高等数值分析

Advanced Numerical Analysis

Dait

目　录

第一章　数学基础知识

1.1　线性空间

定义 1.1.1: 线性空间

设 V 是一个集合，$\mathbb{F}$ 是一个数域，定义加法 $+ : V \times V \to V$ 和数乘 $\cdot : \mathbb{F} \times V \to V$ 满足：

- 加法结合律：$(a+b)+c = a+(b+c)$
- 加法零元：$a+0=a$
- 加法逆元：$a+(-a)=0$
- 加法交换律：$a+b=b+a$
- 数乘单位元：$1a=a$
- 数乘结合律：$(\lambda\mu)a = \lambda(\mu a)$
- 数乘分配率 1：$(\lambda+\mu)a = \lambda a + \mu a$
- 数乘分配率 2：$\lambda(a+b) = \lambda a + \lambda b$

则称 V 在 $\mathbb{F}$ 上构成一个线性空间 (linear space).

例 1.1.1: 线性空间的例子

- 全体在有界闭区间 $[a,b]$ 上连续的函数构成的集合 $C[a,b]$ 是一个线性空间.
- 多项式函数空间
$$P_n = \{a_0 + a_1 x + \cdots + a_n x^n \mid a_i \in \mathbb{F}\}$$
构成一个线性空间，其加法为多项式加法.
- 正实数 $\mathbb{R}_{>0}$ 构成一个线性空间，其加法为实数乘法，数乘为幂次.

定义 1.1.2: 线性无关

设 V 是线性空间，其中 n 个元素 $x_1,\ldots,x_n \in V$，若

$$a_1 x_1 + \cdots + a_n x_n = 0, \tag{1.1}$$

只有零解，则称 $x_1,\ldots,x_n$ 线性无关 (linear independent). 反之，称为线性相关.

例 1.1.2: 线性无关的例子

- P_n 中 $\{1, x, \ldots, x^n\}$ 是线性无关的；
- 所有 $[-\pi, \pi]$ 上的周期函数构成的函数空间是线性空间，其中

$$\{1, \cos x, \sin x, \ldots, \cos nx, \sin nx\}$$

线性无关.

定义 1.1.3: 维度和基

给定线性空间 V 中的一组元素 $\{x_1, \ldots, x_n\}$，若 $\forall x \in V$ 都可以被唯一表示为其线性组合

$$x = a_1 x_1 + \cdots + a_n x_n, \tag{1.2}$$

则称 $\{x_1, \ldots, x_n\}$ 构成一组基 (base)，且 V 的维度 $\dim V = n$.

定理 1.1.1

线性空间的维度与基的选取没有关系.

例 1.1.3: 典型线性空间的维度

- $\dim P_n = n$;
- $\dim C[a, b] = +\infty$.

1.2　范数

定义 1.2.1: 度量空间

设 M 是一个集合，设映射 $d : M \times M \to \mathbb{R}$ 满足：

- $d(x, y) \geqslant 0$，且 $d(x, y) = 0 \iff x = y$；
- $d(x, y) = d(y, x)$；
- 三角不等式：$d(x, y) + d(y, z) \geqslant d(x, z)$.

则称 (M, d) 构成一个度量空间 (metric space) 或距离空间，d 为度量函数或距离函数.

定义 1.2.2: Cauchy 序列

对于序列 $\{x_n\}$，若 $\forall \varepsilon > 0, \exists N \in \mathbb{N}$ 满足：

$$d(x_n, x_{n+p}) < \varepsilon, \quad \forall p \in \mathbb{N}, \forall n > N, \tag{1.3}$$

则称 $\{x_n\}$ 为 Cauchy 序列.

定义 1.2.3: 度量空间的完备性

若对任意 M 中的 Cauchy 序列 $\{x_n\}$，$\exists x \in M$ 满足：

$$\lim_{n\to\infty} d(x_n, x) = 0,$$

则称度量空间 M 是完备的.

例 1.2.1: 实数公理

实数集 $\mathbb{R}$ 中的度量函数 $d(a,b) = |a-b|$ 是完备的.

定理 1.2.1: 完备化定理

若 (M,d) 是一个度量空间，则存在唯一等距同构的完备化空间.

证明. 构造性证明，令 $\tilde{M}$ 是 M 中所有 Cauchy 序列 $x = \{x_n\}$ 的集合. 在 $\tilde{M}$ 中定义等价关系 $\sim$：

$$x \sim y \iff \lim_{n\to\infty} d(x_n, y_n) = 0.$$

令 $[x] = \{y \,|\, x \sim y\}$ 表示 x 的等价类，$\hat{M} = \{[x] \,|\, x \in \tilde{M}\}$ 是 $\tilde{M}$ 中所有元素的等价类构成的集合. 定义 $\hat{M}$ 上的度量为

$$\hat{d}([x],[y]) = \lim_{n\to\infty} d(x_n, y_n),$$

可证 $(\hat{M}, \hat{d})$ 是完备的度量空间. 且存在等距嵌入

$$i : M \to \hat{M},\ \ x \mapsto [\{x, x, \ldots\}].$$

即映射到对应常数序列的等价类. □

定义 1.2.4: 赋范空间

若映射 $\|\cdot\| : S \to \mathbb{R}$ 满足：

- $\|f\| \geqslant 0$，且 $\|f\| = 0 \iff f = 0$；
- $\|\lambda f\| = |\lambda| \, \|f\|$；
- $\|f + g\| \leqslant \|f\| + \|g\|$.

则称 $(S, \|\cdot\|)$ 构成一个赋范空间 (normed space).

显然，赋范空间也是度量空间，只需定义

$$d(f,g) = \|f - g\|,$$

完备的赋范空间称为 Banach 空间.

例 1.2.2: $\mathbb{R}^n$ 的范数

$x = (x_1, \ldots, x_n)^\top \in \mathbb{R}^n$ 的 p - 范数为

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}.$$

如

$$\|x\|_\infty = \max_i |x_i|, \tag{1.4a}$$

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|, \tag{1.4b}$$

$$\|x\|_2 = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2}, \tag{1.4c}$$

证明. 下面给出 (1.4a) 的证明. 记 $k = \arg\max_i |x_i|$，$\forall p > 0$

$$|x_k|^p \leqslant \sum_{i=1}^{n} |x_i|^p \leqslant n\,|x_k|^p,$$

两边开 p 次方，并 $p \to \infty$，即得 $\|x\|_\infty = |x_k|$. □

例 1.2.3: $C[a,b]$ 的范数

$f(x) \in C[a,b]$ 的 p - 范数为

$$\|f\|_p = \left(\int_a^b |f(x)|^p \,\mathrm{d}x\right)^{1/p}.$$

如

$$\|f\|_\infty = \max_{a \leqslant x \leqslant b} |f(x)|, \tag{1.5a}$$

$$\|f\|_1 = \int_a^b |f(x)| \,\mathrm{d}x, \tag{1.5b}$$

$$\|f\|_2 = \left(\int_a^b |f(x)|^2 \,\mathrm{d}x\right)^{1/2}, \tag{1.5c}$$

定义 1.2.5: 范数的等价性

给定线性空间 S 上的两个范数 $\|\cdot\|_\alpha, \|\cdot\|_\beta$，若 $\exists C_1, C_2 > 0$ 满足:

$$C_1 \|x\|_\alpha \leqslant \|x\|_\beta \leqslant C_2 \|x\|_\alpha,$$

则称 $\|\cdot\|_\alpha, \|\cdot\|_\beta$ 等价.

定理 1.2.2

有限维线性空间中，任意两个范数都是等价的.

$C[a,b]$ 中，$\|\cdot\|_\infty$ 是完备的，而 $\|\cdot\|_1$ 不是.

1.3　内积

定义 1.3.1: 内积

给定线性空间 S，内积 (inner product) 是一个映射

$$\langle\cdot,\cdot\rangle : S\times S\to\mathbb{F}, \tag{1.6}$$

满足：

- $\langle x,x\rangle \geqslant 0$，且 $\langle x,x\rangle = 0 \iff x=0$;
- $\langle x,y\rangle = \overline{\langle y,x\rangle}$;
- $\langle ax+by,z\rangle = a\langle x,z\rangle + b\langle y,z\rangle$.

若 $\langle x,y\rangle = 0$，则称 x,y 是正交的 (orthogonal).

内积诱导的范数：

$$\|x\| := \sqrt{\langle x,x\rangle}, \tag{1.7}$$

例 1.3.1: 内积实例

$\mathbb{R}^n$ 的内积为

$$\langle x,y\rangle = \sum_{i=1}^n x_i y_i,$$

$C[a,b]$ 的内积为

$$\langle f,g\rangle = \int_a^b f(x)\overline{g}(x)\,\mathrm{d}x.$$

定理 1.3.1: 内积空间线性无关的判定

给定内积空间 S，$x_1,\dots,x_n\in S$ 是线性无关的 $\iff$ Gramm 矩阵满秩：

$$G_n = \begin{bmatrix} \langle x_1,x_1\rangle & \cdots & \langle x_1,x_n\rangle \\ \vdots & \ddots & \vdots \\ \langle x_n,x_1\rangle & \cdots & \langle x_n,x_n\rangle \end{bmatrix}, \tag{1.8}$$

即 $\det G_n \neq 0$.

证明. 设 $a=(a_1,\dots,a_n)$ 满足 $aG_n=0$，则 $\forall k=1,\dots,n$

$$(aG_n)_k = a_1\langle x_1,x_k\rangle + \cdots + a_n\langle x_n,x_k\rangle = \langle a_1x_1+\cdots, a_nx_n, x_k\rangle = 0,$$

特别的，

$$\langle a_1x_1+\cdots+a_nx_n, a_1x_1+\cdots+a_nx_n\rangle=0, \iff a_1x_1+\cdots, a_nx_n=0,$$

则 $\det G_n\neq 0 \iff \mathrm{N}(G_n^\top)=\{0\} \iff a$ 只有零解. □

例 1.3.2

给定内积空间 S 的一组基 $\{x_1,\dots,x_n\}$，则 $\forall x\in S$ 均可以写成

$$x=a_1x_1+\cdots+a_nx_n,$$

下面计算 $a_1,\dots,a_n$. 两边分别与 x_i 做内积：

$$\langle x,x_i\rangle=a_1\langle x_1,x_i\rangle+\cdots+a_n\langle x_n,x_i\rangle,$$

即

$$(\langle x,x_1\rangle,\dots,\langle x,x_n\rangle)=(a_1,\dots,a_n)G_n,$$

若基是正交的，即 $\forall i\neq j$, $\langle x_i,x_j\rangle=0$，则

$$a_i=\frac{\langle x,x_i\rangle}{\langle x_i,x_i\rangle}.$$

定理 1.3.2: Schmidt 正交化

设 $x_1,\dots,x_n$ 是一组线性无关的基，为得到一组正交基，定义

$$y_i=x_i-\sum_{j=1}^{i-1}\frac{\langle x_i,y_j\rangle}{\langle y_j,y_j\rangle}y_j. \tag{1.9}$$

则 $y_1,\dots,y_n$ 是正交的.

定义 1.3.2: 带权内积

设 $\rho\in C(a,b)$ 是一个几乎处处为正[I]的函数，且

$$\int_a^b\rho(x)\,\mathrm{d}x<+\infty,$$

定义内积

$$\langle f,g\rangle=\int_a^b f(x)g(x)\rho(x)\,\mathrm{d}x. \tag{1.10}$$

[I]即 $\{x|\rho(x)\leqslant 0\}$ 的 Lebesgue 测度为 0.

定义 1.3.3: 正交多项式

已知 $\{1,x,\dots,x^n\}$ 是线性无关的. 考虑 $C[a,b]$ 上的带权内积，Schmidt 正交化得到一

组多项式函数

$$\psi_0(x), \psi_1(x), \ldots, \psi_n(x),$$

显然，$\deg \psi_i = i$.

例 1.3.3: Legendre 多项式

权函数 $\rho = 1$，区间 $[-1, 1]$，得到 Legendre 多项式

$$P_n(x) = \frac{1}{2^n n!} \frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^2 - 1)^n. \tag{1.11}$$

- 内积

$$\langle P_n, P_m \rangle = \frac{2}{2n+1} \delta_{nm}. \tag{1.12}$$

- 递归关系

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x). \tag{1.13}$$

- 奇偶性

$$P_n(-x) = (-1)^n P_n(x). \tag{1.14}$$

例 1.3.4: Chebyshev 多项式

权函数 $\rho(x) = (1 - x^2)^{-1/2}$，区间 $[-1, 1]$，得到 Chebyshev 多项式

$$T_n(x) = \cos(n \arccos x). \tag{1.15}$$

- 内积

$$\langle T_0, T_0 \rangle = \pi; \quad \langle T_n, T_n \rangle = \pi/2, \ \ n \geqslant 1; \tag{1.16}$$

- 递推关系

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x). \tag{1.17}$$

- 奇偶性

$$T_n(-x) = (-1)^n T_n(x); \tag{1.18}$$

- T_n 的 n 个实单根为 $\cos\left(\frac{2k-1}{2n}\pi\right)$，$(n+1)$ 个极值点为 $\cos\left(\frac{k}{n}\pi\right)$
- 当 $|x| \geqslant 1$ 时，

$$T_n(x) = \frac{1}{2}\left[\left(x + \sqrt{x^2 - 1}\right)^k + \left(x - \sqrt{x^2 - 1}\right)^k\right]. \tag{1.19}$$

1.4 矩阵空间

定义 1.4.1: 矩阵范数

矩阵空间 $\mathbb{C}^{n\times n}$ 上的范数 $\|\cdot\|$ 若满足

$$\|AB\| \leqslant \|A\| \, \|B\| , \tag{1.20}$$

则称该范数为矩阵范数.

例 1.4.1: Frobenius 范数

定义 Frobenius 范数

$$\|A\|_{\mathrm{F}} := \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} |A_{ij}|^2}. \tag{1.21}$$

是一个矩阵范数. 注意 $\|I\|_{\mathrm{F}} = \sqrt{n}$.

证明. 由 Cauchy-Schwarz 不等式

$$\begin{aligned}\|AB\|_{\mathrm{F}} &= \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}\left|\sum_{k=1}^{n} A_{ik}B_{kj}\right|^2} \leqslant \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\sum_{k=1}^{n}|A_{ik}|^2\sum_{k=1}^{n}|B_{kj}|^2\right)} \\ &= \sqrt{\sum_{i=1}^{n}\sum_{k=1}^{n}|A_{ik}|^2\sum_{j=1}^{n}\sum_{k=1}^{n}|B_{kj}|^2} = \|A\|_{\mathrm{F}} \, \|B\|_{\mathrm{F}} .\end{aligned}$$

□

定义 1.4.2: 矩阵范数与向量范数的相容

给定矩阵范数 $\|\cdot\|_M$ 和向量范数 $\|\cdot\|_V$，若 $\forall A \in \mathbb{F}^{n\times n}, x \in \mathbb{F}^n$

$$\|Ax\|_V \leqslant \|A\|_M \, \|x\|_V , \tag{1.22}$$

则称他们是相容的. 在不引起混淆的情况下，可以略去下标.

例 1.4.2

Frobenius 范数与向量 2 - 范数相容.

证明.

$$\|Ax\|_2 = \sqrt{\sum_{i=1}^{n}\left|\sum_{j=1}^{n} A_{ij}x_j\right|^2} \leqslant \sqrt{\sum_{i=1}^{n}\left(\sum_{j=1}^{n}|A_{ij}|^2\sum_{j=1}^{n}|x_j|^2\right)} = \|A\|_{\mathrm{F}} \, \|x\|_2 .$$

□

定义 1.4.3: 算子范数

定义矩阵的算子范数 (operate norm)

$$N : \mathbb{C}^{n\times n} \to \mathbb{C},\ \ A \mapsto \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|}. \tag{1.23}$$

称算子范数是该向量范数诱导出来的矩阵范数.

注意 $N(I)=1$. 故 Frobenius 范数不是算子范数.

定理 1.4.1

$N(\cdot)$ 是一个矩阵范数，并与向量范数相容.

证明. 易得 $\forall x \neq 0,\ \|Ax\| \leqslant N(A)\,\|x\|$.

$$N(AB) = \sup_{x\neq 0} \frac{\|ABx\|}{\|x\|} \leqslant \sup_{x\neq 0} \frac{N(A)\,\|Bx\|}{\|x\|} = N(A)N(B). \qquad \square$$

定义 1.4.4: 谱半径

矩阵 A 全体特征值的集合称为 A 的谱，记作 $\sigma(A)$，特征值模的最大值称为谱半径，记作 $\rho(A)$.

例 1.4.3

设 $A\in\mathbb{C}^{n\times n}$，则 p - 向量范数诱导出来的矩阵范数为:

$$\|A\|_\infty = \max_i \sum_{j=1}^n |A_{ij}|\,, \tag{1.24a}$$

$$\|A\|_1 = \max_j \sum_{i=1}^n |A_{ij}|\,, \tag{1.24b}$$

$$\|A\|_2 = \sqrt{\rho(A^\dagger A)}, \tag{1.24c}$$

证明. 先证明 (1.24b)，记 $A=(A_1,\dots,A_n)$，即 A_j 为 A 的第 j 列列向量，令 $k=\arg\max\limits_j \|A_j\|_1$，则 $\forall x\in\mathbb{C}^n$ 且 $\|x\|_1=1$，有

$$\|Ax\|_1 = \left\|\sum_{j=1}^n A_j x_j\right\|_1 \leqslant \sum_{j=1}^n |x_j|\,\|A_j\|_1 \leqslant \|A_k\|_1 \sum_{j=1}^n |x_j| = \|A_k\|_1\,,$$

特别的，取 $x=e_k$ 有 $\|Ae_k\|_1 = \|a_k\|_1$，故

$$\|A\|_1 = \sup_{\|x\|_1=1} \|Ax\|_1 = \|A_k\|_1 = \max_j \sum_{i=1}^n |A_{ij}|\,;$$

然后证明 (1.24a)，记 $k' = \arg\max_i \sum_{j=1}^{n} |A_{ij}|$，则 $\forall x \in \mathbb{C}^n$ 且 $\|x\|_\infty = 1$，有

$$\|Ax\|_\infty = \max_i \left| \sum_{j=1}^{n} A_{ij} x_j \right| =$$

最后证明 (1.24c)，由 2 - 范数的性质：

$$\|A\|_2^2 = \sup_{\|x\|_2=1} \langle Ax, Ax \rangle = \sup_{\|x\|_2=1} \left\langle A^\dagger Ax, x \right\rangle = \rho(A^\dagger A).$$ □

定理 1.4.2

谱半径 $\rho(A)$ 和矩阵范数的关系：

$$\rho(A) \leqslant \|A\|. \tag{1.25}$$

证明. 考虑 A 的一个特征值 λ 和特征向量 x，则

$$|\lambda| \, \|xx'\| = \|Axx'\| \leqslant \|A\| \, \|xx'\|.$$

于是 $\|A\| \geqslant |\lambda|$. □

定理 1.4.3

$\forall A \in \mathbb{C}^{n\times n}, \epsilon > 0$，存在算子范数 $\|\cdot\|$ 满足：

$$\|A\| \leqslant \rho(A) + \epsilon. \tag{1.26}$$

引理. 若 $\|\cdot\|_\alpha$ 是 $\mathbb{C}^n$ 中的向量范数，$P \in \mathbb{C}^{n\times n}$ 非奇异，则

$$\|\cdot\|_{P,\alpha} : x \mapsto \|Px\|_\alpha ,$$

构成另一个向量范数，诱导的算子范数为

$$\|A\|_{P,\alpha} = \|PAP^{-1}\|_\alpha.$$

证明. 令 P 将 A 相似变换为 Jordan 型，即

$$PAP^{-1} = J = \operatorname{diag}(J_1, \ldots, J_r), \quad J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}$$

令 $D_\epsilon = \operatorname{diag}(1, \epsilon, \ldots, \epsilon^{n-1})$，则

$$\hat{J} = D_\epsilon^{-1} J D_\epsilon = \operatorname{diag}(\hat{J}_1, \ldots, \hat{J}_r), \quad \hat{J}_i = \begin{bmatrix} \lambda_i & \epsilon & & \\ & \ddots & \ddots & \\ & & \ddots & \epsilon \\ & & & \lambda_i \end{bmatrix}$$

则
$$\|A\|_{D_\epsilon^{-1}P,\infty} = \|D_\epsilon^{-1}PAP^{-1}D_\epsilon\|_\infty = \|\hat{J}\|_\infty \leqslant \rho(A)+\epsilon. \qquad \square$$

注. 对任意满足 $|\lambda| = \rho(A)$ 的特征值 λ，对应 Jordan 块是对角的，则存在一个算子范数满足
$$\|A\| = \rho(A).$$

任意给定一个矩阵，其为奇异矩阵的概率是很低的．那如何度量矩阵的奇异性？

定理 1.4.4: 扰动定理

给定扰动 B，若 $\|B\| < 1$，则 $I+B$ 非奇异且
$$\|(I+B)^{-1}\| \leqslant \frac{1}{1-\|B\|}. \tag{1.27}$$

证明. 若 $I+B$ 是奇异的，则 $\rho(B) \geqslant 1 > \|B\|$ 矛盾．记 $D = (I+B)^{-1}$，则
$$(I+B)D = I, \iff D = I - BD, \implies \|D\| \leqslant 1 + \|B\|\,\|D\|\,. \qquad \square$$

定理 1.4.5: 扰动定理・二

给定 A, C，若 A 非奇异且
$$\|C-A\| \leqslant \|A^{-1}\|^{-1},$$
则 C 也非奇异，且
$$\|C^{-1}\| \leqslant \frac{1}{\|A^{-1}\|^{-1} - \|C-A\|}. \tag{1.28}$$