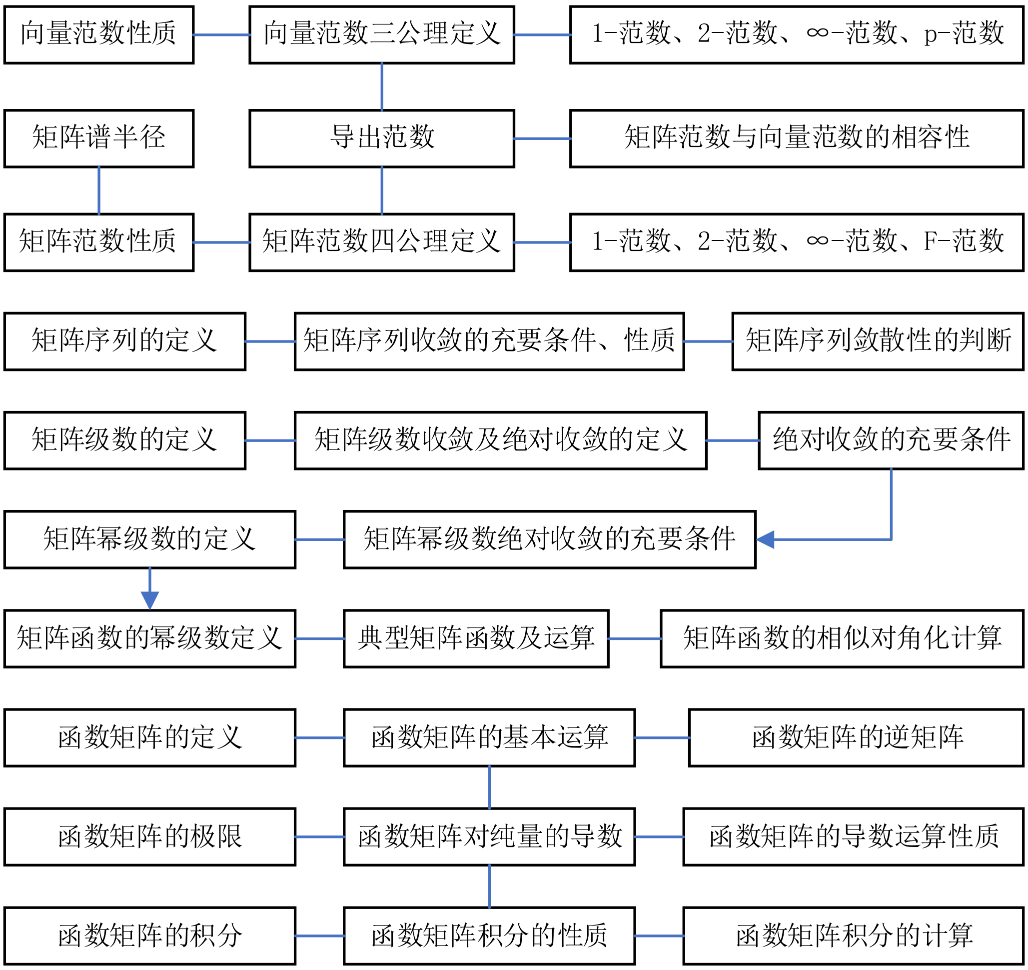
第 6 章　矩阵的微积分

本章首先介绍向量范数和矩阵范数的定义及性质、二者的相容性，给出由向量范数形式推广的矩阵范数和由向量范数导出的矩阵范数．基于矩阵范数概念，介绍矩阵序列及其收敛和极限的概念，研究矩阵序列的敛散性．其次，介绍矩阵级数和矩阵幂级数的定义、性质和收敛性，介绍矩阵函数的幂级数定义和矩阵函数计算．最后，介绍函数矩阵的定义及运算、函数矩阵的极限、微分和积分．

本章的知识网络框图：



6.1 向量和矩阵的范数

6.1.1 向量范数

范数是实数的绝对值或复数的模等表示大小的概念的普遍化. 在第 1 章中把向量概念推广到线性空间并在线性空间中引入了内积运算, 进而对向量赋予向量长度和向量夹角的概念. 但是, 向量范数是比向量长度更一般的概念, 引入内积后向量的长度是唯一的, 但向量的范数并不唯一, 可以定义多种向量范数.

定义 6.1.1 设V是数域$\mathbb{F}$上的线性空间, 用$\|\boldsymbol{x}\|$表示按照某个法则确定的与向量$\boldsymbol{x}$对应的实数, 且满足

(1) 非负性: 当$\boldsymbol{x} \neq \boldsymbol{0}$, $\|\boldsymbol{x}\| > 0$; 当且仅当$\boldsymbol{x} = \boldsymbol{0}$时, $\|\boldsymbol{x}\| = 0$;

(2) 绝对齐次性: $\|k\boldsymbol{x}\| = |k|\|\boldsymbol{x}\|$, 其中$k$为任意数;

(3) 三角不等式: 对于V中任何向量$\boldsymbol{x}, \boldsymbol{y}$都有$\|\boldsymbol{x}+\boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$.

则称实数$\|\boldsymbol{x}\|$是向量$\boldsymbol{x}$的**范数.**

定义 6.1.1 并未给出向量范数的计算方法, 只是规定了向量范数应满足的 3 个条件, 这 3 个条件称为**向量范数公理**.

例 6.1.1 n维欧氏空间中向量的模为$|\boldsymbol{x}| = (\boldsymbol{x}, \boldsymbol{x})^{\frac{1}{2}}$, 它具有如下三个性质:

(1) 若$\boldsymbol{x} \neq \boldsymbol{0}$, 则$|\boldsymbol{x}| > 0$; 若$\boldsymbol{x} = \boldsymbol{0}$, 则$|\boldsymbol{x}| = 0$; (非负性)

(2) $|k\boldsymbol{x}| = |k||\boldsymbol{x}|$, k为任意实数; (绝对齐次性)

(3) 对于任何向量$\boldsymbol{x}$和$\boldsymbol{y}$, 有$|\boldsymbol{x}+\boldsymbol{y}| \leq |\boldsymbol{x}| + |\boldsymbol{y}|$. (三角不等式)

由向量范数定义不难得出如下基本性质.

定理 6.1.1 对任意向量$\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$, 有

(1) $\|-\boldsymbol{x}\| = \|\boldsymbol{x}\|$;

(2) $|\,\|\boldsymbol{x}\| - \|\boldsymbol{y}\|\,| \leq \|\boldsymbol{x}-\boldsymbol{y}\|$.

证: (1) 由齐次性易得.

(2) 因为 $\|\boldsymbol{x}\| = \|\boldsymbol{x}-\boldsymbol{y}+\boldsymbol{y}\| \leq \|\boldsymbol{x}-\boldsymbol{y}\| + \|\boldsymbol{y}\|$

故

$$\|\boldsymbol{x}\| - \|\boldsymbol{y}\| \leq \|\boldsymbol{x}-\boldsymbol{y}\| \tag{*}$$

又

$$\|\boldsymbol{y}\| = \|\boldsymbol{x} - \boldsymbol{x} + \boldsymbol{y}\| \le \|\boldsymbol{x}\| + \|-\boldsymbol{x} + \boldsymbol{y}\| = \|\boldsymbol{x}\| + \|\boldsymbol{x} - \boldsymbol{y}\|$$

故

$$\|\boldsymbol{x}\| - \|\boldsymbol{y}\| \ge -\|\boldsymbol{x} - \boldsymbol{y}\| \tag{**}$$

综合式(*)和(**)，得

$$|\,\|\boldsymbol{x}\| - \|\boldsymbol{y}\|\,| \le \|\boldsymbol{x} - \boldsymbol{y}\|.$$ ■

以下是 Hölder（赫尔德）不等式：设正实数p，q满足$\frac{1}{p} + \frac{1}{q} = 1$，则对任意$\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$，有

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

特别地，若$p = q = 2$，就是 Schwarz（施瓦茨）不等式：

$$\sum_{i=1}^{n} |x_i y_i| \le \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} |y_i|^2\right)^{\frac{1}{2}}$$

以下是 Minkowski（闵可夫斯基）不等式：对任意实数$p \ge 1$和任意$\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$，有

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

由 Minkowski 不等式可以引入p-范数.

定义 6.1.2　对任意正数$p \ge 1$，称向量$\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$的函数

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

为向量$\boldsymbol{x}$的p-**范数**.

显然，$\|\boldsymbol{x}\|_p$满足非负性和齐次性，又由 Minkowski 不等式，$\|\boldsymbol{x} + \boldsymbol{y}\|_p \le \|\boldsymbol{x}\|_p + \|\boldsymbol{y}\|_p$，所以$\|\boldsymbol{x}\|_p$是向量范数.

常用的p-范数有下述三种：

① 1-范数：$\|\boldsymbol{x}\|_1 = \sum_{i=1}^{n} |x_i|$

② 2-范数（欧氏范数）：$\|\boldsymbol{x}\|_2=\left(\sum_{i=1}^{n}|x_i|^2\right)^{\frac{1}{2}}=\left(\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}\right)^{\frac{1}{2}}$

③ ∞-范数: $\|\boldsymbol{x}\|_\infty=\lim_{p\to\infty}\|\boldsymbol{x}\|_p$

定理 6.1.2 对向量$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}\in\mathbb{C}^n$，有$\|\boldsymbol{x}\|_\infty=\max|x_i|$，$i=1,2,\cdots,n$.

证: 令$\alpha=\max\limits_{1\le i\le n}|x_i|$，则

$$\beta_i=\frac{|x_i|}{\alpha}\le 1\qquad(i=1,2,\cdots,n),$$

于是$\|\boldsymbol{x}\|_p=\alpha\left(\sum_{i=1}^{n}\beta_i^p\right)^{\frac{1}{p}}$，由于

$$1\le\left(\sum_{i=1}^{n}\beta_i^p\right)^{\frac{1}{p}}\le n^{\frac{1}{p}},$$

故

$$\lim_{p\to\infty}\left(\sum_{i=1}^{n}\beta_i^p\right)^{\frac{1}{p}}=1.$$

因此

$$\|\boldsymbol{x}\|_\infty=\lim_{p\to\infty}\|\boldsymbol{x}\|_p=\alpha=\max_{1\le i\le n}|x_i|.\qquad\blacksquare$$

在线性空间中可以引进各种范数，按照不同法则规定的向量范数的大小一般不相等. 例如对$\mathbb{R}^n$中的向量$\boldsymbol{x}=(1,1,\cdots,1)^{\mathrm{T}}$，有

$$\|\boldsymbol{x}\|_1=n,\ \|\boldsymbol{x}\|_2=\sqrt{n},\ \|\boldsymbol{x}\|_\infty=1.$$

虽然同一向量的不同范数的值不同，但这些范数之间存在关系. 例如，在研究向量序列收敛性时，不同范数表现出一致性，这种性质称为向量范数的等价性.

定理 6.1.3 设V是n维线性空间，$\|\boldsymbol{x}\|_\alpha$和$\|\boldsymbol{x}\|_\beta$为任意两种向量范数，则总存在正数$c_1$和$c_2$，对$V$中所有向量$\boldsymbol{x}\in V$，恒有

$$c_1\|\boldsymbol{x}\|_\beta\le\|\boldsymbol{x}\|_\alpha\le c_2\|\boldsymbol{x}\|_\beta.$$

6.1.2 矩阵范数

定义 6.1.3 对于矩阵$\boldsymbol{A}\in\mathbb{C}^{m\times n}$，用$\|\boldsymbol{A}\|$表示按照某个法则确定的与$\boldsymbol{A}$对应的实数，且满足

(1) 非负性: 当$\boldsymbol{A}\ne\boldsymbol{0}$时，$\|\boldsymbol{A}\|>0$；当且仅当$\boldsymbol{A}=\boldsymbol{0}$时，$\|\boldsymbol{A}\|=0$;

(2) 绝对齐次性: $\|k\boldsymbol{A}\|=|k|\|\boldsymbol{A}\|$，其中$k$为任意数;

(3) 三角不等式: 对于任何两个可加矩阵$\boldsymbol{A},\boldsymbol{B}$都有$\|\boldsymbol{A}+\boldsymbol{B}\| \le \|\boldsymbol{A}\|+\|\boldsymbol{B}\|$;

(4) 乘法相容性: 若矩阵$\boldsymbol{A}$与$\boldsymbol{B}$可乘, 则$\|\boldsymbol{AB}\| \le \|\boldsymbol{A}\|\|\boldsymbol{B}\|$.

则称实数$\|\boldsymbol{A}\|$是矩阵$\boldsymbol{A}$的**矩阵范数**.

矩阵范数定义中的前 3 个条件与向量范数一致, 因此矩阵范数与向量范数所具有的性质类似, 如$\|-\boldsymbol{A}\|=\|\boldsymbol{A}\|$, $|\,\|\boldsymbol{A}\|-\|\boldsymbol{B}\|\,| \le \|\boldsymbol{A}-\boldsymbol{B}\|$. 矩阵范数的绝对齐次性是针对矩阵与数相乘以后范数的变化特性, 矩阵范数的三角不等式是针对矩阵相加运算以后范数的变化特性, 这都与向量范数相同. 但是, 矩阵运算还包括矩阵与矩阵的乘法, 因此在矩阵范数定义中增加了乘法相容性要求. 这 4 个条件称为**矩阵范数公理**.

由于一个$m \times n$矩阵可以看作一个mn维向量, 因此有些向量范数可以直接推广为矩阵范数, 但矩阵之间还有矩阵乘法运算, 并由乘法相容性公理约束, 因此, 能否推广关键要看是否满足乘法相容性.

将向量范数形式推广到矩阵范数时, 有的符合乘法相容性, 有的则不符合. 例如将向量的 1-范数形式上推广到矩阵范数$\|\boldsymbol{A}\|=\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|$, 或将向量的 2-范数形式上推广到矩阵范数$\|\boldsymbol{A}\|_F=\left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\right)^{1/2}$, 二者都满足乘法相容性. 其中, 由向量的 2-范数形式推广而来的矩阵范数称为 Frobenius（弗罗贝尼乌斯）范数, 简称 F-范数: 若$\boldsymbol{A}=(a_{ij})\in\mathbb{C}^{m\times n}$, 规定$\|\boldsymbol{A}\|_F=\left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\right)^{\frac{1}{2}}$, 可以证明: 该范数定义满足矩阵范数的 4 个性质.

但是, 将向量的∞-范数形式推广到矩阵范数$\|\boldsymbol{A}\|=\max\limits_{i,j}|a_{ij}|$不满足乘法相容性. 例如, 取$\boldsymbol{A}=\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}$, $\boldsymbol{B}=\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}$, 则$\boldsymbol{AB}=\begin{bmatrix}1 & 2\\ 1 & 2\end{bmatrix}$, 若将向量的∞-范数形式推广到矩阵范数, 则$\|\boldsymbol{A}\|=1$, $\|\boldsymbol{B}\|=1$, $\|\boldsymbol{AB}\|=2$, 不满足乘法相容性.

定理 6.1.4　F-范数有如下几个性质:

(1) 若$\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$, 则$\|\boldsymbol{A}\|_F^2=\sum_{i=1}^{n}\|\boldsymbol{\alpha}_i\|_2^2$.

(2) $\|\boldsymbol{A}\|_F^2=\mathrm{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})=\sum_{i=1}^{n}\lambda_{\mathrm{i}}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})$. 其中, $\mathrm{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})$是$n$阶方阵$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$的迹, $\lambda_{\mathrm{i}}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})$表示$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$的第$i$个特征值. 该性质表明矩阵的 F-范数的平方等于矩阵的所有奇异值的平方和.

(3) 对于任何m阶酉矩阵$\boldsymbol{U}$与n阶酉矩阵$\boldsymbol{V}$, 都有

$$\|\boldsymbol{A}\|_F=\|\boldsymbol{UA}\|_F=\|\boldsymbol{A}^{\mathrm{H}}\|_F=\|\boldsymbol{AV}\|_F=\|\boldsymbol{UAV}\|_F.$$

矩阵范数也有等价性定理.

定理 6.1.5　若$\|\boldsymbol{A}\|_\alpha$与$\|\boldsymbol{A}\|_\beta$是任意两种矩阵范数，则总存在正数$c_1$和$c_2$，对任意矩阵$\boldsymbol{A}$，恒有

$$c_1\|\boldsymbol{A}\|_\beta \le \|\boldsymbol{A}\|_\alpha \le c_2\|\boldsymbol{A}\|_\beta. \tag{6.1.1}$$

上述定理表明$\mathbb{C}^{m\times n}$上的任意两个矩阵范数等价.

6.1.3　向量范数与矩阵范数的相容性

设$\boldsymbol{A}\in\mathbb{C}^{m\times n}$，$\boldsymbol{x}\in\mathbb{C}^n$，若已知矩阵范数$\|\boldsymbol{A}\|$，可将$\boldsymbol{x}$与$\boldsymbol{Ax}$作为 $n\times 1$ 与 $m\times 1$ 矩阵，因此根据矩阵范数的相容性应有

$$\|\boldsymbol{Ax}\| \le \|\boldsymbol{A}\|\|\boldsymbol{x}\|$$

但是，$\boldsymbol{x}$与$\boldsymbol{Ax}$终究是向量，若取$\|\boldsymbol{x}\|_\alpha$与$\|\boldsymbol{Ax}\|_\alpha$是向量范数，$\|\boldsymbol{A}\|$是矩阵范数，则不等式

$$\|\boldsymbol{Ax}\|_\alpha \le \|\boldsymbol{A}\|\|\boldsymbol{x}\|_\alpha$$

是否仍能成立？这就是向量范数与矩阵范数的相容性问题.

定义 6.1.4　设$\|\boldsymbol{x}\|_\alpha$是向量范数，$\|\boldsymbol{A}\|_\beta$是矩阵范数，若对于任何矩阵$\boldsymbol{A}$与向量$\boldsymbol{x}$都有

$$\|\boldsymbol{Ax}\|_\alpha \le \|\boldsymbol{A}\|_\beta\|\boldsymbol{x}\|_\alpha \tag{6.1.2}$$

则称$\|\boldsymbol{A}\|_\beta$是与向量范数$\|\boldsymbol{x}\|_\alpha$相容的矩阵范数.

定理 6.1.6　设$\|\boldsymbol{x}\|_\alpha$是向量范数，则

$$\|\boldsymbol{A}\|_i = \max_{\boldsymbol{x}\ne\boldsymbol{0}}\frac{\|\boldsymbol{Ax}\|_\alpha}{\|\boldsymbol{x}\|_\alpha} \tag{6.1.3}$$

满足矩阵范数定义，且$\|\boldsymbol{A}\|_i$是与向量范数$\|\boldsymbol{x}\|_\alpha$相容的矩阵范数.

证：　非负性和齐次性显然满足.

根据向量范数三角不等式可得

$$\|\boldsymbol{A}+\boldsymbol{B}\|_i = \max_{\boldsymbol{x}\ne\boldsymbol{0}}\frac{\|(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{x}\|_\alpha}{\|\boldsymbol{x}\|_\alpha} = \max_{\boldsymbol{x}\ne\boldsymbol{0}}\frac{\|\boldsymbol{Ax}+\boldsymbol{Bx}\|_\alpha}{\|\boldsymbol{x}\|_\alpha}$$

$$\leq \max_{\boldsymbol{x}\neq \boldsymbol{0}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_\alpha + \|\boldsymbol{B}\boldsymbol{x}\|_\alpha}{\|\boldsymbol{x}\|_\alpha} = \max_{\boldsymbol{x}\neq \boldsymbol{0}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_\alpha}{\|\boldsymbol{x}\|_\alpha} + \max_{\boldsymbol{x}\neq \boldsymbol{0}} \frac{\|\boldsymbol{B}\boldsymbol{x}\|_\alpha}{\|\boldsymbol{x}\|_\alpha}$$

$$= \|\boldsymbol{A}\|_i + \|\boldsymbol{B}\|_i.$$

再证矩阵范数的相容性.

设$\boldsymbol{B} \neq 0$，则

$$\|\boldsymbol{A}\boldsymbol{B}\|_i = \max_{\boldsymbol{x}\neq \boldsymbol{0}} \frac{\|\boldsymbol{A}\boldsymbol{B}\boldsymbol{x}\|_\alpha}{\|\boldsymbol{x}\|_\alpha} = \max_{\boldsymbol{x}\neq \boldsymbol{0}} \left(\frac{\|\boldsymbol{A}(\boldsymbol{B}\boldsymbol{x})\|_\alpha}{\|\boldsymbol{B}\boldsymbol{x}\|_\alpha} \frac{\|\boldsymbol{B}\boldsymbol{x}\|_\alpha}{\|\boldsymbol{x}\|_\alpha} \right)$$

$$\leq \max_{\boldsymbol{x}\neq \boldsymbol{0}} \frac{\|\boldsymbol{A}(\boldsymbol{B}\boldsymbol{x})\|_\alpha}{\|\boldsymbol{B}\boldsymbol{x}\|_\alpha} \max_{\boldsymbol{x}\neq \boldsymbol{0}} \frac{\|\boldsymbol{B}\boldsymbol{x}\|_\alpha}{\|\boldsymbol{x}\|_\alpha}$$

$$= \|\boldsymbol{A}\|_i \|\boldsymbol{B}\|_i.$$

因此$\|\boldsymbol{A}\|_i$确实是矩阵范数.

最后证明$\|\boldsymbol{A}\|_i$与$\|\boldsymbol{x}\|_\alpha$相容. 由式(6.1.3)，得$\|\boldsymbol{A}\|_i \geq \frac{\|\boldsymbol{A}\boldsymbol{x}\|_\alpha}{\|\boldsymbol{x}\|_\alpha}$，即

$$\|\boldsymbol{A}\boldsymbol{x}\|_\alpha \leq \|\boldsymbol{A}\|_i \|\boldsymbol{x}\|_\alpha,$$

这表明$\|\boldsymbol{A}\|_i$与$\|\boldsymbol{x}\|_\alpha$相容. ■

由定理 6.1.6 所定义的矩阵范数称为由向量范数$\|\boldsymbol{x}\|_\alpha$所导出的**导出范数**或**从属范数**. 显然，单位矩阵$\boldsymbol{E}$的任何导出范数$\|\boldsymbol{E}\|_i = 1$.

由向量 p-范数$\|x\|_p$所导出的矩阵范数称为**矩阵p-范数**，即

$$\|\boldsymbol{A}\|_p = \max_{\boldsymbol{x}\neq \boldsymbol{0}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_p}{\|\boldsymbol{x}\|_p},$$

常用的矩阵p-范数为$\|\boldsymbol{A}\|_1$，$\|\boldsymbol{A}\|_2$与$\|\boldsymbol{A}\|_\infty$. 这三个范数的计算由下述定理给出.

定理 6.1.7 设矩阵$\boldsymbol{A} = (a_{ij})_{m\times n}$，则

(1) $\|\boldsymbol{A}\|_1 = \max\limits_j \left(\sum_{i=1}^m |a_{ij}|\right)$，$j = 1,2,\cdots,n$. 称$\|\boldsymbol{A}\|_1$是**列和范数**，简称**列范数**或 1−**范数**.

(2) $\|\boldsymbol{A}\|_2 = \max\limits_j (\lambda_j(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}))^{\frac{1}{2}}$，其中，$\lambda_j(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})$表示矩阵$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$的第 j 个特征值. $\|\boldsymbol{A}\|_2$是$\boldsymbol{A}$的奇异值的最大值，称$\|\boldsymbol{A}\|_2$是**谱范数**或 **2−范数**.

(3) $\|\boldsymbol{A}\|_\infty = \max\limits_i \left(\sum_{j=1}^n |a_{ij}|\right)$，$i = 1,2,\cdots,m$. 称$\|\boldsymbol{A}\|_\infty$是**行和范数**，简称**行范数**或

∞−范数.

由向量范数可以按定理 6.1.6 导出矩阵范数，且这两个范数是相容的. 即使一个矩阵范数不是某个向量范数导出的，它们之间也可能是相容的. 给定一个矩阵范数，总可以构造向量范数使这两个范数相容，这就是下述定理.

定理 6.1.8　设$\|\boldsymbol{A}\|_*$是矩阵范数，则存在向量范数$\|\boldsymbol{x}\|$，满足

$$\|\boldsymbol{A}\boldsymbol{x}\| \le \|\boldsymbol{A}\|_*\|\boldsymbol{x}\|.$$

证：　任给非零向量$\boldsymbol{\alpha}$，定义向量范数$\|\boldsymbol{x}\| = \|\boldsymbol{x}\boldsymbol{\alpha}^{\mathrm{H}}\|_*$. 不难验证它满足向量范数的三个性质，且

$$\|\boldsymbol{A}\boldsymbol{x}\| = \|\boldsymbol{A}\boldsymbol{x}\boldsymbol{\alpha}^{\mathrm{H}}\|_* \le \|\boldsymbol{A}\|_*\|\boldsymbol{x}\boldsymbol{\alpha}^{\mathrm{H}}\|_* = \|\boldsymbol{A}\|_*\|\boldsymbol{x}\| \qquad \blacksquare$$

由定理的证明可知，满足相容性条件的向量范数不止一个.

例 6.1.2　已知矩阵范数$\|\boldsymbol{A}\|_* = \|\boldsymbol{A}\|_F = \left(\sum_{i=1}^{m}\sum_{j=1}^{n}\left|a_{ij}\right|^2\right)^{\frac{1}{2}}$，求与之相容的一个向量范数.

解：　取$\boldsymbol{\alpha} = (1,0,\cdots,0)^{\mathrm{T}}$，设$\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$，则

$$\|\boldsymbol{x}\| = \|\boldsymbol{x}\boldsymbol{\alpha}^{\mathrm{H}}\|_* = \left(\sum_{i=1}^{n}|x_i|^2\right)^{\frac{1}{2}} = \|\boldsymbol{x}\|_2.$$

定义 6.1.5　设矩阵$\boldsymbol{A} \in \mathbb{C}^{n\times n}$，$\boldsymbol{A}$的 n 个特征值为$\lambda_1, \lambda_2, \cdots, \lambda_n$，称$\rho(\boldsymbol{A}) = \max\{|\lambda_1|, |\lambda_2|, \cdots, |\lambda_n|\}$是$\boldsymbol{A}$的**谱半径**.

由此定义可知，矩阵的谱半径即该矩阵的所有特征值的模的最大值.

定理 6.1.9　设矩阵$\boldsymbol{A} \in \mathbb{C}^{n\times n}$，则$\rho(\boldsymbol{A}) \le \|\boldsymbol{A}\|$，其中，$\rho(\boldsymbol{A})$是$\boldsymbol{A}$的谱半径，$\|\boldsymbol{A}\|$是$\boldsymbol{A}$的任意一种范数.

证：　设λ是$\boldsymbol{A}$的任意一个特征值，即

$$\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}, \qquad \boldsymbol{x} \ne \boldsymbol{0},$$

故

$$\|\lambda\boldsymbol{x}\| = |\lambda|\|\boldsymbol{x}\| = \|\boldsymbol{A}\boldsymbol{x}\| \le \|\boldsymbol{A}\|\|\boldsymbol{x}\|.$$

又因为$\boldsymbol{x} \ne \boldsymbol{0}$，故$\|\boldsymbol{x}\| > 0$，于是

$$|\lambda| \le \|\boldsymbol{A}\|,$$

由于λ是$\boldsymbol{A}$的任意一个特征值，故

$$\rho(\boldsymbol{A}) \le \|\boldsymbol{A}\|. \qquad \blacksquare$$

6.2 矩阵序列与极限

6.2.1 矩阵序列

定义 6.2.1　设$\{\boldsymbol{A}_k\}$是矩阵序列，其中$\boldsymbol{A}_k=(a_{ij}^{(k)})\in\mathbb{C}^{m\times n}$，$k=1,2,\cdots$，若$m\times n$个数列$\{a_{ij}^{(k)}\}(i=1,2,\cdots,m;j=1,2,\cdots,n)$都收敛，则称**矩阵序列$\{\boldsymbol{A}_k\}$收敛**. 若$\lim\limits_{k\to\infty}a_{ij}^{(k)}=a_{ij}$，则$\lim\limits_{k\to\infty}\boldsymbol{A}_k=\boldsymbol{A}=(a_{ij})_{m\times n}$，称$\boldsymbol{A}$为**矩阵序列$\{\boldsymbol{A}_k\}$的极限**.

若将向量看成矩阵的特例，类似可得向量序列收敛的定义.

定理 6.2.1　矩阵序列$\{\boldsymbol{A}_k\}$收敛于$\boldsymbol{A}$的充要条件是$\lim\limits_{k\to\infty}\|\boldsymbol{A}_k-\boldsymbol{A}\|=0$，其中$\|\boldsymbol{A}_k-\boldsymbol{A}\|$为任何一种矩阵范数.

证：　先取矩阵范数$\|\boldsymbol{A}\|=\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|$.

①必要性　设$\lim\limits_{k\to\infty}\boldsymbol{A}_k=\boldsymbol{A}=(a_{ij})$，由定义知，对于每一个$i,j$都有

$$\lim_{k\to\infty}|a_{ij}^{(k)}-a_{ij}|=0\quad(i=1,2,\cdots,m;j=1,2,\cdots,n),$$

于是

$$\lim_{k\to\infty}\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}^{(k)}-a_{ij}|=0,$$

此即

$$\lim_{k\to\infty}\|\boldsymbol{A}_k-\boldsymbol{A}\|=0.$$

②充分性　$\lim\limits_{k\to\infty}\|\boldsymbol{A}_k-\boldsymbol{A}\|=\lim\limits_{k\to\infty}\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}^{(k)}-a_{ij}|=0$，因此，对于每一个$i,j$都有

$$\lim_{k\to\infty}|a_{ij}^{(k)}-a_{ij}|=0,$$

此即

$$\lim_{k\to\infty}a_{ij}^{(k)}=a_{ij},$$

于是

$$\lim_{k\to\infty}\boldsymbol{A}_k=\boldsymbol{A}.$$

以上证明了对于所取的矩阵范数，定理成立. 若$\|\boldsymbol{A}\|_\alpha$是其他矩阵范数，则由范数等价性定理知

$$c_1\|\boldsymbol{A}_k-\boldsymbol{A}\|\le\|\boldsymbol{A}_k-\boldsymbol{A}\|_\alpha\le c_2\|\boldsymbol{A}_k-\boldsymbol{A}\|$$

由$\lim\limits_{k\to\infty}\|\boldsymbol{A}_k-\boldsymbol{A}\|=0$可得

$$\lim_{k\to\infty}\|\boldsymbol{A}_k-\boldsymbol{A}\|_\alpha=0.$$

因此，对任意一种矩阵范数定理都成立. ■

6.2.2 矩阵序列收敛的性质

定理 6.2.2 收敛的矩阵序列具有以下性质:

(1) 一个收敛矩阵序列的极限是唯一的.

(2) 设$\lim\limits_{k\to\infty}\boldsymbol{A}_k=\boldsymbol{A}$，$\lim\limits_{k\to\infty}\boldsymbol{B}_k=\boldsymbol{B}$，其中$\boldsymbol{A},\boldsymbol{B}\in\mathbb{C}^{m\times n}$，则

$$\lim_{k\to\infty}(a\boldsymbol{A}_k+b\boldsymbol{B}_k)=a\boldsymbol{A}+b\boldsymbol{B},\quad a,b\in\mathbb{C}.$$

(3) 设$\lim\limits_{k\to\infty}\boldsymbol{A}_k=\boldsymbol{A}$，$\lim\limits_{k\to\infty}\boldsymbol{B}_k=\boldsymbol{B}$，其中$\boldsymbol{A},\boldsymbol{B}\in\mathbb{C}^{n\times n}$，则

$$\lim_{k\to\infty}\boldsymbol{A}_k\boldsymbol{B}_k=\boldsymbol{A}\boldsymbol{B}.$$

(4) 设$\lim\limits_{k\to\infty}\boldsymbol{A}_k=\boldsymbol{A}$，其中$\boldsymbol{A}\in\mathbb{C}^{n\times n}$，取$\boldsymbol{P},\boldsymbol{Q}\in\mathbb{C}^{n\times n}$，则

$$\lim_{k\to\infty}\boldsymbol{P}\boldsymbol{A}_k\boldsymbol{Q}=\boldsymbol{P}\boldsymbol{A}\boldsymbol{Q}.$$

(5) 设$\lim\limits_{k\to\infty}\boldsymbol{A}_k=\boldsymbol{A}$，且$\boldsymbol{A}_k$，$\boldsymbol{A}$均可逆，则$\{\boldsymbol{A}_k^{-1}\}$也收敛，且$\lim\limits_{k\to\infty}\boldsymbol{A}_k^{-1}=\boldsymbol{A}^{-1}$.

证: (5) 设$\boldsymbol{A}_k^*$为$\boldsymbol{A}_k$的伴随矩阵，则

$$\boldsymbol{A}_k^{-1}=\frac{\boldsymbol{A}_k^*}{|\boldsymbol{A}_k|}.$$

$\boldsymbol{A}_k^*$的元素是$\boldsymbol{A}_k$元素的代数余子式，也是$\boldsymbol{A}_k$元素的$n-1$次多项式，因此有

$$\lim_{k\to\infty}\boldsymbol{A}_k^*=\boldsymbol{A}^*,\quad \lim_{k\to\infty}|\boldsymbol{A}_k|=|\boldsymbol{A}|\neq 0.$$

于是

$$\lim_{k\to\infty}\boldsymbol{A}_k^{-1}=\lim_{k\to\infty}\frac{\boldsymbol{A}_k^*}{|\boldsymbol{A}_k|}=\frac{\boldsymbol{A}^*}{|\boldsymbol{A}|}=\boldsymbol{A}^{-1}.$$ ■

6.2.3 矩阵序列的敛散性

对于由n阶方阵$\boldsymbol{A}$的幂组成的矩阵序列$\boldsymbol{A},\boldsymbol{A}^2,\boldsymbol{A}^3,\cdots,\boldsymbol{A}^k,\cdots$有如下定理.

定理 6.2.3 若矩阵$\boldsymbol{A}$的某一种范数$\|\boldsymbol{A}\|<1$，则$\lim\limits_{k\to\infty}\boldsymbol{A}^k=\boldsymbol{0}$.

证: 由矩阵范数的相容性，有$\|\boldsymbol{A}^k\|\le\|\boldsymbol{A}\|^k$，又$\|\boldsymbol{A}\|<1$，可得

$$\lim_{k\to\infty}\boldsymbol{A}^k=\boldsymbol{0}.$$ ■

定理 6.2.4 给定矩阵序列$\boldsymbol{A},\boldsymbol{A}^2,\cdots,\boldsymbol{A}^k,\cdots$，则$\lim\limits_{k\to\infty}\boldsymbol{A}^k=\boldsymbol{0}$的充要条件是$\rho(\boldsymbol{A})<1$.

证：设$\boldsymbol{A}$的 Jordan 标准形是

$$\boldsymbol{J}=\begin{bmatrix}\boldsymbol{J}_1(\lambda_1) & & & \\ & \boldsymbol{J}_2(\lambda_2) & & \\ & & \ddots & \\ & & & \boldsymbol{J}_s(\lambda_s)\end{bmatrix},$$

其中, Jordan 块

$$\boldsymbol{J}_i(\lambda_i)=\begin{bmatrix}\lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i\end{bmatrix}_{n_i\times n_i}\quad (i=1,2,\cdots,s),$$

于是有

$$\boldsymbol{A}^k=\boldsymbol{P}\mathrm{diag}(\boldsymbol{J}_1^k(\lambda_1),\boldsymbol{J}_2^k(\lambda_2),\cdots,\boldsymbol{J}_s^k(\lambda_s))\boldsymbol{P}^{-1}.$$

显然，$\lim\limits_{k\to\infty}\boldsymbol{A}^k=\boldsymbol{0}$的充要条件是$\lim\limits_{k\to\infty}\boldsymbol{J}_i^k(\lambda_i)=\boldsymbol{0}$，$i=1,2,\cdots,s$.

又因为

$$\boldsymbol{J}_i^k(\lambda_i)=\begin{bmatrix}\lambda_i^k & \mathrm{C}_k^1\lambda_i^{k-1} & & \cdots & \mathrm{C}_k^{n_i-1}\lambda_i^{k-n_i+1} \\ & \lambda_i^k & \mathrm{C}_k^1\lambda_i^{k-1} & & \\ & & \ddots & & \vdots \\ & & & \ddots & \\ & & & & \mathrm{C}_k^1\lambda_i^{k-1} \\ & & & & \lambda_i^k\end{bmatrix},$$

其中，$\mathrm{C}_k^l=\frac{k(k-1)\cdots(k-l+1)}{l!}$，当$k\geq l$；$\mathrm{C}_k^l=0$，当$k<l$. 于是，$\lim\limits_{k\to\infty}\boldsymbol{J}_i^k(\lambda_i)=\boldsymbol{0}$的充要条件是$|\lambda_i|<1$.

因此，$\lim\limits_{k\to\infty}\boldsymbol{A}^k=\boldsymbol{0}$的充要条件是$\rho(\boldsymbol{A})<1$. ■

例 6.2.1 判断如下矩阵序列$\boldsymbol{A}^k$的敛散性.

(1) $\boldsymbol{A}=\begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}$　　(2) $\boldsymbol{A}=\begin{bmatrix}0.9 & 1\\0 & 0.9\end{bmatrix}$

(3) $\boldsymbol{A}=\begin{bmatrix}1 & 0 & 0\\0 & 0.9 & 1\\0 & 0 & 0.9\end{bmatrix}$　　(4) $\boldsymbol{A}=\begin{bmatrix}0.3 & 0.8\\0.6 & 0.1\end{bmatrix}$

解：(1) $\boldsymbol{A}^k=\begin{bmatrix}1 & k\\0 & 1\end{bmatrix}$，故$\lim\limits_{k\to\infty}\boldsymbol{A}^k$发散.

(2) $\boldsymbol{A}$的特征值$\lambda_1=\lambda_2=0.9<1$，由定理 6.2.4 可知$\boldsymbol{A}^k$收敛，且$\lim\limits_{k\to\infty}\boldsymbol{A}^k=\boldsymbol{0}$.

(3) $\boldsymbol{A}^k=\begin{bmatrix}1 & 0 & 0\\ 0 & 0.9^k & k0.9^{k-1}\\ 0 & 0 & 0.9^k\end{bmatrix}$，由于$\lim\limits_{k\to\infty}0.9^k=0$，$\lim\limits_{k\to\infty}k0.9^{k-1}=0$．故$\boldsymbol{A}^k$收敛，且

$$\lim_{k\to\infty}\boldsymbol{A}^k=\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}.$$

(4) 由于$\|\boldsymbol{A}\|_1=0.9<1$，由定理 6.2.3 可知$\boldsymbol{A}^k$收敛，且$\lim\limits_{k\to\infty}\boldsymbol{A}^k=\boldsymbol{0}$.

例 6.2.2 判断矩阵序列$\boldsymbol{A}^k$的敛散性.

(1) $\boldsymbol{A}=\begin{bmatrix}1 & 1 & 0\\ 0 & 1 & 0\\ 0 & -1 & 1\end{bmatrix}$　　(2) $\boldsymbol{A}=\begin{bmatrix}\frac{2}{3} & 0 & -\frac{1}{3}\\ 0 & 1 & 0\\ -\frac{1}{6} & 0 & \frac{5}{6}\end{bmatrix}$

分析：本题并未要求矩阵序列收敛至零矩阵.

解：(1) $\boldsymbol{A}$的特征值$\lambda_1=\lambda_2=\lambda_3=1$，$\rho(\boldsymbol{A})=1$，若$\boldsymbol{J}$是$\boldsymbol{A}$的 Jordan 标准形，由于$\boldsymbol{A}^k=\boldsymbol{P}\boldsymbol{J}^k\boldsymbol{P}^{-1}$，所以只需判断矩阵序列$\boldsymbol{J}^k$的敛散性.

$$\lambda\boldsymbol{E}-\boldsymbol{A}=\begin{bmatrix}\lambda-1 & -1 & 0\\ 0 & \lambda-1 & 0\\ 0 & 1 & \lambda-1\end{bmatrix}\to\begin{bmatrix}1 & & \\ & \lambda-1 & \\ & & (\lambda-1)^2\end{bmatrix}.$$

$\boldsymbol{A}$的不变因子为$1,\lambda-1,(\lambda-1)^2$，$\boldsymbol{A}$的初等因子为$\lambda-1,(\lambda-1)^2$，所以

$$\boldsymbol{J}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 1\\ 0 & 0 & 1\end{bmatrix},\quad \boldsymbol{J}^k=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & k\\ 0 & 0 & 1\end{bmatrix}\to\infty,$$

所以$\boldsymbol{A}^k$发散.

(2) $\boldsymbol{A}$的特征值$\lambda_1=\lambda_2=1,\lambda_3=\frac{1}{2}$，$\rho(\boldsymbol{A})=1$，需要判断$\boldsymbol{J}^k$的敛散性.

$$\lambda\boldsymbol{E}-\boldsymbol{A}=\begin{bmatrix}\lambda-\frac{2}{3} & 0 & \frac{1}{3}\\ 0 & \lambda-1 & 0\\ \frac{1}{6} & 0 & \lambda-\frac{5}{6}\end{bmatrix}\to\begin{bmatrix}1 & & \\ & \lambda-1 & \\ & & (\lambda-1)(\lambda-\frac{1}{2})\end{bmatrix}.$$

$\boldsymbol{A}$的不变因子为$1,\lambda-1,(\lambda-1)(\lambda-\frac{1}{2})$，$\boldsymbol{A}$的初等因子为$(\lambda-1),(\lambda-1),(\lambda-\frac{1}{2})$，所以

$$\boldsymbol{J}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & \frac{1}{2}\end{bmatrix},\quad \boldsymbol{J}^k=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & (\frac{1}{2})^k\end{bmatrix}\to\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{bmatrix},$$

所以$\boldsymbol{A}^k$收敛.

6.3 矩阵级数与矩阵函数

6.3.1 矩阵级数

定义 6.3.1 设$\boldsymbol{A}_k=\left(a_{ij}^{(k)}\right)_{m\times n}\in\mathbb{C}^{m\times n}$，若$m\times n$个常数项级数

$$\sum_{k=0}^{+\infty}a_{ij}^{(k)}=a_{ij}^{(0)}+a_{ij}^{(1)}+\cdots+a_{ij}^{(k)}+\cdots$$

$$(i=1,2,\cdots,m;\ j=1,2,\cdots,n) \tag{6.3.1}$$

都收敛，则称矩阵级数

$$\textstyle\sum_{k=0}^{+\infty}\boldsymbol{A}_k=\boldsymbol{A}_0+\boldsymbol{A}_1+\cdots+\boldsymbol{A}_k+\cdots \tag{6.3.2}$$

收敛.若常数项级数(6.3.1)的和为a_{ij}，则矩阵级数(6.3.2)的和为$\boldsymbol{A}=\left(a_{ij}\right)_{m\times n}$.

不收敛的矩阵级数称为**发散**的.

定义 6.3.2 若$m\times n$个常数项级数(6.3.1)都绝对收敛，则称矩阵级数(6.3.2)**绝对收敛**.

例 6.3.1 已知

$$\boldsymbol{A}_k=\begin{bmatrix}\frac{1}{2^k}&\frac{\pi}{4^k}\\0&\frac{1}{(k+1)(k+2)}\end{bmatrix},\quad(k=0,1,\cdots)$$

讨论矩阵级数$\sum_{k=0}^{+\infty}\boldsymbol{A}_k$的敛散性.

解：因为

$$\boldsymbol{S}_N=\sum_{k=0}^{N}\boldsymbol{A}_k=\begin{bmatrix}\sum\limits_{k=0}^{N}\frac{1}{2^k}&\sum\limits_{k=0}^{N}\frac{\pi}{4^k}\\0&\sum\limits_{k=0}^{N}\frac{1}{(k+1)(k+2)}\end{bmatrix}$$

$$=\begin{bmatrix}2-\frac{1}{2^N}&\frac{\pi}{3}(4-\frac{1}{4^N})\\0&1-\frac{1}{N+2}\end{bmatrix},$$

因此所给矩阵级数收敛，其和为

$$\boldsymbol{S}=\lim_{N\to+\infty}\boldsymbol{S}_N=\begin{bmatrix}2&\frac{4}{3}\pi\\0&1\end{bmatrix}.$$

利用矩阵范数，可以将判断矩阵级数是否绝对收敛的问题转化为判断一个

正项级数是否收敛的问题.

定理 6.3.1 设矩阵序列$\boldsymbol{A}_k=\left(a_{ij}^{(k)}\right)_{m\times n}\in\mathbb{C}^{m\times n}$，则矩阵级数$\sum_{k=0}^{+\infty}\boldsymbol{A}_k$绝对收敛的充要条件是正项级数$\sum_{k=0}^{+\infty}\|\boldsymbol{A}_k\|$收敛，其中$\|\boldsymbol{A}\|$为$\mathbb{C}^{m\times n}$上的任一矩阵范数.

证：①充分性 取矩阵范数$\|\boldsymbol{A}_k\|=\sum_{i=1}^{m}\sum_{j=1}^{n}\left|a_{ij}^{(k)}\right|$，对于每一个$i,j$都有

$$\|\boldsymbol{A}_k\|\geq\left|a_{ij}^{(k)}\right|,$$

因此，若$\sum_{k=0}^{+\infty}\|\boldsymbol{A}_k\|$收敛，则对于每一个$i,j$，常数项级数$\sum_{k=0}^{+\infty}\left|a_{ij}^{(k)}\right|$都收敛，于是$\sum_{k=0}^{+\infty}\boldsymbol{A}_k$绝对收敛.

②必要性 若矩阵级数$\sum_{k=0}^{+\infty}\boldsymbol{A}_k$绝对收敛，则对每一个$i,j$都有$\sum_{k=0}^{+\infty}\left|a_{ij}^{(k)}\right|<+\infty$，于是

$$\sum_{k=0}^{+\infty}\|\boldsymbol{A}_k\|=\sum_{k=0}^{+\infty}\left(\sum_{i=1}^{m}\sum_{j=1}^{n}\left|a_{ij}^{(k)}\right|\right)=\sum_{i=1}^{m}\sum_{j=1}^{n}\left(\sum_{k=0}^{+\infty}\left|a_{ij}^{(k)}\right|\right)<+\infty$$

即正项级数$\sum_{k=0}^{+\infty}\|\boldsymbol{A}_k\|$收敛.

最后，根据范数等价性定理可知此结论对任一矩阵范数都正确. ■

6.3.2 矩阵幂级数

矩阵幂级数是一类特殊的矩阵级数，它是研究矩阵函数的重要工具.

定义 6.3.3 设矩阵$\boldsymbol{A}\in\mathbb{C}^{n\times n}$，$c_k\in\mathbb{C}$，称形式为

$$\sum_{k=0}^{+\infty}c_k\boldsymbol{A}^k=c_0\boldsymbol{E}+c_1\boldsymbol{A}+c_2\boldsymbol{A}^2+\cdots+c_k\boldsymbol{A}^k+\cdots$$

的矩阵级数为矩阵$\boldsymbol{A}$的**幂级数**.

利用定义判断矩阵幂级数的敛散性需要判断n^2个常数项级数的敛散性，当矩阵阶数 n 较大时很不方便. 由于矩阵幂级数是复变量 z 的幂级数$\sum_{k=0}^{+\infty}c_kz^k$的推广，如果幂级数$\sum_{k=0}^{+\infty}c_kz^k$的收敛半径是 R，则对收敛圆$|z|<R$内的所有 z，$\sum_{k=0}^{+\infty}c_kz^k$都是绝对收敛的，因此，矩阵幂级数$\sum_{k=0}^{+\infty}c_k\boldsymbol{A}^k$的敛散性与复变量 z 的幂级数$\sum_{k=0}^{+\infty}c_kz^k$的收敛半径存在联系，这就是下述定理.

定理 6.3.2 设幂级数$\sum_{k=0}^{+\infty}c_kx^k$的收敛半径为 R，矩阵$\boldsymbol{A}\in\mathbb{C}^{n\times n}$，若$\rho(\boldsymbol{A})<R$，则矩阵幂级数$\sum_{k=0}^{+\infty}c_k\boldsymbol{A}^k$绝对收敛；若$\rho(\boldsymbol{A})>R$，则$\sum_{k=0}^{+\infty}c_k\boldsymbol{A}^k$发散.

例 6. 3. 2　由于幂级数

$$1+x+\frac{1}{2!}x^2+\cdots+\frac{1}{k!}x^k+\cdots$$

$$1-\frac{1}{2!}x^2+\frac{1}{4!}x^4-\cdots+(-1)^k\frac{1}{(2k)!}x^{2k}+\cdots$$

$$x-\frac{1}{3!}x^3+\frac{1}{5!}x^5-\cdots+(-1)^k\frac{1}{(2k+1)!}x^{2k+1}+\cdots$$

的收敛半径$R=\infty$，所以对于任意 n 阶矩阵$\boldsymbol{A}$，矩阵幂级数

$$\boldsymbol{E}+\boldsymbol{A}+\frac{1}{2!}\boldsymbol{A}^2+\cdots+\frac{1}{k!}\boldsymbol{A}^k+\cdots$$

$$\boldsymbol{E}-\frac{1}{2!}\boldsymbol{A}^2+\frac{1}{4!}\boldsymbol{A}^4-\cdots+(-1)^k\frac{1}{(2k)!}\boldsymbol{A}^{2k}+\cdots$$

$$\boldsymbol{A}-\frac{1}{3!}\boldsymbol{A}^3+\frac{1}{5!}\boldsymbol{A}^5-\cdots+(-1)^k\frac{1}{(2k+1)!}\boldsymbol{A}^{2k+1}+\cdots$$

都绝对收敛.

由定理 6.3.2 可得下述定理.

定理 6. 3. 3　设幂级数$\sum_{k=0}^{+\infty}c_kx^k$的收敛半径为 R，矩阵$\boldsymbol{A}\in\mathbb{C}^{n\times n}$，若$\boldsymbol{A}$的某一范数$\|\boldsymbol{A}\|$在幂级数$\sum_{k=0}^{+\infty}c_kx^k$的收敛域内，即$\|\boldsymbol{A}\|<R$，则矩阵幂级数$\sum_{k=0}^{+\infty}c_k\boldsymbol{A}^k$绝对收敛.

例 6. 3. 3　若$\boldsymbol{A}=\begin{bmatrix}0.2 & 0.5 & 0.1\\ 0.1 & 0.5 & 0.3\\ 0.2 & 0.4 & 0.2\end{bmatrix}$，试证明$\boldsymbol{E}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots+\boldsymbol{A}^k+\cdots$ 绝对收敛.

证：因为级数$1+x+x^2+\cdots+x^k+\cdots$的收敛半径为 1，而$\|\boldsymbol{A}\|_\infty=0.9<1$，故矩阵幂级数$\boldsymbol{E}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots+\boldsymbol{A}^k+\cdots$ 绝对收敛.

最后，介绍一个特殊的矩阵幂级数——Neumann 级数.

定理 6. 3. 4　设$\boldsymbol{A}\in\mathbb{C}^{n\times n}$，矩阵幂级数

$$\sum_{k=0}^{+\infty}\boldsymbol{A}^k=\boldsymbol{E}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots+\boldsymbol{A}^k+\cdots\qquad\text{（Neumann 级数）}$$

绝对收敛的充要条件是$\rho(\boldsymbol{A})<1$，且其和是$(\boldsymbol{E}-\boldsymbol{A})^{-1}$.

证：①充分性　由于幂级数$\sum_{k=0}^{+\infty}x^k=1+x+x^2+\cdots+x^k+\cdots$的收敛半径$R=1$．故由定理 6.3.2，当$\rho(\boldsymbol{A})<1$时，$\sum_{k=0}^{+\infty}\boldsymbol{A}^k=\boldsymbol{E}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots+\boldsymbol{A}^k+\cdots$绝对收敛.

②必要性　若矩阵幂级数$\sum_{k=0}^{+\infty}\boldsymbol{A}^k$绝对收敛，由定理 6.3.1，则正项级数$\|\boldsymbol{E}\|+\|\boldsymbol{A}\|+\|\boldsymbol{A}^2\|+\cdots+\|\boldsymbol{A}^k\|+\cdots$ 收敛，故$\|\boldsymbol{A}^k\|\to 0$，$\boldsymbol{A}^k\to\boldsymbol{0}$，再由定理6.2.4,

可得$\rho(\boldsymbol{A}) < 1$.

再求其和. 因为

$$(\boldsymbol{E}-\boldsymbol{A})(\boldsymbol{E}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots+\boldsymbol{A}^k+\cdots)=\boldsymbol{E},$$

故

$$\boldsymbol{E}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots+\boldsymbol{A}^k+\cdots=(\boldsymbol{E}-\boldsymbol{A})^{-1}. \quad \blacksquare$$

例 6. 3. 4　设$\boldsymbol{A}=\begin{bmatrix}0.2 & 0.5 & 0.1\\ 0.1 & 0.5 & 0.3\\ 0.2 & 0.4 & 0.2\end{bmatrix}$, 求$\boldsymbol{E}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots+\boldsymbol{A}^k+\cdots$ 的和.

解:　因$\|\boldsymbol{A}\|_\infty = 0.9 < 1$, 故$\rho(\boldsymbol{A}) < 1$. 由定理 6.3.4 知所求矩阵幂级数绝对收敛, 且其和是$(\boldsymbol{E}-\boldsymbol{A})^{-1}$. 因此

$$\boldsymbol{E}-\boldsymbol{A}=\begin{bmatrix}0.8 & -0.5 & -0.1\\ -0.1 & 0.5 & -0.3\\ -0.2 & -0.4 & 0.8\end{bmatrix},$$

于是

$$\boldsymbol{E}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots+\boldsymbol{A}^k+\cdots=(\boldsymbol{E}-\boldsymbol{A})^{-1}=\begin{bmatrix}2 & \frac{22}{7} & \frac{10}{7}\\ 1 & \frac{31}{7} & \frac{25}{14}\\ 1 & 3 & \frac{5}{2}\end{bmatrix}.$$

6. 3. 3　矩阵函数的幂级数定义

矩阵函数与通常的函数类似, 不同之处在于矩阵函数的自变量和因变量都是 n 阶方阵. 矩阵函数中最简单的是矩阵多项式, 矩阵多项式是研究其他矩阵函数的基础. 矩阵函数一般用幂级数表示.

由高等数学的相关知识及定理 6.3.2, 可利用矩阵幂级数来定义矩阵函数.

定义 6. 3. 4　设幂级数$\sum_{k=0}^{+\infty} c_k z^k$的收敛半径是$R$, 且当$|z| < R$时, 幂级数收敛于$f(z)$, 即

$$f(z)=\sum_{k=0}^{+\infty} c_k z^k,\quad 当|z| < R,$$

如果矩阵$\boldsymbol{A}\in\mathbb{C}^{n\times n}$的谱半径$\rho(\boldsymbol{A}) < R$, 则称收敛矩阵幂级数$\sum_{k=0}^{+\infty} c_k \boldsymbol{A}^k$的和为**矩阵函数**, 记为$f(\boldsymbol{A})$, 即$f(\boldsymbol{A})=\sum_{k=0}^{+\infty} c_k \boldsymbol{A}^k$.

定义 6.3.4 称为矩阵函数的幂级数定义或幂级数表示. 根据这个定义, 可得到形式上和高等数学中的一些函数类似的矩阵函数.

如下的幂级数在各自的收敛域内均收敛:

$$e^z = 1 + z + \frac{1}{2!}z^2 + \cdots + \frac{1}{k!}z^k + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!}z^k \qquad (|z| < +\infty)$$

$$\sin z = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} z^{2k+1} \qquad (|z| < +\infty)$$

$$\cos z = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} z^{2k} \qquad (|z| < +\infty)$$

$$\frac{1}{1-z} = \sum_{k=0}^{\infty} z^k \qquad (|z| < 1)$$

$$\ln(1+z) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} z^{k+1} \qquad (|z| < 1)$$

因此, 对于 n 阶方阵$\boldsymbol{A}$, 有

$$e^{\boldsymbol{A}} = \sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{A}^k \qquad (\forall \boldsymbol{A} \in \mathbb{C}^{n\times n})$$

$$\sin\boldsymbol{A} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} \boldsymbol{A}^{2k+1} \qquad (\forall \boldsymbol{A} \in \mathbb{C}^{n\times n})$$

$$\cos\boldsymbol{A} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} \boldsymbol{A}^{2k} \qquad (\forall \boldsymbol{A} \in \mathbb{C}^{n\times n})$$

$$\frac{1}{\boldsymbol{E}-\boldsymbol{A}} = \sum_{k=0}^{\infty} \boldsymbol{A}^k \qquad (\rho(\boldsymbol{A}) < 1)$$

$$\ln(\boldsymbol{E}+\boldsymbol{A}) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} \boldsymbol{A}^{k+1} \qquad (\rho(\boldsymbol{A}) < 1)$$

对于 n 阶方阵$\boldsymbol{A}$, 称$e^{\boldsymbol{A}}$为**矩阵指数函数**, 称$\sin\boldsymbol{A}$为**矩阵正弦函数**, 称$\cos\boldsymbol{A}$为**矩阵余弦函数**.

定理 6. 3. 5　对于任意方阵$\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{n\times n}$, $k, l \in \mathbb{C}$, 有

(1) $e^{k\boldsymbol{A}}e^{l\boldsymbol{A}} = e^{(k+l)\boldsymbol{A}}$

(2) $(e^{\boldsymbol{A}})^{-1} = e^{-\boldsymbol{A}}$

(3) 当$\boldsymbol{AB} = \boldsymbol{BA}$时, $e^{\boldsymbol{A}}e^{\boldsymbol{B}} = e^{\boldsymbol{B}}e^{\boldsymbol{A}} = e^{\boldsymbol{A}+\boldsymbol{B}}$

(4) $\frac{d}{dt}(e^{\boldsymbol{A}t}) = \boldsymbol{A}e^{\boldsymbol{A}t} = e^{\boldsymbol{A}t}\boldsymbol{A}$

(5) $\frac{d}{dt}(\sin(\boldsymbol{At})) = \boldsymbol{A}\cos(\boldsymbol{At}) = \cos(\boldsymbol{At}) \cdot \boldsymbol{A}$

(6) $\frac{d}{dt}(\cos(\boldsymbol{At})) = -\boldsymbol{A}\sin(\boldsymbol{At}) = -\sin(\boldsymbol{At}) \cdot \boldsymbol{A}$

6.4　函数矩阵的微分与积分

6.4.1　函数矩阵的定义及运算

定义 6.4.1　以实变量x的实函数$a_{ij}(x)$为元素的矩阵

$$\boldsymbol{A}(x) = \begin{bmatrix} a_{11}(x) & a_{12}(x) & \cdots & a_{1n}(x) \\ a_{21}(x) & a_{22}(x) & \cdots & a_{2n}(x) \\ \vdots & \vdots & & \vdots \\ a_{m1}(x) & a_{m2}(x) & \cdots & a_{mn}(x) \end{bmatrix}$$

称为**函数矩阵**，其中所有元素$a_{ij}(x)$ $(i = 1,2,\cdots,m; j = 1,2,\cdots,n)$都是定义在区间$[a,b]$上的实函数. 当$m = 1$时，$\boldsymbol{A}(x)$是**函数行向量**，当$n = 1$时，$\boldsymbol{A}(x)$是**函数列向量**.

定义 6.4.2　函数矩阵的运算定义如下:

①**加法**　设$\boldsymbol{A}(x) = \left(a_{ij}(x)\right)_{m\times n}$，$\boldsymbol{B}(x) = \left(b_{ij}(x)\right)_{m\times n}$，它们的和$\boldsymbol{A}(x) + \boldsymbol{B}(x)$定义为

$$\boldsymbol{A}(x) + \boldsymbol{B}(x) = \left(a_{ij}(x) + b_{ij}(x)\right)_{m\times n}.$$

②**数乘**　设$k(x)$为x的函数，$\boldsymbol{A}(x) = \left(a_{ij}(x)\right)_{m\times n}$，$k(x)$与$\boldsymbol{A}(x)$的数乘定义为

$$k(x)\boldsymbol{A}(x) = \left(k(x)a_{ij}(x)\right)_{m\times n}.$$

③**矩阵乘法**　设$\boldsymbol{A}(x) = \left(a_{ij}(x)\right)_{m\times k}$，$\boldsymbol{B}(x) = \left(b_{ij}(x)\right)_{k\times n}$，它们的积$\boldsymbol{A}(x)\boldsymbol{B}(x)$定义为

$$\boldsymbol{A}(x)\boldsymbol{B}(x) = \left(c_{ij}(x)\right)_{m\times n}.$$

其中，$c_{ij}(x) = \sum_{l=1}^{k} a_{il}(x)b_{lj}(x)$ $(i = 1,2,\cdots m; j = 1,2,\cdots,n)$，称$\boldsymbol{A}(x)$与$\boldsymbol{B}(x)$是可乘的.

④**转置**　设$\boldsymbol{A}(x)=\begin{bmatrix} a_{11}(x) & a_{12}(x) & \cdots & a_{1n}(x) \\ a_{21}(x) & a_{22}(x) & \cdots & a_{2n}(x) \\ \vdots & \vdots & & \vdots \\ a_{m1}(x) & a_{m2}(x) & \cdots & a_{mn}(x) \end{bmatrix}_{m\times n}$，它的转置矩阵$\boldsymbol{A}^{\mathrm{T}}(x)$定义为

$$\boldsymbol{A}^{\mathrm{T}}(x)=\begin{bmatrix} a_{11}(x) & a_{21}(x) & \cdots & a_{m1}(x) \\ a_{12}(x) & a_{22}(x) & \cdots & a_{m2}(x) \\ \vdots & \vdots & & \vdots \\ a_{1n}(x) & a_{2n}(x) & \cdots & a_{mn}(x) \end{bmatrix}_{n\times m}.$$

由定义 6.4.2 可知，函数矩阵的运算及运算性质与常数矩阵相同.

定义 6.4.3　设$\boldsymbol{A}(x)=\left(a_{ij}(x)\right)$为$n$阶函数矩阵，若存在$n$阶函数矩阵$\boldsymbol{B}(x)=\left(b_{ij}(x)\right)$使得对于任何$x\in[a,b]$都有$\boldsymbol{A}(x)\boldsymbol{B}(x)=\boldsymbol{B}(x)\boldsymbol{A}(x)=\boldsymbol{E}$，则称$\boldsymbol{A}(x)$在$[a,b]$上可逆，$\boldsymbol{B}(x)$是$\boldsymbol{A}(x)$的**逆矩阵**，记为$\boldsymbol{A}^{-1}(x)$.

定理 6.4.1　n阶矩阵$\boldsymbol{A}(x)$在$[a,b]$上可逆的充要条件是$|\boldsymbol{A}(x)|$在$[a,b]$上处处不为零，且$\boldsymbol{A}^{-1}(x)=\frac{\boldsymbol{A}^*(x)}{|\boldsymbol{A}(x)|}$，其中$\boldsymbol{A}^*(x)$是$\boldsymbol{A}(x)$的伴随矩阵，即

$$\boldsymbol{A}^*(x)=\begin{bmatrix} \boldsymbol{A}_{11}(x) & \boldsymbol{A}_{21}(x) & \cdots & \boldsymbol{A}_{n1}(x) \\ \boldsymbol{A}_{12}(x) & \boldsymbol{A}_{22}(x) & \cdots & \boldsymbol{A}_{n2}(x) \\ \vdots & \vdots & & \vdots \\ \boldsymbol{A}_{1n}(x) & \boldsymbol{A}_{2n}(x) & \cdots & \boldsymbol{A}_{nn}(x) \end{bmatrix},$$

式中的$\boldsymbol{A}_{ij}(x)$是$\boldsymbol{A}(x)$中元素$a_{ij}(x)$的代数余子式.

例如，函数矩阵$\boldsymbol{A}(x)=\begin{bmatrix} x & 1 \\ 1 & x \end{bmatrix}$在[2,3]上的逆矩阵为$\boldsymbol{A}^{-1}(x)=\frac{1}{x^2-1}\begin{bmatrix} x & -1 \\ -1 & x \end{bmatrix}$，在[0,2]上$\boldsymbol{A}(x)$不可逆，这是因为在$x=1$时，$|\boldsymbol{A}(x)|=0$.

6.4.2　函数矩阵的极限

定义 6.4.4　若$\boldsymbol{A}(x)=(a_{ij}(x))_{m\times n}$的所有元素$a_{ij}(x)$在$x=x_0$处有极限，即

$$\lim_{x\to x_0} a_{ij}(x)=a_{ij}\quad (i=1,2,\dots,m;\ j=1,2,\dots,n),$$

其中a_{ij}为固定常数，则称$\boldsymbol{A}(x)$在$x=x_0$处有极限，并记为$\lim\limits_{x\to x_0}\boldsymbol{A}(x)=\boldsymbol{A}$，其中

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

若$\boldsymbol{A}(x)$的所有元素$a_{ij}(x)$在$x=x_0$处连续，即

$$\lim_{x\to x_0} a_{ij}(x) = a_{ij}(x_0) \qquad (i = 1,2,\dots,m; j = 1,2,\dots,n),$$

则称$\boldsymbol{A}(x)$在$x = x_0$处连续，并记为$\lim\limits_{x\to x_0} \boldsymbol{A}(x) = \boldsymbol{A}(x_0)$，其中

$$\boldsymbol{A}(x_0) = \begin{bmatrix} a_{11}(x_0) & a_{12}(x_0) & \cdots & a_{1n}(x_0) \\ a_{21}(x_0) & a_{22}(x_0) & \cdots & a_{2n}(x_0) \\ \vdots & \vdots & & \vdots \\ a_{m1}(x_0) & a_{m2}(x_0) & \cdots & a_{mn}(x_0) \end{bmatrix}_{m\times n}.$$

容易验证如下性质:

设$\lim\limits_{x\to x_0} \boldsymbol{A}(x) = \boldsymbol{A}$，$\lim\limits_{x\to x_0} \boldsymbol{B}(x) = \boldsymbol{B}$，则

(1) $\lim\limits_{x\to x_0} (\boldsymbol{A}(x) \pm \boldsymbol{B}(x)) = \boldsymbol{A} \pm \boldsymbol{B}$;

(2) $\lim\limits_{x\to x_0} (k\boldsymbol{A}(x)) = k\boldsymbol{A}$;

(3) 当$\boldsymbol{A}(x)$与$\boldsymbol{B}(x)$可乘时，$\lim\limits_{x\to x_0} (\boldsymbol{A}(x)\boldsymbol{B}(x)) = \boldsymbol{AB}$.

6.4.3 函数矩阵的导数

定义 6.4.5 若$\boldsymbol{A}(x) = \left(a_{ij}(x)\right)_{m\times n}$的所有元素$a_{ij}(x)(i = 1,2,\cdots,m; j = 1,2,\cdots,n$）在点$x = x_0$处可导，则称函数矩阵$\boldsymbol{A}(x)$在点$x = x_0$处可导，并记为

$$\boldsymbol{A}'(x_0) = \left.\frac{\mathrm{d}\boldsymbol{A}(x)}{\mathrm{d}x}\right|_{x=x_0} = \lim_{\Delta x\to 0} \frac{\boldsymbol{A}(x_0+\Delta x) - \boldsymbol{A}(x_0)}{\Delta x}$$

$$= \begin{bmatrix} a'_{11}(x_0) & a'_{12}(x_0) & \cdots & a'_{1n}(x_0) \\ a'_{21}(x_0) & a'_{22}(x_0) & \cdots & a'_{2n}(x_0) \\ \vdots & \vdots & & \vdots \\ a'_{m1}(x_0) & a'_{m2}(x_0) & \cdots & a'_{mn}(x_0) \end{bmatrix}.$$

函数矩阵的导数运算具有如下性质:

(1) $\boldsymbol{A}(x)$是常数矩阵的充要条件是$\frac{\mathrm{d}\boldsymbol{A}(x)}{\mathrm{d}x} = \boldsymbol{0}$.

(2) 设$\boldsymbol{A}(x) = \left(a_{ij}(x)\right)_{m\times n}$，$\boldsymbol{B}(x) = \left(b_{ij}(x)\right)_{m\times n}$均可导，则

$$\frac{\mathrm{d}}{\mathrm{d}x}[\boldsymbol{A}(x) + \boldsymbol{B}(x)] = \frac{\mathrm{d}\boldsymbol{A}(x)}{\mathrm{d}x} + \frac{\mathrm{d}\boldsymbol{B}(x)}{\mathrm{d}x}.$$

(3) 设$k(x)$是x的标量函数，$\boldsymbol{A}(x)$是函数矩阵，$k(x)$与$\boldsymbol{A}(x)$均可导，则

$$\frac{\mathrm{d}}{\mathrm{d}x}[k(x)\boldsymbol{A}(x)] = \frac{\mathrm{d}k(x)}{\mathrm{d}x}\boldsymbol{A}(x) + k(x)\frac{\mathrm{d}\boldsymbol{A}(x)}{\mathrm{d}x}.$$

特别地，当$k(x)$是常数k时，有

$$\frac{\mathrm{d}}{\mathrm{d}x}[k\boldsymbol{A}(x)] = k\frac{\mathrm{d}\boldsymbol{A}(x)}{\mathrm{d}x}.$$

(4) 设$\boldsymbol{A}(x)$与$\boldsymbol{B}(x)$均可导，且$\boldsymbol{A}(x)$与$\boldsymbol{B}(x)$是可乘的，则

$$\frac{\mathrm{d}}{\mathrm{d}x}[\boldsymbol{A}(x)\boldsymbol{B}(x)] = \frac{\mathrm{d}\boldsymbol{A}(x)}{\mathrm{d}x}\boldsymbol{B}(x) + \boldsymbol{A}(x)\frac{\mathrm{d}\boldsymbol{B}(x)}{\mathrm{d}x}.$$

因为矩阵乘法没有交换律，所以$\frac{\mathrm{d}}{\mathrm{d}x}\boldsymbol{A}^2(x) \neq 2\boldsymbol{A}(x)\frac{\mathrm{d}\boldsymbol{A}(x)}{\mathrm{d}x}$，$\frac{\mathrm{d}}{\mathrm{d}x}\boldsymbol{A}^3(x) \neq 3\boldsymbol{A}^2(x)\frac{\mathrm{d}\boldsymbol{A}(x)}{\mathrm{d}x}$.

(5) 若$\boldsymbol{A}(x)$与$\boldsymbol{A}^{-1}(x)$都可导，则

$$\frac{\mathrm{d}\boldsymbol{A}^{-1}(x)}{\mathrm{d}x} = -\boldsymbol{A}^{-1}(x)\frac{\mathrm{d}\boldsymbol{A}(x)}{\mathrm{d}x}\boldsymbol{A}^{-1}(x).$$

证： 因为$\boldsymbol{A}^{-1}(x)\boldsymbol{A}(x) = \boldsymbol{E}$，所以

$$\frac{\mathrm{d}}{\mathrm{d}x}[\boldsymbol{A}^{-1}(x)\boldsymbol{A}(x)] = \frac{\mathrm{d}\boldsymbol{A}^{-1}(x)}{\mathrm{d}x}\boldsymbol{A}(x) + \boldsymbol{A}^{-1}(x)\frac{\mathrm{d}\boldsymbol{A}(x)}{\mathrm{d}x} = \frac{\mathrm{d}\boldsymbol{E}}{\mathrm{d}x} = \boldsymbol{0},$$

于是

$$\frac{\mathrm{d}\boldsymbol{A}^{-1}(x)}{\mathrm{d}x} = -\boldsymbol{A}^{-1}(x)\frac{\mathrm{d}\boldsymbol{A}(x)}{\mathrm{d}x}\boldsymbol{A}^{-1}(x).$$

(6) 设$\boldsymbol{A}(x)$为函数矩阵，$x = f(t)$是t的标量函数，$\boldsymbol{A}(x)$与$f(t)$均可导，则

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{A}(x)) = \frac{\mathrm{d}\boldsymbol{A}(x)}{\mathrm{d}x}f'(t) = f'(t)\frac{\mathrm{d}\boldsymbol{A}(x)}{\mathrm{d}x}.$$

函数矩阵的导数也是一个函数矩阵，它可以再求导，因此就有函数矩阵的高阶导数.

例 6.4.1 已知$\boldsymbol{A} = \begin{bmatrix} x & 1 \\ 1 & x \end{bmatrix}$在$[2,3]$上有逆矩阵$\boldsymbol{A}^{-1}(x) = \frac{1}{x^2-1}\begin{bmatrix} x & -1 \\ -1 & x \end{bmatrix}$，试计算$\frac{\mathrm{d}\boldsymbol{A}^{-1}(x)}{\mathrm{d}x}$.

解 1： 由性质(3)知

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}x}\left\{\frac{1}{x^2-1}\begin{bmatrix} x & -1 \\ -1 & x \end{bmatrix}\right\} \\
&= \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{x^2-1}\right)\cdot\begin{bmatrix} x & -1 \\ -1 & x \end{bmatrix} + \frac{1}{x^2-1}\frac{\mathrm{d}}{\mathrm{d}x}\begin{bmatrix} x & -1 \\ -1 & x \end{bmatrix} \\
&= -\frac{2x}{(x^2-1)^2}\begin{bmatrix} x & -1 \\ -1 & x \end{bmatrix} + \frac{1}{x^2-1}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
&= \frac{1}{(x^2-1)^2}\begin{bmatrix} -x^2-1 & 2x \\ 2x & -x^2-1 \end{bmatrix}.
\end{aligned}$$

解 2:　由性质(5)知

$$\begin{aligned}\frac{\mathrm{d}\boldsymbol{A}^{-1}(x)}{\mathrm{d}x} &= -\boldsymbol{A}^{-1}(x)\frac{\mathrm{d}\boldsymbol{A}(x)}{\mathrm{d}x}\boldsymbol{A}^{-1}(x) \\ &= -\frac{1}{x^2-1}\begin{bmatrix} x & -1 \\ -1 & x \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\frac{1}{x^2-1}\begin{bmatrix} x & -1 \\ -1 & x \end{bmatrix} \\ &= -\frac{1}{(x^2-1)^2}\begin{bmatrix} x^2+1 & -2x \\ -2x & x^2+1 \end{bmatrix} \\ &= \frac{1}{(x^2-1)^2}\begin{bmatrix} -x^2-1 & 2x \\ 2x & -x^2-1 \end{bmatrix}.\end{aligned}$$

6.4.4　函数矩阵的积分

定义 6.4.6　若函数矩阵$\boldsymbol{A}(x) = \left(a_{ij}(x)\right)_{m\times n}$的所有元素$a_{ij}(x)$ $(i = 1,2,\cdots,m; j = 1,2,\cdots,n)$都在$[a,b]$上可积，则称$\boldsymbol{A}(x)$在$[a,b]$上可积，且

$$\int_a^b \boldsymbol{A}(x)\mathrm{d}x = \begin{bmatrix} \int_a^b a_{11}(x)\mathrm{d}x & \int_a^b a_{12}(x)\mathrm{d}x & \cdots & \int_a^b a_{1n}(x)\mathrm{d}x \\ \int_a^b a_{21}(x)\mathrm{d}x & \int_a^b a_{22}(x)\mathrm{d}x & \cdots & \int_a^b a_{2n}(x)\mathrm{d}x \\ \vdots & \vdots & & \vdots \\ \int_a^b a_{m1}(x)\mathrm{d}x & \int_a^b a_{m2}(x)\mathrm{d}x & \cdots & \int_a^b a_{mn}(x)\mathrm{d}x \end{bmatrix}.$$

函数矩阵的定积分有如下性质:

(1) $\int_a^b k\boldsymbol{A}(x)\mathrm{d}x = k\int_a^b \boldsymbol{A}(x)\mathrm{d}x,\quad k\in\mathbb{R}.$

(2) $\int_a^b [\boldsymbol{A}(x)+\boldsymbol{B}(x)]\mathrm{d}x = \int_a^b \boldsymbol{A}(x)\mathrm{d}x + \int_a^b \boldsymbol{B}(x)\mathrm{d}x.$

例 6.4.2　已知$\boldsymbol{A} = \begin{bmatrix} \sin x & -\cos x \\ \cos x & \sin x \end{bmatrix}$，求$\int_0^x \boldsymbol{A}(x)dx$和$\int_0^\pi \boldsymbol{A}(x)dx$.

解:
$$\begin{aligned}\int_0^x \boldsymbol{A}(x)dx &= \begin{bmatrix} \int_0^x \sin x dx & \int_0^x (-\cos x)dx \\ \int_0^x \cos x dx & \int_0^x \sin x dx \end{bmatrix} \\ &= \begin{bmatrix} 1-\cos x & -\sin x \\ \sin x & 1-\cos x \end{bmatrix}.\end{aligned}$$

$$\int_0^\pi \boldsymbol{A}(x)dx = \begin{bmatrix} 1-\cos\pi & -\sin\pi \\ \sin\pi & 1-\cos\pi \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

本 章 小 结

本章介绍了矩阵的微积分，包括向量和矩阵的范数、矩阵序列与极限、矩阵级数、矩阵函数的定义及计算、函数矩阵的微分与积分.

学习完本章内容后，应能达到如下基本要求：

1. 掌握向量范数、矩阵范数的定义及性质. 对于向量范数，应掌握向量的1−范数、2−范数、∞−范数和p−范数的定义及性质；对于矩阵范数，应掌握矩阵的列和范数（1−范数）、谱范数（2−范数）、行和范数（∞−范数）、Frobenius−范数的定义及性质. 掌握导出范数的定义、矩阵范数与向量范数的相容性、矩阵谱半径的定义及性质.

2. 掌握矩阵序列的定义、矩阵序列收敛的充要条件和性质、矩阵序列敛散性的判断方法.

3. 掌握矩阵级数的定义、矩阵级数收敛和绝对收敛的定义、矩阵级数绝对收敛的充要条件；矩阵幂级数的定义、矩阵幂级数绝对收敛的充要条件.

4. 掌握矩阵函数的幂级数定义、典型矩阵函数，了解矩阵函数的相似对角化计算方法.

5. 掌握函数矩阵的定义和基本运算，掌握函数矩阵的极限、函数矩阵对标量的导数、函数矩阵的积分.

习题 6

6-1　设$\boldsymbol{\alpha} \in \mathbb{C}^n$，试证明：$\frac{1}{n}\|\boldsymbol{\alpha}\|_1 \leq \|\boldsymbol{\alpha}\|_\infty \leq \|\boldsymbol{\alpha}\|_2 \leq \|\boldsymbol{\alpha}\|_1$.

6-2　设$a_1, a_2, \cdots, a_n \in \mathbb{R}^+$，$\boldsymbol{\alpha} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}} \in \mathbb{R}^n$，证明$\|\boldsymbol{\alpha}\| = (\sum_{i=1}^n a_i |x_i|^2)^{\frac{1}{2}}$是向量范数.

6-3　试证：对于$\boldsymbol{A} = (a_{ij}) \in \mathbb{C}^{m \times n}$，$\|\boldsymbol{A}\| = \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|$是矩阵范数.

6-4　证明矩阵$\boldsymbol{A}$的 Frobenius 范数满足矩阵范数定义的 4 个条件.

6-5　证明矩阵的 Frobenius 范数与向量的 2-范数相容.

6-6　对下列矩阵$\boldsymbol{A}$，求$\|\boldsymbol{A}\|_2$和$\|\boldsymbol{A}\|_\infty$.

(1)　$\boldsymbol{A} = \begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix}$　　(2)　$\boldsymbol{A} = \begin{bmatrix} i & 1+i \\ -i & 1-i \end{bmatrix}$

6-7　设$\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{C}^n$，$\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{n \times n}$，试证：

(1)　$\|\boldsymbol{\alpha} - \boldsymbol{\beta}\| \geq \|\boldsymbol{\alpha}\| - \|\boldsymbol{\beta}\|$.

(2)　$\|\boldsymbol{A} - \boldsymbol{B}\| \geq \|\boldsymbol{A}\| - \|\boldsymbol{B}\|$.

6-8　设 $\boldsymbol{A} = (a_{ij}) \in \mathbb{C}^{n \times n}$，试证：

(1)　$\|\boldsymbol{A}\| = [\mathrm{tr}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})]^{\frac{1}{2}}$ 是矩阵范数.

(2)　$\|\boldsymbol{A}\| = n\max_{i,j}|a_{ij}|$是矩阵范数.

6-9　对$\boldsymbol{\alpha} \in \mathbb{C}^n, \boldsymbol{A} \in \mathbb{C}^{n \times n}$，设$\|\boldsymbol{A}\|$是导出范数，且$\det\boldsymbol{A} \neq 0$，试证：

(1)　$\|\boldsymbol{A}^{-1}\| \geq \|\boldsymbol{A}\|^{-1}$.

(2)　$\|\boldsymbol{A}^{-1}\|^{-1} = \min_{\boldsymbol{\alpha} \neq \boldsymbol{0}} \frac{\|\boldsymbol{A}\boldsymbol{\alpha}\|}{\|\boldsymbol{\alpha}\|}$.

6-10　设$\boldsymbol{A} = \begin{bmatrix} 0 & a & a \\ a & 0 & a \\ a & a & 0 \end{bmatrix}$，求 a 为何值时有$\lim\limits_{k \to \infty} \boldsymbol{A}^k = \boldsymbol{0}$.

6-11　验证下列向量序列是否收敛，若收敛则试求其极限.

(1)　$\boldsymbol{x}_k = (\frac{1}{k}, \sin\left(\frac{1}{k}\right), e^{-k})^{\mathrm{T}}$;

(2)　$\boldsymbol{x}_k = (\frac{\cos(k)}{k}, e^{1/k}, 1)^{\mathrm{T}}$;

(3)　$\boldsymbol{x}_k = (k, 0, \frac{k+1}{k-1}, \sin(k))^{\mathrm{T}}$.

6-12　考察矩阵序列$\boldsymbol{A}_k = \begin{bmatrix} 1 & \frac{1}{2^k} \\ 0 & \sum_{i=1}^{k} \frac{1}{i} \end{bmatrix}$的敛散性.

6-13　讨论矩阵级数$\sum_{k=1}^{\infty} \boldsymbol{A}_k$的敛散性，其中

$$A_k = \begin{bmatrix} \frac{1}{k(k+1)} & (-1)^{k-1}\frac{1}{k} \\ 0 & \frac{1}{2^k} \end{bmatrix} \quad (k = 1,2,\dots).$$

6-14　判断矩阵幂级数$\sum_{k=0}^{+\infty} \begin{bmatrix} 0.1 & 0.7 \\ 0.3 & 0.6 \end{bmatrix}^k$的敛散性，若收敛，则试求其和.

6-15　判断矩阵幂级数$\sum_{k=0}^{+\infty} \frac{k}{6^k} \begin{bmatrix} 1 & -8 \\ -2 & 1 \end{bmatrix}^k$的敛散性.

6-16　已知函数矩阵$\boldsymbol{A}(x) = \begin{bmatrix} x+1 & 2x^2 \\ 1 & 2(x-1) \end{bmatrix}$，试求：$\frac{d}{dx}\boldsymbol{A}(x)$.

6-17　设函数矩阵$\boldsymbol{A}(x)$与题 6-16 相同，试求：$\frac{d}{dx}\boldsymbol{A}^{-1}(x)$.

6-18　已知函数矩阵$\boldsymbol{A}(x) = \begin{bmatrix} e^{2x} & x^2 \\ 3x & 0 \end{bmatrix}$，试求：$\int_0^1 \boldsymbol{A}(x)dx$.

6-19　已知函数矩阵$\boldsymbol{A}(x) = \begin{bmatrix} e^{2x} & xe^x & x^2 \\ e^{-x} & 2e^{2x} & 0 \\ 3x & 0 & 0 \end{bmatrix}$，试求：$\int_0^1 \boldsymbol{A}(x)dx$.

6-20　设函数矩阵$\boldsymbol{A}(x)$与题 6-19 相同，试求：$\frac{d}{dx}\left[\int_0^{x^2} \boldsymbol{A}(t)dt\right]$.