数): $\|A\|_1 = \max\limits_{1\leq j\leq n} \left(\sum\limits_{i=1}^{n} |a_{ij}| \right)$

(2) ∞-范数 (行范数): $\|A\|_\infty = \max\limits_{1\leq i\leq n} \left(\sum\limits_{j=1}^{n} |a_{ij}| \right)$

(3) 2-范数 (谱范数): $\|A\|_2 = \sqrt{\rho(A^* A)}$

(证明留作练习)

✍ 另一个常用范数 *F*-范数:

$$\|A\|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2}$$

矩阵范数的等价性

定理 定义在 $\mathbb{R}^{n\times n}$ 上的所有范数都是等价的, 特别地, 有

$$
\begin{aligned}
&\frac{1}{\sqrt{n}}\|A\|_2 \le \|A\|_1 \le \sqrt{n}\,\|A\|_2,\\
&\frac{1}{\sqrt{n}}\|A\|_2 \le \|A\|_\infty \le \sqrt{n}\,\|A\|_2,\\
&\frac{1}{n}\|A\|_\infty \le \|A\|_1 \le n\,\|A\|_\infty,\\
&\frac{1}{\sqrt{n}}\|A\|_1 \le \|A\|_F \le \sqrt{n}\,\|A\|_1.
\end{aligned}
$$

(证明留作练习)

矩阵范数的一基本些性质

- 对任意的算子范数 $\|\cdot\|$, 有 $\|I\| = 1$
- 对任意的相容范数 $\|\cdot\|$, 有 $\|I\| \geq 1$
- F-范数不是算子范数
- $\|A^*\|_2 = \|A\|_2$, $\|A^*\|_1 = \|A\|_\infty$
- 若 A 是正规矩阵, 则 $\|A\|_2 = \rho(A)$
- $\|\cdot\|_2$ 和 $\|\cdot\|_F$ 是 **酉不变范数**, 即

$$\|UA\|_2 = \|A\|_2 = \|AV\|_2, \quad \|UA\|_F = \|A\|_F = \|AV\|_F,$$

其中 U, V 为任意酉矩阵.

谱半径与范数之间的关系

定理 (谱半径与范数的关系) 设 $A \in \mathbb{C}^{n\times n}$, 则

(1) 对任意算子范数, 有 $\rho(A) \leq \|A\|$;

(2) 反之, 对任意 $\varepsilon > 0$, 都存在一个算子范数 $\|\cdot\|_\varepsilon$, 使得 $\|A\|_\varepsilon \leq \rho(A) + \varepsilon$, 其中范数 $\|\cdot\|_\varepsilon$ 依赖于 A 和 ε. 所以, 若 $\rho(A) < 1$, 则存在算子范数 $\|\cdot\|_\varepsilon$, 使得 $\|A\|_\varepsilon < 1$;

(板书, 只证明结论 (1), 结论 (2) 可查询相关资料)

✍ 事实上, 上述定理中的结论 (1) 对任意矩阵范数都成立.

推论 设 $A \in \mathbb{R}^{n\times n}$, 则 $\|A\|_2^2 \leq \|A\|_1 \cdot \|A\|_\infty$, 且

$$\max_{1\leq i,j\leq n}\{|a_{ij}|\} \leq \|A\|_2 \leq n \max_{1\leq i,j\leq n}\{|a_{ij}|\}.$$

(证明留作练习)

谱半径与范数之间的一个非常重要的性质, 有时也用来定义谱半径.

引理 设 $G \in \mathbb{R}^{n\times n}$, 则对任意矩阵范数 $\|\cdot\|$, 有

$$\rho(G) = \lim_{k\to\infty} \left\|G^k\right\|^{\frac{1}{k}}.$$

6 矩阵标准型

计算矩阵特征值的一个基本思想

通过相似变换, 将其转化成一个形式尽可能简单的矩阵, 使得其特征值更易于计算.

两个非常有用的特殊矩阵是 **Jordan 标准型** 和 **Schur 标准型**

Jordan 标准型

定理 设 $A \in \mathbb{C}^{n\times n}$ 有 p 个不同特征值, 则存在非奇异矩阵 $X \in \mathbb{C}^{n\times n}$, 使得

$$X^{-1}AX = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_p \end{bmatrix} \triangleq J,$$

其中 J_i 的维数等于 λ_i 的代数重数, 且具有下面的结构

$$J_i = \begin{bmatrix} J_{i1} & & & \\ & J_{i2} & & \\ & & \ddots & \\ & & & J_{i\nu_i} \end{bmatrix}, \quad J_{ik} = \begin{bmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_i & 1 \\ & & & \lambda_i \end{bmatrix},$$

这里 ν_i 为 λ_i 的几何重数, J_{ik} 称为 **Jordan 块**, 每个 Jordan 块对应一个特征向量.

另一种描述：设 $A \in \mathbb{C}^{n\times n}$, 则存在非奇异矩阵 $X \in \mathbb{C}^{n\times n}$ 使得

$$X^{-1}AX = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_\ell \end{bmatrix},$$

其中

$$J_k = \begin{bmatrix} \lambda_k & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_k & 1 \\ & & & \lambda_k \end{bmatrix}_{n_k\times n_k}, \quad n_k \geq 1, k = 1, 2, \ldots, \ell, \text{ 且 } \sum_{k=1}^{\ell} n_k = n.$$

矩阵 J_k 就称为 **Jordan 块**, $\lambda_1, \lambda_2, \ldots, \lambda_\ell$ 是 A 的特征值 (不要求互异).

✍ Jordan 标准型在理论研究中非常有用, 但数值计算比较困难, 目前还没有找到稳定的数值算法.

基于 Jordan 标准型, 我们可以得到下面的非常重要的结论.

推论 所有可对角化矩阵组成的集合在所有矩阵组成的集合中是稠密的.

Schur 标准型

定理 设 $A\in\mathbb{C}^{n\times n}$, 则存在一个酉矩阵 $U\in\mathbb{C}^{n\times n}$ 使得

$$U^*AU=\begin{bmatrix}\lambda_1 & r_{12} & \cdots & r_{1n}\\ 0 & \lambda_2 & \cdots & r_{2n}\\ \vdots & & \ddots & \vdots\\ 0 & \cdots & 0 & \lambda_n\end{bmatrix}\triangleq R \quad \text{或} \quad A=URU^*, \tag{1.2}$$

其中 $\lambda_1,\lambda_2,\ldots,\lambda_n$ 是 A 的特征值 (排序任意). (板书)

关于 Schur 标准型的几点说明

- Schur 标准型可以说是酉相似变化下的最简形式
- U 和 R 不唯一, R 的对角线元素可按任意顺序排列
- A 是正规矩阵当且仅当 (1.2) 中的 R 是对角矩阵;
- A 是 Hermite 矩阵当且仅当 (1.2) 中的 R 是实对角矩阵.

实 Schur 标准型

定理 设 $A \in \mathbb{R}^{n\times n}$, 则存在正交矩阵 $Q \in \mathbb{R}^{n\times n}$, 使得

$$Q^{\intercal} A Q = T,$$

其中 $T \in \mathbb{R}^{n\times n}$ 是 拟上三角矩阵, 即 T 是块上三角的, 且对角块为 1×1 或 2×2 的块矩阵. 若对角块是 1×1 的, 则其就是 A 的一个特征值, 若对角块是 2×2 的, 则其特征值是 A 的一对共轭复特征值.

(留作自习)

7 对称正定矩阵

设 $A \in \mathbb{C}^{n\times n}$.

A 是 **半正定** $\iff \mathrm{Re}(x^*Ax) \geq 0, \forall\, x \in \mathbb{C}^n$

A 是 **正定** $\iff \mathrm{Re}(x^*Ax) > 0, \forall\, x \in \mathbb{C}^n, x \neq 0$

A 是 **Hermite 半正定** $\iff A$ Hermite 且半正定

A 是 **Hermite 正定** $\iff A$ Hermite 且正定

✍ 正定和半正定矩阵不要求是对称或 Hermite 的

对称正定矩阵的判别

定理 设 $A \in \mathbb{C}^{n\times n}$, 则 A 正定 (半正定) 的充要条件是矩阵

$$H = \frac{1}{2}(A + A^*)$$

正定 (半正定). (证明留作练习)

定理 设 $A \in \mathbb{R}^{n\times n}$, 则 A 正定 (或半正定) 的充要条件是对任意非零向量 $x \in \mathbb{R}^n$ 有

$$x^\intercal Ax > 0 \quad (\text{或 } x^\intercal Ax \geq 0),$$

即实矩阵只要针对实向量成立即可。 (证明留作练习)

矩阵平方根

定理 设 $A \in \mathbb{C}^{n\times n}$ 是 Hermite 半正定, k 是正整数. 则存在 **唯一** 的 Hermite 半正定矩阵 $B \in \mathbb{C}^{n\times n}$ 使得

$$B^k = A.$$

同时, 我们还有下面的性质:

(1) $BA = AB$, 且存在一个多项式 $p(t)$ 使得 $B = p(A)$;

(2) $\mathrm{rank}(B) = \mathrm{rank}(A)$, 因此, 若 A 是正定的, 则 B 也正定;

(3) 如果 A 是实矩阵的, 则 B 也是实矩阵.

✍ 特别地, 当 $k = 2$ 时, 称 B 为 A 的**平方根**, 通常记为 $A^{\frac{1}{2}}$.

Hermite 正定矩阵与内积之间的关系

定理 设 $(\cdot,\cdot)$ 是 $\mathbb{C}^n$ 上的一个内积, 则存在一个 Hermite 正定矩阵 $A\in\mathbb{C}^{n\times n}$ 使得

$$(x,y)=y^*Ax.$$

反之, 若 $A\in\mathbb{C}^{n\times n}$ 是 Hermite 正定矩阵, 则

$$f(x,y)\triangleq y^*Ax$$

是 $\mathbb{C}^n$ 上的一个内积. (证明留作练习)

✍ 该结论说明内积与 Hermite 正定矩阵是一一对应的.

✍ 该结论在实数域中也成立.