Chapter 7

Norm

对于实数和复数，由于定义了它们的绝对值(模)，可用模来表示其大小而带来许多方便.在希耳伯特空间中，可以定义内积.但是，并不是所有的元素都能够定义内积的.可以把模或者向量内积的概念进行推广.对于向量、矩阵甚至算子，引入在某种意义上表示其“大小”的纯量，这就是范数的概念.内积与角度有关.范数只与长度有关，与角度无关.或者说，可以把范数就理解为一种长度.不过，相比较长度而言，范数的概念要普遍得多.本章介绍向量、矩阵和算子的范数.向量范数和矩阵范数在矩阵分析中占有要的地位.算子的范数在泛函分析中占有重要的地位.而且，有了算子范数的概念，便于对各种算子进行理论分析.

§7.1 巴拿赫空间

7.1.1 巴拿赫空间

(1) 范数三公理

先来回顾关于向量的模的两个例子.

例 1 我们知道，平面向量 $\boldsymbol{x}=a\boldsymbol{i}+b\boldsymbol{j}$ 的模是用量 $\sqrt{a^2+b^2}$ 来描述的.如果用 $\|\boldsymbol{x}\|$ 来表示 $\boldsymbol{x}$ 的模，那么 $\|\boldsymbol{x}\|$ 具有下面三条性质：

(1) 若 $\boldsymbol{x}\neq 0$，则 $\|\boldsymbol{x}\|>0$；当且仅当 $\boldsymbol{x}=0$ 时，有 $\|\boldsymbol{x}\|=0$.

(2) $\|k\boldsymbol{x}\|=|k|\|\boldsymbol{x}\|$， k 任意实数.

(3) 对于任意平面向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$，有三角不等式.

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

例 2 由线性代数知，n 维欧氏空间中向量的模定义为

$$\|\boldsymbol{x}\| = (\boldsymbol{x}, \boldsymbol{x})^{1/2}$$

用θ表示零向量.这样定义的模也具有下面三条性质：

(1) 若 $\boldsymbol{x} \neq 0$,则$\|\boldsymbol{x}\| > 0$；当且仅当 $\boldsymbol{x} = 0$ 时，$\|\theta\| = 0$.

(2) 对任意实数 k 和任意向量 x,有

$$\|k\boldsymbol{x}\| = |k|\|\boldsymbol{x}\|$$

(3) 对任意向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$,有，

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

对于一般的线性空间，引入满足上述三条性质的纯量(或函数)，用它来描述向量的大小，并称之为范数.

定义 1 设 K 为一数域(等于 R 或 C)，V 是 K 上的向量空间.如果对于 V 中的任何一个元，指定一个正的实数，记作$\|f\|$，它满足以下关系：

(1) 三角不等式，对于所有 $f, g \in V$，$\|f+g\| \le \|f\| + \|g\|$ (7.1.1)

(2) 非负性，当且仅当 $f \equiv 0$ 时，$\|f\| = 0$ (7.1.2)

(3) 齐次性，对于数域 K 中的任何纯量 a，$\|af\| = |a| \cdot \|f\|$ (7.1.3)

则称$\|f\|$为 f 的**范数**.这三个条件称为**范数三公理**，也就是范数所应具有的三个性质.能够定义范数的空间称为**赋范空间**. 能够定义范数的向量空间称为**赋范向量空间**.当 K=R 或 C 时，分别称 V 为**实赋范向量空间**或**复赋范向量空间**.

注意，这些条件只是定义范数所必须满足的条件，或者说，这是范数概念的定义，并不是范数的数值的具体定义.如果用某种方法定义了一种关系，它满足这些条件，那么，这种关系就可以成为范数了.

定义 2 设 V 是数域 K 上的赋范空间.V 中任意两个元 f 和 g 之间的距离可以按照(7.1.5)定义为这两个元的差值的范数.如果 V 关于距离 $\rho(f,g)$ 是完备的，则称 V 是 K 上的**巴拿赫空间**.当 K=R 或 C 时，分别称 V 为**实巴拿赫空间**或**复巴拿赫空间**.

一般简单地称：完备的赋范向量空间就是**巴拿赫空间**.下面用 B 表示巴拿赫空间.

注意，巴拿赫空间可以没有内积.因此，巴拿赫空间的外延比希尔伯特空间的外延大.定义了内积之后一定可以定义范数.因而，任何一个希尔伯特空间都是巴拿赫空间.

当距离就是用范数来定义的时候，我们说空间关于距离是完备性的，就可以等价地说空间关于范数是完备性的.有的文献上定义的希尔伯特空间就是关于范数是完备的. 希尔伯特空间定义的另一个表述是，定义了内积的完备的赋范线性空间.

例 3 在有理数向量空间中，选择数列

$$S_n = \sum_{m=1}^{n} \frac{1}{m^2}$$

它是有理数的柯西序列，是可赋范的.可定义任何一个元的绝对值作为其范数，这样定义的范数满足范数三公理.并可定义两个元之差的范数作为这两个元之间的距离.由此，上述有理数的

柯西序列是有极限的，但是极限$\lim\limits_{n\to\infty}S_n=\lim\limits_{n\to\infty}\sum\limits_{m=1}^{n}\dfrac{1}{m^2}\to\dfrac{\pi^2}{6}$是个无理数，不在有理数向量空间之内.因此，有理数向量空间不是一个完备的赋范空间，即，它不是一个巴拿赫空间.

例 4 在区间$[a,b]$上的光滑连续函数组成空间，记作$C_1[a,b]$.光滑连续表示函数至少是一次可导.对于这个空间，一种定义范数的方法是

$$\|f(x)\|=\max_{x\in[a,b]}|f(x)| \tag{7.1.9}$$

容易证明，由此定义的范数满足范数三公理.因此这是一个赋范空间.由于闭区间上收敛的光滑连续函数序列在闭区间上也是一致收敛的，因此它的极限函数也是光滑连续函数，故$C_1[a,b]$是一个巴拿赫空间.

要注意的是，在区间$[a,b]$上的连续函数组成空间$C[a,b]$是不完备的.在第二章习题 3 已经给出了一个这样的例子.因此，如果在闭区间$[a,b]$上的连续函数不光滑，不能构成一个巴拿赫空间.

例 5 由长度为 n 的实数组成的向量空间R^n，对于向量空间中的元素

$x=(x_1,x_2,\cdots,x_n)\in R^n$，定义范数

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

由长度为 n 的复数组成的向量空间C^n，对于向量空间中的元素$z=(z_1,z_2,\cdots,z_n)\in C^n$，定义范数

$$\|z\|=\left(\sum_{i=1}^{n}|z_i|^2\right)^{1/2}$$

按此定义范数之后，R^n是实巴拿赫空间，C^n是复巴拿赫空间.

在进一步定义具体的范数之前，我们先介绍两个不等式.

7.1.2 赫尔德不等式

定理 1 设

$$p>1,\ q=\frac{p}{p-1} \tag{7.1.10}$$

则

$$\sum_{k=1}^{n}a_kb_k\le(\sum_{k=1}^{n}a_k^p)^{1/p}(\sum_{k=1}^{n}b_k^q)^{1/q} \tag{7.1.11}$$

其中$a_k,b_k\ge0$.本节以下的 p 和 q 总是指(7.1.10)式.式(7.1.11)称为**赫尔德不等式**.

证明 先来证明：如果 u 和 v 均非负，则总有

$$uv \le \frac{u^p}{p} + \frac{v^q}{q} \qquad (7.1.12)$$

考虑函数 $v = u^{p-1}$，不失一般性，考虑如图 7.1 所示的图形.定义以下两个积分.

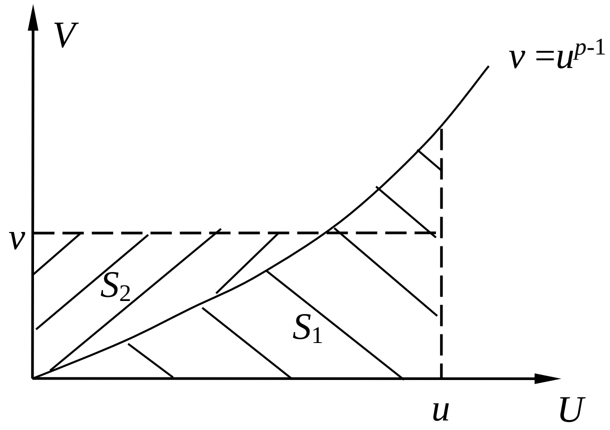

图 7.1

$$S_1 = \int_0^u U^{p-1} dU = \frac{1}{p} u^p, \qquad S_2 = \int_0^v V^{\frac{1}{p-1}} dV = \int_0^v V^{q-1} dV = \frac{1}{q} v^q$$

$S_1 + S_2 \ge uv$,且等号只在 $v = u^{p-1}$ 时成立.于是

$$uv \le \frac{u^p}{p} + \frac{v^q}{q}$$

这就是(7.1.12)式.式(7.1.10)意味着

$$\frac{1}{p} + \frac{1}{q} = 1 \qquad (7.1.13)$$

现在，如果 u 和 v 各为某一个向量的分量 u_k 和 v_k，那么

$$u_k v_k \le \frac{u_k^p}{p} + \frac{v_k^q}{q} \qquad (7.1.12a)$$

并且使得

$$\sum_{k=1}^{n} u_k^p = 1, \sum_{k=1}^{n} v_k^q = 1 \qquad (7.1.14)$$

那么在(7.1.12)式两边做求和 $\sum_{k=1}^{n}$ ，就可得

$$\sum_{k=1}^{n} u_k v_k \le \frac{1}{p} + \frac{1}{q} = 1 \qquad (7.1.15)$$

令

$$u_k = \frac{a_k}{(\sum_{j=1}^{n} a_j^p)^{1/p}}, v_k = \frac{b_k}{(\sum_{j=1}^{n} b_j^q)^{1/q}}$$

显然，这样定义的 u_k 和 v_k 是符合(7.1.14)式的，因此由(7.1.15)式得到

$$\sum_{k=1}^{n}\frac{a_k}{(\sum_{j=1}^{n}a_j^p)^{1/p}}\frac{b_k}{(\sum_{j=1}^{n}b_j^q)^{1/q}}\le 1$$

于是得到赫尔德不等式(7.1.11).**证明完毕**.

特例，令 $p=q=2,$ 代入式(7.1.11)，有

$$\sum_{k=1}^{n}a_k b_k \le \sqrt{\sum_{k=1}^{n}a_k^2}\sqrt{\sum_{k=1}^{n}b_k^2} \tag{7.1.16}$$

这就是第二章 2.1.2 小节例 17 的柯西不等式.

可以把(7.1.14)和(7.1.15)式中的求和改成积分.即

$$u(x)\ v(x)\le\frac{u^p(x)}{p}+\frac{v^q(x)}{q}$$

如果

$$\int u^p(x)\ \mathrm{d}x=1,\ \int v^q(x)\ \mathrm{d}x=1 \tag{7.1.17}$$

那么那么在(7.1.12)式两边做积分 $\int \mathrm{d}x$，就可得

$$\int u(x)\ v(x)\ \mathrm{d}x\le 1 \tag{7.1.18}$$

现在令

$$u(x)=\frac{f(x)}{(\int f^p(x)\mathrm{d}x)^{1/p}},v(x)=\frac{g(x)}{(\int g^q(x)\mathrm{d}x)^{1/q}}$$

其中 $f(x)>0,g(x)>0$.这样定义的 u 和 v 是符合(7.1.17)式的，因此由(7.1.18)式得到

$$\int \mathrm{d}x\frac{f(x)}{(\int f^p(x)\mathrm{d}x)^{1/p}}\frac{g(x)}{(\int g^q(x)\mathrm{d}x)^{1/q}}\le 1$$

由此我们便得到积分形式的赫尔德不等式：

$$\int f(x)g(x)\mathrm{d}x\le(\int f^p(x)\mathrm{d}x)^{1/p}(\int g^q(x)\mathrm{d}x)^{1/q} \tag{7.1.19}$$

在本章习题中，证明了加权赫尔德不等式.

要特别强调的是，以上不等式只能适用于 p>1，不能用于 p=1.

7.1.3 闵可夫斯基不等式

定理 2 对任何 $p\ge 1$，有

$$(\sum_{i=1}^{n}|a_i+b_i|^p)^{1/p}\le(\sum_{i=1}^{n}|a_i|^p)^{1/p}+(\sum_{i=1}^{n}|b_i|^p)^{1/p} \tag{7.1.20}$$

此式称为**闵可夫斯基不等式**.

证明 对于 p=1，不等式是显然的.下面证明 p>1 的情况.以 $q=p/(p-1)$ 代入下式

$$\sum_{i=1}^{n}|a_i+b_i|^p=\sum_{i=1}^{n}|a_i+b_i||a_i+b_i|^{p-1}$$

等式右边有

$$\sum_{i=1}^{n}|a_i+b_i||a_i+b_i|^{p/q}\le\sum_{i=1}^{n}|a_i||a_i+b_i|^{p/q}+\sum_{i=1}^{n}|b_i||a_i+b_i|^{p/q}$$

由赫尔德不等式，得

$$\sum_{i=1}^{n}|a_i+b_i|^p\le(\sum_{i=1}^{n}|a_i|^p)^{1/p}(\sum_{i=1}^{n}|a_i+b_i|^p)^{1/q}+(\sum_{i=1}^{n}|b_i|^p)^{1/p}(\sum_{i=1}^{n}|a_i+b_i|^p)^{1/q}$$

$$=\left[(\sum_{i=1}^{n}|a_i|^p)^{1/p}+(\sum_{i=1}^{n}|b_i|^p)^{1/p}\right](\sum_{i=1}^{n}|a_i+b_i|^p)^{1/q}$$

两边除以$(\sum_{i=1}^{n}|a_i+b_i|^p)^{1/q}$，注意到(7.1.13)式，即得到(7.1.20)式.**证明完毕**.

积分形式的闵可夫斯基不等式是

$$\left(\int|f(x)+g(x)|^p\,\mathrm{d}x\right)^{1/p}\le\left(\int|f(x)|^p\,\mathrm{d}x\right)^{1/p}+\left(\int|g(x)|^p\,\mathrm{d}x\right)^{1/p}\qquad(7.1.21)$$

在本章习题中，证明了加权闵可夫斯基不等式.

赫尔德不等式和闵可夫斯基不等式还有一些推广的形式.

例 6 L_p空间，它定义为由函数$f(x)$构成的向量空间.这个空间中的所有元$f(x)$满足条件

$$\int_D|f(x)|^p\,\mathrm{d}x<\infty$$

其中D是实轴的某个子集(并不必须有限).这个空间的一种范数由

$$\|f\|=[\int_D|f(x)|^p\,\mathrm{d}x]^{1/p}\qquad(7.1.22)$$

给出.可以证明这一范数的定义满足范数三公理.L_p空间也称为p次可积空间，它是完备的，因此是一个巴拿赫空间.

§7.2 向量范数

7.2.1 向量范数

闵可夫斯基不等式的形式正好与范数的三角不等式是相同的.因而，由闵可夫斯基不等式，我们可以引入常用的p-范数的概念.

定义 1 设$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$，对任意$p\ge1$，称量

$$\left(\sum_{i=1}^{n}|x_i|^p\right)^{1/p}\qquad(7.2.1)$$

为向量$\boldsymbol{x}$的***p*-范数**，记作$\|\boldsymbol{x}\|_p$.

易知，$\|\boldsymbol{x}\|_p$满足非负性和齐次性条件.由式(7.1.20),又有$\|\boldsymbol{x}+\boldsymbol{y}\|_p \le \|\boldsymbol{x}\|_p + \|\boldsymbol{y}\|_p$.因此，$\|\boldsymbol{x}\|_p$满足范数的三条性质.

常用的 p-范数有下面三种：

(1) 1-范数$\|\boldsymbol{x}\|_1$，即 $p=1$，这时

$$\|\boldsymbol{x}\|_1 = |x_1| + |x_2| + \cdots + |x_n|$$

(2) 2-范数$\|\boldsymbol{x}\|_2$，即 $p=2$，这就是 7.1.1 小节例 5 中定义的欧氏空间中的向量范数，称为欧氏范数.

(3) ∞-范数$\|\boldsymbol{x}\|_\infty$，即

$$\|\boldsymbol{x}\|_\infty = \lim_{p\to\infty}\|\boldsymbol{x}\|_p = \lim_{p\to\infty}(|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{1/p}$$

命题 1 $\|\boldsymbol{x}\|_\infty = \max\limits_i |x_i|$

证明 令$\alpha = \max\limits_i |x_i|$，$\beta_i = \left|\dfrac{x_i}{\alpha}\right| \le 1, (i=1,2,\cdots,n)$.于是

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

在$\beta_1,\beta_2,\cdots,\beta_n$中至少有一个等于 1，故

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

因为$\lim\limits_{p\to\infty} n^{1/p} = 1$故

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

因此

$$\|\boldsymbol{x}\|_\infty = \lim_{p\to\infty}\|\boldsymbol{x}\|_p = \alpha = \max_i |x_i|$$

证明完毕.

以上证明简捷写为如下过程：

$$\|\boldsymbol{x}\|_\infty = \lim_{p\to\infty}\|\boldsymbol{x}\|_p = \lim_{p\to\infty}[\alpha^p(\frac{|x_1|^p}{\alpha^p} + \frac{|x_2|^p}{\alpha^p} + \cdots + \frac{|x_n|^p}{\alpha^p})]^{1/p}$$

$$= \alpha\lim_{p\to\infty}(\frac{|x_1|^p}{\alpha^p} + \frac{|x_2|^p}{\alpha^p} + \cdots + \frac{|x_n|^p}{\alpha^p})^{1/p} = \alpha(0+0+\cdots+1+\cdots+0)^{1/p} = \alpha$$

注意，即使$\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$的分量中，有 k 个元素符合$\alpha = \max\limits_i |x_i|$，仍然有$\|\boldsymbol{x}\|_\infty = \alpha$.

与有限维向量空间相仿，对于连续函数空间可同样定义 p-范数的概念.设有 L_p $(1 \le p < \infty)$ 空间，即 p 次可积空间.它定义为由函数 $f(x) \in C[a,b]$ 构成的向量空间.这个空间中的所有元 $f(x)$ 满足条件

$$\int_a^b |f(x)|^p \mathrm{d}x < \infty \tag{7.2.2}$$

定义 2 对于 $f(x) \in C[a,b]$，且满足(7.2.2)式，则

$$\|f\|_p = [\int_a^b |f(x)|^p \mathrm{d}x]^{1/p} < \infty \tag{7.2.3}$$

称为函数 $f(x)$ 的 ***p*-范数**，也称**赫尔德范数**.

$$\|f\|_\infty = \max_{x\in[a,b]} |f(x)|$$

称为 $C[a,b]$ 上的**一致范数**或者**切比雪夫范数**.

显然，一致范数与有限维数域空间中向量的∞-范数是相对应的.因此也称为函数 $f(x)$ 的 **∞-范数**.它也就是(7.2.3)式中取 $p=\infty$ 的特例.

式(7.2.3)的另外两个特例是

1-范数：$\|f\|_1 = \int_a^b |f(x)| \mathrm{d}x$；　2-范数：　$\|f\|_2 = [\int_a^b | f(x)|^2 \mathrm{d}x]^{1/2}$

容易证明，由(7.2.3)式定义的函数的范数满足范数三公理. L_p 空间也称为 p 次可积空间，它是完备的，因此是一个巴拿赫空间.

范数 $\|\boldsymbol{x}\|_a$ 是变量 $x_1, x_2, \cdots, x_n$ 的函数，记为

$$\|\boldsymbol{x}\|_a = \varphi(x_1, x_2, \cdots, x_n) \tag{7.2.4}$$

定理 1 向量 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$ 的范数 $\|\boldsymbol{x}\|_a = \varphi(x_1, x_2, \cdots, x_n)$ 是 $x_1, x_2, \cdots, x_n$ 的连续函数.

证明 设 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ 是空间 V 的一组基底，则对任意 $\boldsymbol{x} \in V$，有

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \cdots + x_n\boldsymbol{e}_n \tag{7.2.5}$$

和

$$\boldsymbol{x}' = x_1'\boldsymbol{e}_1 + x_2'\boldsymbol{e}_2 + \cdots + x_n'\boldsymbol{e}_n$$

向量 $\boldsymbol{x}'$ 的范数就是

$$\|\boldsymbol{x}'\|_a = \varphi(x_1', x_2', \cdots, x_n')$$

那么

$$\begin{aligned}|\varphi(x_1',x_2',\cdots,x_n')-\varphi(x_1,x_2,\cdots,x_n)|&=\big|\|\boldsymbol{x}'\|_a-\|\boldsymbol{x}\|_a\big|\le\|\boldsymbol{x}'-\boldsymbol{x}\|_a\\&=\|(x_1'-x_1)\boldsymbol{e}_1+(x_2'-x_2)\boldsymbol{e}_2+\cdots+(x_n'-x_n)\boldsymbol{e}_n\|_a\\&\le|x_1'-x_1|\|\boldsymbol{e}_1\|_a+|x_2'-x_2|\|\boldsymbol{e}_2\|_a+\cdots+|x_n'-x_n|\|\boldsymbol{e}_n\|_a\end{aligned}$$

因为$\|\boldsymbol{e}_i\|_a$是常数，所以当$|x_i'-x_i|$充分小时，有

$$|\varphi(x_1',x_2',\cdots,x_n')-\varphi(x_1,x_2,\cdots,x_n)|<\delta$$

这就证明了$\varphi(x_1,x_2,\cdots,x_n)$是$x_1,x_2,\cdots,x_n$的连续函数.**证明完毕**.

最后我们给出带权范数的形式

$$\|\boldsymbol{x}\|_p=\left(\sum_{i=1}^n w_i|x_i|^p\right)^{1/p}\quad 和\quad \|f\|_p=[\int_a^b\rho(x)|f(x)|^p\mathrm{d}x]^{1/p}$$

其中权都是大于零的.

§7.3 矩阵范数

7.3.1 矩阵范数

(1) 矩阵的诱导范数

在这一节，我们进一步把范数的概念推广到$m\times n$矩阵上.

定义 1 设$A\in F^{m\times n}$(F是实数域R或复数域C)是任一个$m\times n$矩阵,按照某个法则在$F^{m\times n}$上规定A的一个实函数，记作$\|A\|$，此函数具有下述性质：

(i) 非负性.若$A\ne 0$, 则$\|A\|>0,\|\theta\|=0$.

(ii) 齐次性.对任意纯量k， $\|kA\|=|k|\|A\|$.

(iii) 三角不等式.对任意$A,B\in F^{m\times n},\|A+B\|\le\|A\|+\|B\|$.

则称$\|A\|$为矩阵A的**范数**.

在矩阵理论和应用中，矩阵的乘法占有特殊重要的地位，矩阵和向量常以乘积的形式出现.因此，在讨论矩阵范数时，应考虑矩阵作用后的向量范数与矩阵范数及原向量范数的协调.为此，有下面的定义.

定义 2 (范数相容性) 设在$F^{m\times n},F^{n\times p},F^{m\times p}$定义了三种范数$\|\bullet\|_{m\times n}$， $\|\bullet\|_{n\times p}$， $\|\bullet\|_{m\times p}$对任意的矩阵$A\in F^{m\times n}$， $B\in F^{n\times p}$， 恒有

$$\|AB\|_{m\times p}\le\|A\|_{m\times n}\|B\|_{n\times p}\qquad(7.3.1)$$

则称范数$\|\bullet\|_{m\times n}$， $\|\bullet\|_{n\times p}$和$\|\bullet\|_{m\times p}$是**相容的**.

如果 $p=1$，这时 B 是 n 维列向量，式(7.3.1)就可写成

$$\|A\boldsymbol{x}\|_{m\times 1} \le \|A\|_{m\times n}\|\boldsymbol{x}\|_{n\times 1} \qquad \text{或写成} \qquad \|A\boldsymbol{x}\|_{\alpha} \le \|A\|_{m\times n}\|\boldsymbol{x}\|_{\beta}$$

其中 $\|\bullet\|_{\alpha}$ 和 $\|\bullet\|_{\beta}$ 表示 R^m 和 R^n 中的两种向量范数.

为了在矩阵分析的讨论中不致产生困难，我们所指的矩阵范数通常总是相容的.

下面我们来讨论几种具体的矩阵范数.我们总是设 $\boldsymbol{x}\in R^m, \boldsymbol{y}\in R^n$ 在相应的 m 和 n 维线性空间中规定了向量的某种范数 $\|\boldsymbol{x}\|_{\alpha}$ 和 $\|\boldsymbol{y}\|_{\beta}$.

定义 3 设 $A\in F^{m\times n}$，规定矩阵 A 的范数如下：

$$\|A\|_{\beta} = \max_{\boldsymbol{x}\neq 0}\frac{\|A\boldsymbol{x}\|_{\beta}}{\|\boldsymbol{x}\|_{\beta}} = \max_{\|\boldsymbol{x}\|_{\beta}=1}\|A\boldsymbol{x}\|_{\beta} \tag{7.3.2}$$

即矩阵 A 的范数取自当 $\|\boldsymbol{x}\|_{\beta}=1$ 时，所有向量范数 $\|A\boldsymbol{x}\|_{\beta}$ 的最大值.由式(7.3.2)定义的矩阵范数 $\|A\|_{\beta}$ 称为**诱导范数**.它表示矩阵范数 $\|\bullet\|_{\beta}$ 由向量范数 $\|\bullet\|_{\beta}$ 导出.

由式(7.3.2)定义的矩阵范数，满足定义 1 的三条性质和相容性条件.事实上，$\|\theta\|=0$，且 $A\neq 0$ 时，$\|A\|>0$.又

$$\|kA\| = \max_{\|\boldsymbol{x}\|_{\beta}=1}\|kA\boldsymbol{x}\|_{\beta} = |k|\max_{\|\boldsymbol{x}\|_{\beta}=1}\|A\boldsymbol{x}\|_{\beta} = |k|\|A\|$$

再有，对 $A,B\in F^{m\times n}$

$$\|A+B\| = \max_{\|\boldsymbol{x}\|_{\beta}=1}\|(A+B)\boldsymbol{x}\|_{\beta} \le \max_{\|\boldsymbol{x}\|_{\beta}=1}\|A\boldsymbol{x}\|_{\beta} + \max_{\|\boldsymbol{x}\|_{\beta}=1}\|B\boldsymbol{x}\|_{\beta} = \|A\|+\|B\|$$

一般来说，有

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

即矩阵范数 $\|A\|$ 与向量范数 $\|\boldsymbol{x}\|_{\alpha}$、$\|\boldsymbol{y}\|_{\beta}$ 相容.

当 $\|\boldsymbol{x}\|_{\beta}$ 分别取 $\|\boldsymbol{x}\|_1$，$\|\boldsymbol{x}\|_2$ 以及 $\|\boldsymbol{x}\|_{\infty}$ 时，我们得到三种矩阵范数，分别记为 $\|A\|_1, \|A\|_2, \|A\|_{\infty}$.

定理 1 设 $A\in F^{m\times n}$，则

(i) $$\|A\|_1 = \max_j\left(\sum_{i=1}^m |a_{ij}|\right), (j=1,2,\cdots,n) \tag{7.3.3}$$

(ii) $$\|A\|_2 = \max_j\left(\lambda_j(A^{+}A)\right)^{1/2}, \tag{7.3.4}$$

其中 $\lambda_j(A^{+}A)$ 表示矩阵 $A^{+}A$ 的第 j 个特征值.

(iii) $\|A\|_{\infty}=\max_i\left(\sum_{j=1}^{n}|a_{ij}|\right),(i=1,2,\cdots,m)$ (7.3.5)

证明 (i) 令 $w=\max_j\left(\sum_{i=1}^{m}|a_{ij}|\right),(j=1,2,\cdots,n)$

设 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 且 $\|\boldsymbol{x}\|_1=1$.又设 $A=[\alpha_1,\alpha_2,\cdots,\alpha_n]$ 则有

$w=\max_j\|\alpha_j\|_1$

$$A\boldsymbol{x}=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{1n}&a_{n2}&\cdots&a_{nn}\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\begin{pmatrix}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n\\a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n\\\vdots\\a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n\end{pmatrix}$$

$$=\begin{pmatrix}a_{11}x_1\\a_{21}x_1\\\vdots\\a_{n1}x_1\end{pmatrix}+\begin{pmatrix}a_{12}x_2\\a_{22}x_2\\\vdots\\a_{n2}x_2\end{pmatrix}+\cdots+\begin{pmatrix}a_{1n}x_n\\a_{2n}x_n\\\vdots\\a_{nn}x_n\end{pmatrix}=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n$$

于是

$$\begin{aligned}\|A\boldsymbol{x}\|_1&=\|x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n\|_1\\&\le|x_1|\|\alpha_1\|_1+|x_2|\|\alpha_2\|_1+\cdots+|x_n|\|\alpha_n\|_1\\&\le\left(|x_1|+|x_2|+\cdots+|x_n|\right)w=\|\boldsymbol{x}\|_1 w=w\end{aligned}$$

现在把 $\|\alpha_j\|_1$ 明确写出为

$$\|\alpha_j\|_1=\sum_{i=1}^{m}|a_{ij}|$$

如果第 r 列的 $\|\alpha_r\|_1$ 最大，就取 $\boldsymbol{x}_r=(0,0\cdots,1_{(r)},0\cdots,0)^{\mathrm{T}}$.则 $\|\boldsymbol{x}_r\|=1$，且

$$A\boldsymbol{x}_r=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1,r-1}&a_{1,r}&a_{1,r+1}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2,r-1}&a_{2,r}&a_{2,r+1}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots&\vdots&\vdots&\ddots&\vdots\\a_{s-1,1}&a_{s-1,2}&\cdots&a_{s-1,r-1}&a_{s-1,r}&a_{s-1,r+1}&\cdots&a_{s-1,n}\\a_{s,1}&a_{s,2}&\cdots&a_{s,r-1}&a_{s,r}&a_{s,r+1}&\cdots&a_{s,n}\\a_{s+1,1}&a_{s+1,2}&\cdots&a_{s+1,r-1}&a_{s+1,r}&a_{s+1,r+1}&\cdots&a_{s+1,n}\\\vdots&\vdots&\ddots&\vdots&\vdots&\vdots&\ddots&\vdots\\a_{m1}&a_{m2}&\cdots&a_{m,r-1}&a_{m,r}&a_{m,r+1}&\cdots&a_{mn}\end{pmatrix}\begin{pmatrix}0\\0\\\vdots\\0\\1_{(r)}\\0\\\vdots\\0\end{pmatrix}=\begin{pmatrix}a_{1r}\\a_{2,r}\\\vdots\\a_{s-1,r}\\a_{s,r}\\a_{s+1,r}\\\vdots\\a_{m,r}\end{pmatrix}=\alpha_r$$

$$\|A\boldsymbol{x}_r\|_1=\sum_{i=1}^{m}|a_{ir}|=\max_j\|\alpha_j\|_1=w$$

所以

$$\|A\|_1=\sum_{i=1}^{m}|a_{ir}|=\max_j\left(\sum_{i=1}^{m}|a_{ij}|\right),\quad (j=1,2,\cdots,n)$$

(ii) 由向量范数$\|\boldsymbol{x}\|_2$的定义知

$$\|A\boldsymbol{x}\|_2=(A\boldsymbol{x},A\boldsymbol{x})^{1/2}=(\boldsymbol{x},A^{+}A\boldsymbol{x})^{1/2}$$

因为$A^{+}A$是正定或半正定矩阵，所以$A^{+}A$的n个特征值非负，于是

$$\|A\boldsymbol{x}\|_2=\max_{\|\boldsymbol{x}\|_2=1}(A\boldsymbol{x},A\boldsymbol{x})^{1/2}=\max_{\|\boldsymbol{x}\|_2=1}(\boldsymbol{x},A^{+}A\boldsymbol{x})^{1/2}=\max_i\left(\lambda_i(A^{+}A)\right)^{1/2}$$

具体证明如下. 令$\boldsymbol{y}_i$是矩阵$P=A^{+}A$的属于第i个特征值λ_i的特征向量且模为 1.

$$Py_i=A^{+}Ay_i=\lambda_i y_i$$

由n个$\boldsymbol{y}_i$作为列矢量构成P的本征向量矩阵，记为V.由P的n个特征值构成对角矩阵Λ.即

$$PV=V\Lambda$$

可以得到

$$\|A\boldsymbol{x}\|_2=\max_{\|\boldsymbol{x}\|_2=1}(\boldsymbol{x},P\boldsymbol{x})^{1/2}=\max_{\|\boldsymbol{x}\|_2=1}(\boldsymbol{x},V\Lambda V^{-1}\boldsymbol{x})^{1/2}\le\max_{\|\boldsymbol{x}\|_2=1,i}(\boldsymbol{x},V\lambda_i IV^{-1}\boldsymbol{x})^{1/2}$$

$$=\max_{\|\boldsymbol{x}\|_2=1,i}\lambda_i^{1/2}(\boldsymbol{x},VV^{-1}\boldsymbol{x})^{1/2}=\max_{\|\boldsymbol{x}\|_2=1,i}\lambda_i^{1/2}(\boldsymbol{x},\boldsymbol{x})^{1/2}=\max_i\left(\lambda_i(A^{+}A)\right)^{1/2}$$

现在我们来寻找左边的最大值.这只要令$\boldsymbol{x}$就是P的最大的那个特征值的的特征向量.那么，

$$\|A\boldsymbol{x}\|_2=\max_{\|\boldsymbol{y}_i\|_2=1}(\boldsymbol{y}_i,A^{+}A\boldsymbol{y}_i)^{1/2}=\max_i(\boldsymbol{y}_i,\lambda_i\boldsymbol{y}_i)^{1/2}$$

$$=\max_i\lambda_i^{1/2}(\boldsymbol{y}_i,\boldsymbol{y}_i)^{1/2}=\max_i\left(\lambda_i(A^{+}A)\right)^{1/2}$$

因此，(7.3.4)的证.

(iii) 令$w=\max_i\left(\sum_{j=1}^{n}|a_{ij}|\right),(i=1,2,\cdots,m)$.对任意的$x\in R^n$，有

$$\|A\boldsymbol{x}\|_\infty=\max_{1\le i\le m}\left|\sum_{j=1}^{n}|a_{ij}x_j|\right|\le\max_{1\le i\le m}\left(\sum_{j=1}^{n}|a_{ij}|\right)\|\boldsymbol{x}\|_\infty=w\|\boldsymbol{x}\|_\infty$$

此式对于任意$\|\boldsymbol{x}\|_\infty$都成立.两边取$\|\boldsymbol{x}\|_\infty\to 1$的极限.则有$\|A\|_\infty\le w$.现在假定矩阵$A$的第$r$行的各列矩阵元的绝对值之和最大.

$$\max_{1\le i\le m}\sum_{j=1}^{n}|a_{ij}|=\sum_{j=1}^{n}|a_{rj}|$$

取$\boldsymbol{x}$是如下的向量

$$\boldsymbol{x}=\left(\frac{a_{r1}^{*}}{|a_{r1}|},\frac{a_{r2}^{*}}{|a_{r2}|},\cdots,\frac{a_{rn}^{*}}{|a_{rn}|}\right)^{\mathrm{T}}$$

其中 a_{rj}^* 是 a_{rj} 的复共轭.则 $\|\boldsymbol{x}\|_\infty = 1$.在向量 $A\boldsymbol{x}$ 的 n 个分量中，显然第 r 个分量的绝对值最大.结果就是

$$A\boldsymbol{x} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1,s-1} & a_{1,s} & a_{1,s+1} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2,s-1} & a_{2,s} & a_{2,s+1} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{r-1,1} & a_{r-1,2} & \cdots & a_{r-1,s-1} & a_{r-1,s} & a_{r-1,s+1} & \cdots & a_{r-1,n} \\ a_{r,1} & a_{r,2} & \cdots & a_{r,s-1} & a_{r,s} & a_{r,s+1} & \cdots & a_{r,n} \\ a_{r+1,1} & a_{r+1,2} & \cdots & a_{r+1,s-1} & a_{r+1,s} & a_{r,s+1} & \cdots & a_{r+1,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{m,s-1} & a_{m,s} & a_{m,s+1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} a_{r,1}^* / |a_{r,1}| \\ a_{r,2}^* / |a_{r,2}| \\ \vdots \\ a_{r,s-1}^* / |a_{r,s-1}| \\ a_{r,s}^* / |a_{r,s}| \\ a_{r,s+1}^* / |a_{r,s+1}| \\ \vdots \\ a_{r,n}^* / |a_{r,n}| \end{pmatrix}$$

$$\|A\boldsymbol{x}\|_\infty = \sum_{j=1}^n |a_{rj}| = \max_{1\le i\le m} \sum_{j=1}^n |a_{ij}| = w$$

证明完毕.

范数 $\|A\|_1$ 称为**列和范数**，$\|A\|_\infty$ 称为**行和范数**，$\|A\|_2$ 称为**谱范数**.关于谱范数在下一小节将作进一步的讨论.

若 $A \in F^{n\times n}, B \in F^{n\times p}$，因为对任意 $\boldsymbol{y} \in R^p$，有

$$\|AB\boldsymbol{y}\|_2 \le \|A\|_2 \|B\boldsymbol{y}\|_2 \le \|A\|_2 \|B\|_2 \|\boldsymbol{y}\|_2$$

此式对于任意 $\|\boldsymbol{y}\|_2$ 都是成立的.两边取 $\|\boldsymbol{y}\|_2 \to 1$ 的极限.

$$\|AB\|_2 = \max_{\|\boldsymbol{y}\|_2=1} \|AB\boldsymbol{y}\|_2 \le \max_{\|\boldsymbol{y}\|_2=1} \|A\|_2 \|B\|_2 \|\boldsymbol{y}\|_2 = \|A\|_2 \|B\|_2$$

故有

$$\|AB\|_2 \le \|A\|_2 \|B\|_2 \qquad (7.3.6)$$

即矩阵范数 $\|\bullet\|_2$ 是相容的.

命题 1 单位矩阵的诱导范数总是等于 1，$\|I\|_\beta = 1$.

直接从定义式(7.3.2)是得到

$$\|I\|_\beta = \max_{\boldsymbol{x}\ne 0} \frac{\|I\boldsymbol{x}\|_\beta}{\|\boldsymbol{x}\|_\beta} = \max_{\boldsymbol{x}\ne 0} \frac{\|\boldsymbol{x}\|_\beta}{\|\boldsymbol{x}\|_\beta} = 1, \qquad \|I\|_\beta = \max_{\|\boldsymbol{x}\|_\beta=1} \|I\boldsymbol{x}\|_\beta = \max_{\|\boldsymbol{x}\|_\beta=1} \|\boldsymbol{x}\|_\beta = 1$$

注意，诱导范数是范数中的一类.还可以定义除了诱导范数之外其它的范数.单位矩阵的诱导范数总是 1，但是其它的范数不见得是 1.例如定义一个矩阵 A 的范数是 $\|A\| = 2\|A\|_\beta$，其中 $\|A\|_\beta$ 是诱导范数，则这样定义的范数是满足范数三公理的.可是在这样的定义下，单位矩阵的范数不是 1.

(2) 矩阵的弗罗贝尼乌斯范数

除了上面讨论的三种矩阵范数以外，还有一种常用的矩阵范数.做叫弗罗贝尼乌斯范数.

定义 4 设 $A \in F^{m\times n}$，由下式定义的范数

$$\|A\|_{\mathrm{F}} = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2} \tag{7.3.7}$$

叫做矩阵的**弗罗贝尼乌斯范数**，简称 **F 范数**.

这一定义也可以写成如下形式：

$$\|A\|_{\mathrm{F}} = \left(\sum_{i=1}^{m} (A^{+}A)_{ii} \right)^{1/2} = \left(\sum_{j=1}^{n} (AA^{+})_{jj} \right)^{1/2}$$

或者简写成

$$\|A\|_{\mathrm{F}} = \left(\mathrm{Tr}(A^{+}A) \right)^{1/2} = \left(\mathrm{Tr}(AA^{+}) \right)^{1/2}$$

把矩阵明确写出如下。

$$\mathrm{Tr}(AA^{+}) = \mathrm{Tr}\left[\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{1n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ a_{41} & a_{42} & a_{43} & \cdots & a_{4n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} a_{11}^{*} & a_{21}^{*} & a_{31}^{*} & a_{41}^{*} & \cdots & a_{m1}^{*} \\ a_{12}^{*} & a_{22}^{*} & a_{32}^{*} & a_{42}^{*} & \cdots & a_{m2}^{*} \\ a_{13}^{*} & a_{23}^{*} & a_{33}^{*} & a_{43}^{*} & \cdots & a_{m3}^{*} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1n}^{*} & a_{2n}^{*} & a_{11}^{*} & a_{11}^{*} & \cdots & a_{nm}^{*} \end{pmatrix} \right]$$

$$= \mathrm{Tr} \begin{pmatrix} \sum_{j=1}^{n} |a_{1j}|^2 & & & & & \\ & \sum_{j=1}^{n} |a_{2j}|^2 & & & & \\ & & \sum_{j=1}^{n} |a_{3j}|^2 & & & \\ & & & \sum_{j=1}^{n} |a_{4j}|^2 & & \\ & & & & \ddots & \\ & & & & & \sum_{j=1}^{n} |a_{mj}|^2 \end{pmatrix} = \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2$$

$$\mathrm{Tr}(A^{+}A)=\mathrm{Tr}\left[\begin{pmatrix} a_{11}^{*} & a_{21}^{*} & a_{31}^{*} & a_{41}^{*} & \cdots & a_{m1}^{*} \\ a_{12}^{*} & a_{22}^{*} & a_{32}^{*} & a_{42}^{*} & \cdots & a_{m2}^{*} \\ a_{13}^{*} & a_{23}^{*} & a_{33}^{*} & a_{43}^{*} & \cdots & a_{m3}^{*} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1n}^{*} & a_{2n}^{*} & a_{11}^{*} & a_{11}^{*} & \cdots & a_{nm}^{*} \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{1n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ a_{41} & a_{42} & a_{43} & \cdots & a_{4n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix}\right]$$

$$=\mathrm{Tr}\begin{pmatrix} \sum_{i=1}^{m}|a_{i1}|^2 & & & & & \\ & \sum_{i=1}^{m}|a_{i2}|^2 & & & & \\ & & \sum_{i=1}^{m}|a_{i3}|^2 & & & \\ & & & \sum_{i=1}^{m}|a_{i4}|^2 & & \\ & & & & \ddots & \\ & & & & & \sum_{i=1}^{m}|a_{in}|^2 \end{pmatrix}=\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2$$

当 $A=\boldsymbol{x}\in F^{m\times 1}$ 时，$\|A\|_{\mathrm{F}}=\|\boldsymbol{x}\|_2$,所以弗罗贝尼乌斯范数是向量范数 $\|\boldsymbol{x}\|_2$ 的一个自然推广.范数 $\|A\|_{\mathrm{F}}$ 实际上是把 A 看作 $F^{m\times n}$ 中的向量而建立的向量范数.我们可以把矩阵 A 写成行向量组成的矩阵

$$A=\begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix}$$

也可以写成是由列向量组成的矩阵

$$A=[B_1,B_2,\cdots,B_n]$$

那么，矩阵 A 的 F-范数的平方就是行向量的 2-范数的平方之和，也等于列向量的 2-范数的平方之和.

$$\|A\|_{\mathrm{F}}^2=\sum_{i=1}^{m}\|A_i\|_2^2=\sum_{j=1}^{n}\|B_j\|_2^2$$

由(7.3.7)定义的矩阵范数显然满足定义 1 的条件.下面我们来证明范数 $\|A\|_{\mathrm{F}}$ 满足相容性.

考虑两个矩阵 A 与 B.前者是 $m\times n$ 的，写成行向量组成的形式.后者是 $n\times k$ 的.它们的乘积 $C=AB$ 可以写成列向量的形式，$C=\begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_m \end{bmatrix}$，其中第 i 行的矩阵元是：$C_i=[c_{i1},c_{i2},\cdots,c_{ik}]$.向量 C_i 的 2-范数的平方是

$$\|C_i\|_2^2 |= \sum_{p=1}^{k}\left|c_{ip}\right|^2 = \sum_{p=1}^{k}\left|\sum_{j=1}^{n} a_{ij}b_{jp}\right|^2 \le \sum_{p=1}^{k}\sum_{j=1}^{n}\left|a_{ij}\right|^2\sum_{j=1}^{n}\left|b_{jp}\right|^2 = \sum_{j=1}^{n}\left|a_{ij}\right|^2\sum_{p=1}^{k}\sum_{j=1}^{n}\left|b_{jp}\right|^2 = \|A_i\|_F^2\|B\|_F^2$$

其中用到了赫尔德不等式.两边对 i 求和.这样就得到了 C 矩阵的 F 范数的平方：

$$\|AB\|_F^2 = \|C\|_F^2 = \sum_{i=1}^{m}\|C_i\|_2^2 \le \sum_{i=1}^{m}\|A_i\|_F^2\|B\|_F^2 = \|A\|_F^2\|B\|_F^2 \qquad (7.3.8)$$

对照(7.3.1)式，可知弗罗贝尼乌斯范数满足相容性条件.可以验证，$\|A\|_F$ 与 $\|\boldsymbol{x}\|_2$ 相容.

应该指出，尽管弗罗贝尼乌斯范数是 2-范数的一个推广，它不是由向量的范数诱导而得到的范数，因此不属于诱导范数.

7.3.2 矩阵的谱范数

(1) 矩阵的谱范数

由前一小节定义的矩阵范数 $\|A\|_2$ 称为谱范数，它是矩阵分析和系统理论中很有用的一种矩阵范数.由(7.3.2)式，谱范数可以写为

$$\|A\|_2 = \max_{\boldsymbol{x}\ne 0}\frac{\|A\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2}$$

它的几何意义是：变换后向量的长 $\|A\boldsymbol{x}\|_2$ 与原向量长 $\|\boldsymbol{x}\|_2$ 之比的最大值.

下面讨论谱范数的性质.

定理 4 设 $A\in C^{m\times n}$，则

(i) $\max\limits_{\|\boldsymbol{x}\|_2=\|\boldsymbol{y}\|_2=1}\left|\boldsymbol{y}^+ A\boldsymbol{x}\right| = \|A\|_2, \boldsymbol{x}\in F^n, \boldsymbol{y}\in F^m$.

(ii) $\left\|A^+\right\|_2 = \|A\|_2$

(iii) $\left\|A^+A\right\|_2 = \|A\|_2^2$

证明 (i) 对 $\|\boldsymbol{x}\|_2 = \|\boldsymbol{y}\|_2 = 1$，有

$$\max_{\|\boldsymbol{x}\|_2=\|\boldsymbol{y}\|_2=1}\left|\boldsymbol{y}^+A\boldsymbol{x}\right| \le \max_{\|\boldsymbol{x}\|_2=\|\boldsymbol{y}\|_2=1}\|\boldsymbol{y}\|_2\|A\boldsymbol{x}\|_2 = \max_{\|\boldsymbol{x}\|_2=1}\|A\boldsymbol{x}\|_2 = \|A\|_2$$

此处第一个不等号是利用了施瓦兹不等式,见第二章式(2.1.8).也可以将矩阵元写出，容易证明.

$$|\boldsymbol{p}^+\boldsymbol{q}| = \left(\sum_{i}^{n}|p_i^* q_i|\right)^{1/2} \le \left(\sum_{i}^{n}|p_i^*|^2\right)^{1/2}\left(\sum_{i}^{n}|q_i|^2\right)^{1/2} = |\boldsymbol{p}^+||\boldsymbol{q}| = |\boldsymbol{p}||\boldsymbol{q}|$$

现在我们来找不等式左边的最大值.为此，设 $\|\boldsymbol{x}\|_2 = 1$，令 $\boldsymbol{y} = A\boldsymbol{x}/\|A\boldsymbol{x}\|_2$，显然 $\|\boldsymbol{y}\|_2 = 1$，并且使 $\|A\boldsymbol{x}\|_2 = \|A\|_2 \ne 0$，就有

$$\max_{\|\boldsymbol{x}\|_2=\|\boldsymbol{y}\|_2=1}\left|\boldsymbol{y}^+A\boldsymbol{x}\right| = \max_{\|\boldsymbol{x}\|_2=1}\left|\frac{(A\boldsymbol{x})^+}{\|A\boldsymbol{x}\|_2}A\boldsymbol{x}\right| = \max_{\|\boldsymbol{x}\|_2=1}\frac{\|A\boldsymbol{x}\|_2^2}{\|A\boldsymbol{x}\|_2} = \max_{\|\boldsymbol{x}\|_2=1}\|A\boldsymbol{x}\|_2 = \|A\|_2 \quad (7.3.10)$$

(ii) $\|A\|_2=\max\limits_{\|\boldsymbol{x}\|_2=\|\boldsymbol{y}\|_2=1}|\boldsymbol{y}^+A\boldsymbol{x}|=\max\limits_{\|\boldsymbol{x}\|_2=\|\boldsymbol{y}\|_2=1}|\boldsymbol{x}^+A^+\boldsymbol{y}|=\|A^+\|_2$ (7.3.11)

其中用到性质(i).

(iii) 一般情况下，由范数相容性(7.3.6)，$\|A^+A\|_2\le\|A^+\|_2\|A\|_2, \|A^+\|_2=\|A\|_2$，知

$$\|A^+A\|_2\le\|A\|_2^2$$

但实际上，令$\|\boldsymbol{x}\|_2=1$，$\max\limits_{\|\boldsymbol{x}\|_2=1}\|A\boldsymbol{x}\|_2=\|A\|_2$，于是，根据性质(i)，

$$\|A^+A\|_2=\max_{\|\boldsymbol{x}\|_2=1}|\boldsymbol{x}^+A^+A\boldsymbol{x}|=\max_{\|\boldsymbol{x}\|_2=1}\|A\boldsymbol{x}\|_2^2=\|A\|_2^2 \qquad (7.3.12)$$

证明完毕.

定理 5 设$A\in C^{m\times n}$, $U\in C^{m\times m}, V\in C^{n\times n}$, 且$U^+U=I_m, V^+V=I_n$，则

$$\|UAV\|_2=\|A\|_2 \qquad (7.3.13)$$

证明 令$\boldsymbol{v}=V\boldsymbol{x}, \boldsymbol{u}=U^+\boldsymbol{y}$，则：$\|\boldsymbol{x}\|_2=1$当且仅当$\|\boldsymbol{v}\|_2=1$；因为

$$\|\boldsymbol{v}\|_2^2=\|V\boldsymbol{x}\|_2^2=|\boldsymbol{x}^+V^+V\boldsymbol{x}|=|\boldsymbol{x}^+\boldsymbol{x}|=\|\boldsymbol{x}\|_2^2$$

同理，$\|\boldsymbol{y}\|_2=1$当且仅当$\|\boldsymbol{u}\|_2=1$.于是，由性质(i)，

$$\|A\|_2=\max_{\|\boldsymbol{x}\|_2=\|\boldsymbol{y}\|_2=1}|\boldsymbol{y}^+A\boldsymbol{x}|=\max_{\|\boldsymbol{v}\|_2=\|\boldsymbol{u}\|_2=1}|\boldsymbol{u}^+UAV\boldsymbol{v}|=\|UAV\|_2$$

证明完毕.

定理 6 若$\|A\|_2<1$，则$I-A$为非奇异，且

$$\|(I-A)^{-1}\|_2\le\left(1-\|A\|_2\right)^{-1} \qquad (7.3.14)$$

证明 设 $\boldsymbol{x}$ 为任一非零向量，则

$$\|(I-A)\boldsymbol{x}\|_2=\|\boldsymbol{x}-A\boldsymbol{x}\|_2\ge\|\boldsymbol{x}\|_2-\|A\boldsymbol{x}\|_2\ge\|\boldsymbol{x}\|_2-\|A\|_2\|\boldsymbol{x}\|_2=(1-\|A\|_2)\|\boldsymbol{x}\|_2>0$$

于是，若$\boldsymbol{x}\ne0$，则$(I-A)\boldsymbol{x}\ne0$，设$(I-A)\boldsymbol{x}=\boldsymbol{y}$.解为

$$\boldsymbol{x}=(I-A)^{-1}\boldsymbol{y}$$

说明矩阵$I-A$非奇异.因为$I-A$非奇异，故有

$$I=(I-A)(I-A)^{-1}=(I-A)^{-1}-A(I-A)^{-1}$$

于是：$$(I-A)^{-1}=I+A(I-A)^{-1}$$

从而：$$\|(I-A)^{-1}\|_2\le\|I\|_2+\|A\|_2\|(I-A)^{-1}\|_2=1+\|A\|_2\|(I-A)^{-1}\|_2$$

即：$$\|(I-A)^{-1}\|_2\le\frac{1}{1-\|A\|_2}$$

证明完毕.

7.4.1 算子的范数

(1) 有界线性算子

我们已经在第二章 2.2 节中定义了算子的概念.其中定义 1 和 2 是对于任意的集合来说的.定义 3 的线性变换是在内积空间中定义的，只要把其中的内积空间换成赋范空间，就是定义了赋范空间中的线性变换.

对于任意一个算子 T，我们遇到需要一般地界定它的大小的问题.因而需要引入**算子范数**的概念.算子的范数，就是要用某种长度的概念来度量算子的大小.可是，对于一般的算子定义其范数是不容易的.下面我们只对有界线性算子来定义其范数.为此先定义有界线性算子的概念.

定义 1 设 V 和 U 是同一域 K 上的两个赋范线性空间，T 是 V 到 U 的算子，如果算子 T 满足以下两个条件：

(i) 对所有 $f,g\in V$，$T(f+g)=Tf+Tg$ (7.4.1)

(ii) 对所有 $f\in V$，数域 V 中所有纯量 $\alpha\in K$，$T(\alpha f)=\alpha Tf$ (7.4.2)

则称 T 为 V 到 U 的**线性算子**或者**线性变换**.(这两条性质已经在第二章的 2.2 节提到过了.)如果一个算子 T 满足(i)(ii)两条，还满足

(iii) 对 $f,g\in V$，与每个 $\varepsilon>0$，存在 $\delta>0$，使当

$\|f-g\|<\varepsilon$ 时，$\|Tf-Tg\|<\delta$， (7.4.3)

则称 T 为 V 到 U 的**连续线性算子**.条件(iii)刻画了算子 T 作用结果的连续性，简称为算子 T 的连续性.如果一个算子 T 满足(i)(ii)两条，还满足

(iv) 存在一个常数 M，使对于所有的

$f\in V$，$\|Tf\|\le M\|f\|$ (7.4.4)

则称 T 为 V 到 U 的**有界线性变换**或者**有界线性算子**.条件(iv)刻画了算子 T 作用结果的有界的性质，简称为算子 T 的有界性.

可以把有界线性算子的这个性质与函数的连续性做一个类比.算子 T 是作用在向量 f 上的.如果把向量看成是函数概念中的自变量，而算子看成是函数，则可写出 $Tf=T(f)$.函数的连续性要求，当自变量的变化充分小时，函数值的差能够任意小.有界线性算子正是具有这样的性质.当然，现在我们说到函数值的差任意小，实际上是指函数值的差的范数任意小.

定理 1 赋范空间 V 到赋范空间 U 的线性算子为连续的，其充分必要条件是：存在一个常数 M，使对于所有的 $f\in V$，$\|Tf\|\le M\|f\|$.

这一定理表明，对赋范空间之间的线性算子来说，条件(7.4.3)和(7.4.4)是互为充分必要的.或者说，有界性和连续性是等价的.因此以下我们可以只称有界算子或者连续算子.

例 1 对于线性积分算子 K，

$$Kf\equiv\int_a^b k(x,y)f(y)\mathrm{d}y \tag{7.4.5}$$

如果函数 f 和 k 是闭区间 $[a,b]$ 上的有界函数，那么，算子 K 就是**有界线性积分算子**或者**连续线性积分算子**.

令 $B(V,U)$ 是 V 到 U 的一切连续线性算子构成的集.对于任意两个元 $S,T\in B(V,U)$ 和数域 K 中的纯量α，令 $S+T$ 为由 $(S+T)f=Sf+Tf$ 确定的线性算子，αS 为由 $(\alpha S)f=\alpha Sf$ 确定的线性算子，则 $B(V,U)$ 关于这样定义的加法和纯量乘法成为 K 上的一个线性空间.即，有界线性变换的集合可以构成一个线性空间.在这个线性空间中的元都是连续线性算子.注意，与前面讲的由向量组成的空间不同.现在组成这个空间的元素都是线性变换.如果该空间中的元素都是可以定义范数的，就是赋范空间；如果是可以定义内积的，就是内积空间.

(2) 有界线性算子的范数

下面我们对于有界线性算子定义一个范数.

回顾矩阵范数的定义.矩阵的范数本身不容易定义.向量范数是容易定义的.因此，用矩阵先作用于向量，这样得到的是向量.再根据向量的范数的定义来定义矩阵的范数.对于算子，也采用这样的思路.先把算子作用在向量上.得到的是向量.再根据向量范数的定义，来定义算子的范数.

定义 2 设 V 是一个赋范向量空间，T 是作用于 V 上的一个有界线性变换，对于所有的 $f\in V$，满足

$$\|Tf\|\le M\|f\| \tag{7.4.6}$$

对于不同 f，M 的数值可能是不同的.我们把其中最小的 M 记为T_M，并将T_M 定义为 T 的**范数**，记作$\|T\|=T_M$.即如果有(7.4.6)，那么

$$\|T\|=T_M\le M \tag{7.4.7}$$

由这个定义

$$\|Tf\|\le\|T\|\|f\| \tag{7.4.8}$$

容易证明，由此定义的范数满足范数三公理.例如，

$$\|(T_1+T_2)f\|=\|T_1f+T_2f\|\le\|T_1f\|+\|T_2f\|\le\left(\|T_1\|+\|T_2\|\right)\|f\|$$

对于所有 f 都成立，而算子 (T_1+T_2) 的范数 $\|T_1+T_2\|=M$ 是满足 $\|(T_1+T_2)f\|\le M\|f\|$ 的最小的 M.故有 $M\le\|T_1\|+\|T_2\|$，即

$$\|T_1+T_2\|\le\|T_1\|+\|T_2\|$$

满足三角不等式.容易证明另外两条公理也满足.

有界线性算子的范数还有一个性质：由 (7.4.8)式，我们还得到两个有界线性变换T_1 和 T_2 乘积的范数与范数的乘积之间的不等式.把 T_1 作用在任何函数 f 上之后，将 T_1f 看做是一个函数.那么，反复运用(7.4.6)式，得到

$$\|T_2T_1f\|\le\|T_2\|\cdot\|T_1f\|\le\|T_2\|\cdot\|T_1\|\|f\| \tag{7.4.9a}$$

由此，乘积T_2T_1的范数一定是小于等于$\|T_2\|\cdot\|T_1\|$的.即

$$\|T_2T_1\|\le\|T_2\|\cdot\|T_1\| \tag{7.4.9b}$$

说明由此定义的算子的范数是满足相容性条件的.此式的一个特例是：

$$\|T^n\|\le\|T\|^n \tag{7.4.10}$$

由式(7.4.6)的定义，还是难于立即找到算子 T 的范数.但是，如果对于某一向量 f，有$\|Tf\|=m\|f\|$，那么$\|T\|\ge m$，这是因为$m\|f\|=\|Tf\|\le\|T\|\|f\|$.这对于我们确定有界线性算子的范数是有帮助的.

有了关于算子范数的定义，就可以比较明确地从数学上来表示线性算子的连续性了：当$\|f-g\|\to 0$，$\|Kf-Kg\|=\|K(f-g)\|\le\|K\|\|f-g\|\le M\|f-g\|\to 0$.

(3) 赋范线性变换空间的完备性

定义了有界线性算子的范数之后，从赋范线性空间 V 到 U 的一切有界线性变换集合构成的空间$B(V,U)$是一个赋范线性空间.那么这个空间是否完备的呢？我们有以下的定理.

定理 2 如果 V 和 U 是巴拿赫空间，从 V 到 U 的一切有界线性变换集合构成的空间$B(V,U)$关于定义 2 所定义的范数是一个巴拿赫空间.

我们把这一结论简单地说成：巴拿赫空间中一切有界线性变换的集合也是一个巴拿赫空间.

定义 3 设 V 和 U 是两个巴拿赫空间.如果存在 V 到 U 的有界线性变换 T，对于 $f\in V$，$u\in U$，有$Tf=u$，且$\|u\|=\|Tf\|=\|f\|$，则称这两个巴拿赫空间是**同构**的.V 和 U 同构记作$V\cong U$.

定理 3 当

$$\frac{1}{p}+\frac{1}{q}=1,\quad p\ge 1,\ (\text{当 } p=1 \text{时理解为} q=+\infty) \tag{7.4.11}$$

时

$$(L_q[a,b])^*\cong L_q[a,b] \tag{7.4.12}$$

即，在式(7.4.11)的条件下，p 次可积空间的对偶空间(对偶空间的概念见下)与 q 次可积空间是同构的.

我们定义巴拿赫空间中序列的收敛性的概念.

定义 4 设 V 是巴拿赫空间，$\{f_n\}$是 V 中的序列.如果存在 $f\in V$，使得

$$\lim_{n\to\infty}\|f_n-f\|=0 \tag{7.4.13}$$

则称$\{f_n\}$**强收敛**于 f.如果存在 $f\in V$，使对每一个$g\in V^*$，有

$$\lim_{n\to\infty}g(f_n)=g(f) \tag{7.4.14}$$

则称$\{f_n\}$**弱收敛**于 f.

我们可以把(7.4.13)(7.4.14)两式与(7.4.3)对照，式(7.4.13)对应于函数自变量的连续性，而(7.4.14)对应于函数的连续性.对照函数的连续性的有关理论，容易理解以下的定理.

定理 4 强收敛一定蕴含弱收敛，反之则不然.

进一步，我们定义有界线性算子的收敛性.

定义 5 设 V 是巴拿赫空间，$\{T_n\}$ 是 $B(V,V)$ 中的序列，如果存在 $T\in B(V,V)$，使得

$$\lim_{n\to\infty}\|T_n - T\| = 0，\text{或记为}\|T_n - T\|\to 0 \tag{7.4.15}$$

则称 $\{T_n\}$ **一致收敛**或者**依范数收敛**于 T.设存在 $T\in B(V,V)$，对于每一个 $f\in V$，若序列 $\{T_n f\}$ 在 V 中强收敛于 Tf，即

$$\lim_{n\to\infty}\|T_n f - Tf\| = 0，\text{或记为}\|T_n f\|\to\|Tf\| \tag{7.4.16}$$

则称 $\{T_n\}$ **强收敛**于 T.若序列 $\{T_n f\}$ 在 V 中弱收敛于 Tf，则称 $\{T_n\}$ **弱收敛**于 T.

注意，式(7.4.15)是与 f 无关的，即对于任意 f，此式都成立，才能叫做依范数收敛.如果把向量 f 看成是函数概念中的自变量，而算子看成是函数，$Tf = T(f)$.式(7.4.15)表示，无论“自变量 f”是什么，算子序列 $\{T_n\}$ 都依范数收敛于 T.这与第二章讲的函数的一致收敛的概念是类似的.而式(7.4.16)是算子作用到某一具体 f 上之后的效果.这与第二章讲的函数的逐点收敛的概念是类似的.可以说，依范数收敛和强收敛的概念，是把函数的一致收敛和逐点收敛的概念推广到巴拿赫空间.在函数空间中，收敛性用绝对值来判别.在巴拿赫空间中，收敛性用范数来判别.

既然函数的一致收敛性比逐点收敛是更强的要求，容易理解，依范数收敛是比强收敛更强的要求.因此有以下定理.

定理 5 对于 X 到 X 的算子序列，一致收敛蕴含强收敛，强收敛蕴含弱收敛.，但反之不然.

后面我们将以投影算子为例说明，依范数收敛是比强收敛更强的要求.

7.4.2 伴随算子

算子作用在原像上，得到的像也称为算子的值.算子的值是某种空间中的元素.

定义 6 算子的值为实数或者复数时，称这样的算子为**泛函**.巴拿赫空间 V 到 U 的有界线性变换，且当空间 U 为实数或复数空间时，称为 V 上的**连续线性泛函**或者**有界线性泛函**.V 上的一切连续线性泛函构成的巴拿赫空间 $B(V,U)$，称为 V 的**对偶空间**或者**共轭空间**，记作 V^* 或者 V'.

我们可以从这样几个例子来理解对偶空间：对于 n 维欧几里德空间(n 维实列向量空间)E，若列向量 $\boldsymbol{y}\in E$，那么一切 $\boldsymbol{x}=\boldsymbol{y}^{\mathrm{T}}$，即所有 $\boldsymbol{y}$ 的转置得到的行向量构成的 n 维实行向量空间 E^* 就是 E 的对偶空间；对于 n 维复列向量空间 U，若列向量 $\boldsymbol{y}\in U$，那么一切 $\boldsymbol{x}=\boldsymbol{y}^{+}$，即所有 $\boldsymbol{y}$ 的转置共轭得到的行向量构成的 n 维复行向量空间 U^* 就是 U 的对偶空间；对于 $m\times n$ 维实矩

阵空间 $R_{m\times n}$，若矩阵 $A\in R_{m\times n}$，那么一切 $B=A^{\mathrm{T}}$，即所有 A 的转置得到的矩阵构成的空间 $R_{n\times m}$ 就是 $R_{m\times n}$ 的对偶空间 $R_{m\times n}^{*}$；对于 $m\times n$ 维复矩阵空间 $C_{m\times n}$，若矩阵 $A\in C_{m\times n}$，那么一切 $B=A^{+}$，即所有 A 的转置共轭得到的矩阵构成的空间 $C_{n\times m}$ 就是 $C_{m\times n}$ 的对偶空间 $C_{m\times n}^{*}$.从一个空间得到它的对偶空间的过程中，我们实际上对其中的每一个向量都进行了一个操作，也就是进行了一个线性变换.

容易验证，对于上述的这几个空间及对偶空间，按照 7.3.1 小节定义了诱导范数之后，符合定理 3.

有了对偶空间的这个概念，我们可以进一步来理解内积的定义 (g,f)：它是在空间 V 中的一个向量 f 与在 V 的共轭空间 V^{*} 中的一个向量 g 共同以某种方式构成的一个纯量.

在第二章定义伴随算子的时候，是根据内积来定义.现在我们要在赋范空间内给予更为一般的定义. 因为在一个巴拿赫空间中，可能还没有定义内积.

由定义 4，巴拿赫空间 V 到 U 有界线性变换，称为 V 上的连续线性泛函或者有界线性泛函.有界连续算子 T 作用在元 f 上，记为 $T(f)$，它的效果，记为 $F(f)$.$T(f)$ 和 $F(f)$ 的差别是：当我们强调算子的作用时，写成 $T(f)$，这是一个算子.当我们强调算子作用的效果时，写成 $F(f)$，相当于是 f 的函数.对于 $F(f)$，可以将元 f 看成是自变量，在希尔伯特空间中，元 f 本身可以是一个函数.因此，$F(f)$ 正是第一章中讲过的泛函的含义.

当我们讲到通过有界线性算子 T 的作用来产生一个泛函的时候，并没有定义泛函的具体的形式.内积可以是泛函的一种形式. 在第五章中已经用到了这种形式，见(5.1.6)式.

定理 6 对赋范空间 X 上的每个连续线性泛函 F，存在唯一的 $g_T\in X$，使对每个 $f\in X$，有 $F(f)=(f,g_T)$，并且 $\|F\|=\|g_T\|$.

注意，其中的 (f,g_T) 表示泛函，内积是其特例.当然，因为内积的具体形式相对简单，在具体举例时往往采用内积的形式.

此定理表明，一个赋范空间上的每一个连续线性泛函 F，它作用在元 f 上的效果，就是 f 与这个空间中一个特定的元 g_T 形成的泛函，对于每一个 f 都是如此.并且，F 的范数与 g_T 的范数相等.

定义 7 设 X 是希尔伯特空间，T 是 X 到 X 的连续线性算子.对于每一个固定的 $g\in V$，作 Tf 与 g 的泛函，$f\in X$，将结果记为 $F_g(f)$，

$$F_g(f)=(Tf,g)$$

由此定义的$F_g(f)$是X上的连续线性泛函.由定理 6，存在$u\in X$，使得对一切$f\in X$，有$F_g(f)=(f,u)$.现在令$u=T^*g$，则T^*也是X到X的连续线性算子.称T^*是T的**伴随算子**或**共轭算子**.联系T与T^*的基本关系是：对任何$f,g\in X$，有

$$(Tf,g)=(f,T^*g) \tag{7.4.17}$$

有了前面关于共轭空间的概念，我们对于伴随算子就比较容易理解了.式(7.4.17)显然表示了，如果向量f属于空间V，那么向量g是属于对偶空间V^*的.一个向量f被线性算子T作用之后得到的向量，与g做的泛函，等于f与伴随算子T^*作用在g上之后得到的向量做的泛函.可见，若算子T作用于空间V中的向量上，那么伴随算子T^*是作用在对偶空间的向量上的.正因为如此，对偶空间又称共轭空间，伴随算子又称共轭算子.

例 2 对于n维欧几里得空间E，列向量用x表示，共轭空间内的行向量x^{T}组成的空间E^*.式(7.4.17)在本例中的表示就是

$$(Bx)^{\mathrm{T}}y=x^{\mathrm{T}}B^{\mathrm{T}}y \tag{7.4.18}$$

例 3 对于n维酉空间U，酉空间U的共轭空间U^*由所有行向量x^+所组成.式(7.4.17)在本例中的表示就是

$$(Ax)^+y=x^+A^+y \tag{7.4.19}$$

例 2 和例 3 与 2.2 节中的例 2 和例 3 一样.只是这儿增加了对共轭空间的理解.

定理 7 设X是希尔伯特空间，T是X到X的连续线性算子.那么，$\|T^*\|=\|T\|$.

证明 将向量的内积定义为范数的平方，利用范数的相容性，有

$$\|Tx\|^2=(Tx,Tx)=(T^*Tx,x)\le\|T^*Tx\|\|x\|\le\|T^*\|\|Tx\|\|x\|$$

其中第一个不等号是用了施瓦兹不等式.这就得到，$\|Tx\|\le\|T^*\|\|x\|$.又由(7.4.8)式，$\|Tx\|\le\|T\|\|x\|$.可见$\|T^*\|\ge\|T\|$.将T换成T^*， 则得到$\|T^*x\|\le\|T\|\|x\|$和$\|T^*x\|\le\|T^*\|\|x\|$，则可得$\|T^*\|\le\|T\|$，因此，必有$\|T^*\|=\|T\|$.

此定理表明，连续线性算子T与它的伴随算子的范数总是相等的.例如，幺正矩阵与它的转置矩阵的范数是相等的，酉矩阵与它的转置共轭矩阵的范数是相等的.

定义 8 设X是希尔伯特空间，T是X到X的连续线性算子.如果对于任何$x\in X$，有

$$\|Tx\|=\|x\| \tag{7.4.20}$$

则称T是**保范变换**.算子T就称为**保范算子**，或者称算子T是**保范**的.

显然，保范变换这一定义是对第二章中等距变换定义的扩展.虽然(7.4.20)与(2.2.31)形式一样，现在是在范数的意义上定义的.总之，保范变换保持向量的范数不变.

保范变换的定义式(7.4.20)蕴含以下两个等式.

$$T^*T=TT^*=I \tag{7.4.21}$$

$$(Tf, Tg) = (f, g) \tag{7.4.22}$$

证明：定义一个向量与自身的内积的根号作为这个向量的范数.那么

$$\|Tx\|^2 = (Tx, Tx) = (T^*Tx, x) = \|x\|^2 = (x, x)$$

因此必有(7.4.21).另根据自伴算子的定义，易得$(Tf, Tg) = (T^*Tf, g) = (f, g)$，这就是(7.4.22)式.

式(7.4.20)-(7.4.22)是两两相互等价.

对于逆算子，正规算子，自伴算子等概念的定义与与第二章中的一样，只要把内积理解为泛函即可.此处不再重复.

由(7.4.21)式看到，保范算子一定存在逆.逆算子就是伴随算子，且也是保范的.

定理 8 设 T 是希尔伯特空间 X 到自身的连续线性算子，则 T 是正规算子的充分必要条件是，对每一个 $x \in X$，有$\|T^*x\| = \|Tx\|$.

7.4.3 投影算子

定义 9 对 V 中任一正交归一的向量集合$\{\phi_i\}$，算子 P_n 作用到 V 中任一函数 f 上的效果是

$$P_n f \equiv \sum_{i=1}^{n} (\phi_i, f)\phi_i \,. \tag{7.4.23}$$

那么，称 P_n 是投影到由集合$\{\phi_i\}$的前 n 个元组成的 V 的子空间上的**投影算子**.

若这一正交向量集合的维数是 M 那么，向量 f 可用这一组完备基展开如下

$$f \equiv \sum_{i=1}^{M} (\phi_i, f)\phi_i$$

恒等算子的作用效果如下。

$$If \equiv \sum_{i=1}^{M} (\phi_i, f)\phi_i \tag{7.4.23a}$$

定义 10 一个算子 K 如果满足条件

$$K^2 = K \tag{7.4.24}$$

则称算子 K 是**等幂**的，或者称 K 具有**幂等性**.

关于投影算子有以下定理.

定理 9 希尔伯特空间 X 到自身的连续线性算子 T 为投影算子的充分必要条件是 T 为自伴算子且是等幂的$T^2 = TT = T$.

这一定理说明了投影算子的两个性质：

(i) P_n 是厄米算子.由定义式(7.4.23)显而易见，对于所有 $f, g \in H$，$(f, P_n g) = (P_n f, g)$ .P_n 是厄米算子.

(ii) P_n 是等幂的.

命题 1 投影算子 P_n 的范数是 1.

证明 首先证明投影算子 P_n 的范数满足

$$\|P_n\| \le 1 \tag{7.4.25}$$

为此考察 P_n 的范数

$$\|P_n f\| = (P_n f, P_n f)^{1/2} = (f, P_n f)^{1/2} = \left[\sum_{i=1}^{n}(f,\phi_i)(\phi_i, f)\right]^{1/2} \le \|f\| \tag{7.4.26}$$

最后用到贝塞尔不等式.对于所有 f，$\|P_n f\| \le \|f\|$，因此 $\|P_n\| \le 1$.

对于算子 $(I - P_n)$，同样可证

$$\|I - P_n\| \le 1. \tag{7.4.27}$$

下面进一步证明，(7.4.25)和(7.4.27)中的不等号可以去掉，只剩下等号.证明如下.至少存在一个 f，对于它满足 $\|P_n f\| = \|f\|$，事实上，每当 $f = \phi_1, \phi_2, \cdots, \phi_n$ 时都有 $\|P_n f\| = \|f\|$ ，例如当 $f = \phi_i, 1 \le i \le n$ 时，

$$P_n f = \sum_{j=1}^{n}(\phi_j, \phi_i)\phi_j = \sum_{j=1}^{n}\delta_{ij}\phi_j = \phi_i = f$$

可见，至少有一个 f，使得 $\|P_n f\| = \|f\|$.同时应该有 $\|f\| = \|P_n f\| \le \|P_n\|\|f\|$.可见，此时有 $\|P_n\| \ge 1$.结合(7.4.25)式可得

$$\|P_n\| = 1 \tag{7.4.28}$$

类似地，

$$\|I - P_n\| = 1 \tag{7.4.29}$$

结论，投影算子的范数是 1.**证明完毕**.

当 $n \to \infty$ 时，(7.4.26)中的不等号可写成等号.这就是完备性关系.由(7.4.23)式，可以用投影算子来表示这一完备性关系.

$$\lim_{n\to\infty}(f, P_n f) = \sum_{i=1}^{\infty}(f, \phi_n)(\phi_n, f) = (f, f) \tag{7.4.30}$$

命题 2 投影算子序列 $\{P_n\}$ 强收敛于单位算子 I.

证明 **利用**(7.4.23)式和投影算子的等幂性，

$$\|P_n f - If\| = (P_n f - f, P_n f - f)^{1/2}$$
$$= [(P_n f, P_n f) - (P_n f, f) - (f, P_n f) + (f, f)]^{1/2}$$
$$= [(f, P_n f) - 2(f, P_n f) + (f, f)]^{1/2} = [(f, f) - (f, P_n f)]^{1/2}$$

由(7.4.30)式知，

$$\lim_{n\to\infty} \| P_n f - If \| = 0$$

对照(7.4.16)式知，序列 $\{P_n\}$ 强收敛于 I.**证明完毕**.

另一种证明方法是，因为对于所有 $f\in V$，有(7.4.23)式，由于

$$\|P_n f\|=(P_n f,P_n f)^{1/2}=(f,P_n^2 f)^{1/2}=(f,P_n f)^{1/2}$$

所以$\lim\limits_{n\to\infty}\|P_n f\|=(f,f)^{1/2}=\|If\|$，因此 P_n 趋于 *I*.并且，按照(7.4.16)式知，序列$\{P_n\}$强收敛于 *I*.

命题 3 投影算子序列$\{P_n\}$并不依范数收敛于 *I*.

证明 如果序列$\{P_n\}$按范数收敛于 *I* 时，那么，给定任意ε，必存在一个 *N*，使只要$n\geq N$，对于所有 $f\in V$ 就都有$\|(P_n-I)f\|\leq\varepsilon\|f\|$，即 *N* 与 *f* 是无关的.但是，事实上，对于任何给定的 *N*，总可以选出某些函数$h\in V$，例如$h=\phi_{N+1}$，使得$(P_N-I)h=(P_N-I)\phi_{N+1}=-\phi_{N+1}$.

$$(P_N-I)h\equiv\sum_{i=1}^{N}(\phi_i,h)\phi_i-\sum_{i=1}^{\infty}(\phi_i,h)\phi_i=-\sum_{i=N+1}^{\infty}(\phi_i,h)\phi_i$$

$$(P_N-I)\phi_{N+1}=-\sum_{i=N+1}^{\infty}(\phi_i,\phi_{N+1})\phi_i=-\phi_{N+1}$$

这时，$\|(P_N-I)h\|=\|h\|$，所以序列$\{P_n\}$并不按范数收敛于 *I*.**证明完毕**.

以上两个命题说，P_n 的范数是 1，但是序列$\{P_n\}$并不按范数收敛于 *I*.从表面上看来，这两个命题似乎是矛盾的.既然这个算子本身的范数就是 1 了，为什么还不能按范数收敛于单位算符？

实际上这两个命题不矛盾.为了解释这一点，我们看范数的三角不等式：

$$\|P_n\|-\|I\|\leq\|P_n-I\|$$

左边等于零并不能推出右边趋于零.事实上由(7.4.29)式，右边是 1.这个问题的关键在于依范数收敛指的是$\|P_n-I\|\to 0$而不是$\|P_n\|\to\|I\|$.

由此可见，按范数收敛是比强收敛更强的要求.

设在空间 X 中，投影算子 P_1 将 X 中的向量投影到 X 的子空间 X_1，投影算子 P_2 将 X 中的向量投影到 X 的子空间 X_2.为明确起见，我们将这两个投影算子分别记为 P_{X_1} 和 P_{X_2}.

定义 11 设 P_{X_1} 和 P_{X_2} 是同一空间中的两个投影算子.若 $P_{X_1}P_{X_2}=\theta$，则称 P_{X_1} 和 P_{X_2} 是**正交**的.若 $P_{X_1}P_{X_2}=P_{X_2}$，则称 P_{X_2} 是 P_{X_1} 的**部分**.

此处的θ既表示零向量，也表示零算子.零算子作用在任何向量上得到零向量.由于投影算子是自伴的，当 $P_{X_1}P_{X_2}=\theta$ 时，必有 $P_{X_1}P_{X_2}=(P_{X_2}P_{X_1})^{\dagger}=\theta$.

由 $P_{X_1}P_{X_2}=P_{X_2}$，必有 $P_{X_2}P_{X_1}=P_{X_2}$.

P_{X_1} 和 P_{X_2} 这两个投影算子的加、积与差之后仍为投影算子的充要条件如下.

定理 10 (i) P_{X_1} 与 P_{X_2} 之和为投影算子的充要条件是 P_{X_1} 与 P_{X_2} 正交.此时

$$P_{X_1} + P_{X_2} = P_{X_1+X_2}$$

投影算子 $P_{X_1+X_2}$ 将 X 中的向量投影到 X 的子空间 X_1 与 X_2 的直和空间.

(ii) P_{X_1} 与 P_{X_2} 之积为投影算子的充要条件是 P_{X_1} 与 P_{X_2} 可交换.此时

$$P_{X_1} P_{X_2} = P_{X_1 \cap X_2}$$

投影算子 $P_{X_1 \cap X_2}$ 将 X 中的向量投影到 X 的子空间 X_1 与 X_2 的交集.

(iii) P_{X_1} 与 P_{X_2} 之差 $P_{X_1} - P_{X_2}$ 为投影算子的充要条件是 P_{X_2} 是 P_{X_1} 的部分.此时投影算子 $P_{X_1} - P_{X_2}$ 将 X 中的向量投影到子空间 X_1 中去掉 X_2 之后剩下的子空间.

本章重点

范数的定义。巴拿赫空间的定义。
赫尔德不等式和闵可夫斯基不等式。
向量范数、矩阵范数、线性算子范数的定义。
向量和矩阵的三种最常用的范数，1、2、∞范数的定义。
连续线性算子的概念。
投影算子的概念。幂等性。

小贴士

本章提到，柯西不等式是赫尔德不等式的特例。第二章中 2.1.2 小节例 17 介绍的是：柯西不等式是施瓦兹不等式(2.1.8)的特例。请判断，施瓦兹不等式是否是赫尔德不等式的特例。在这个意义上，是否可以把赫尔德不等式看做是施瓦兹不等式的一个推广？如果是，那么(2.1.9)式是否可以做相应的推广？

布置习题：
做以下习题 8,10,12，15，24，25，此外，再任选 5 题。

习题

1. 证明(7.1.7)和(7.1.8)两式成立，即从等式右边可以得到左边.提示，这两式定义本身说明：$(f, f) = \|f\|^2$

2. 证明：(7.1.22)式定义的范数满足范数三公理.

3. 积分形式的赫尔德不等式的一种证明.

(1) 证明，当$0<\alpha<1$时，函数$f(z)=z^{\alpha}-\alpha z-\beta$在$z=1$取极大值，并在$\beta=1-\alpha$时这个极大值为零.因此当$\alpha<1$和$\beta=1-\alpha$时，$z^{\alpha}\le\alpha z+\beta$.

(2) 进行变量置换$z=x/y$，证明，当$\alpha\le1$和$\beta=1-\alpha$时，$x^{\alpha}y^{\beta}\le\alpha x+\beta y$.

(3) 现在设$|f(x)|^p$和$|g(x)|^q$都是可积函数，$1/p+1/q=1$.利用 b)的结论证明

$|fg|\le\frac{1}{p}|f|^p+\frac{1}{q}|g|^q$，所以$|fg|$也是可积的.

(4) 假如$|f(x)|^p$和$|g(x)|^q$都是可积的，那么利用 c) 的结论证明，

$$|\int f(x)g(x)\mathrm{d}x|\le[\int|f(x)|^p\,\mathrm{d}x]^{1/p}[\int|g(x)|^q\,\mathrm{d}x]^{1/q}$$

4. 积分形式的闵可夫斯基不等式的一种证明.

(1) 证明$|f+g|^p\le2^p[|f|^p+|g|^p]$，因此,假如$f$和$g$都是$p$次可积的,那么$f+g$也是$p$次可积的.

(2) 证明，对于$p>1$和$1/p+1/q=1$，

$$\int|f(x)+g(x)|^p\,\mathrm{d}x\le\int|f(x)||f(x)+g(x)|^{p-1}\mathrm{d}x+\int|g(x)||f(x)+g(x)|^{p-1}\,\mathrm{d}x$$

如果函数f和g都是p次可积的，属于L_p，那么函数$|f+g|^{p-1}$就是q次可积的，属于L_q.

(3) 将赫尔德不等式用于 b)中两边的积分上，证明闵可夫斯基不等式.

5. 加权赫尔德不等式.

(1) 在第二章 2.1.2 小节中，定义了加权内积的概念.现在假如在n维向量空间中有$\gamma_i,(i=1,2,\cdots,n)$是一组给定的正数，将它作为权.考虑如何将(7.1.14)式作适当的修改，并重新定义u和v，以此加权赫尔德不等式：

$$\sum_{k=1}^{n}\gamma_k a_k b_k<(\sum_{k=1}^{n}\gamma_k a_k^p)^{1/p}(\sum_{k=1}^{n}\gamma_k b_k^q)^{1/q}$$

(2) 在第二章 2.1.2 小节中的例 14 还引入了积分形式的加权内积的概念.现在假如在函数空间中有权函数$\rho(x)\ge0$，证明如下的积分形式的加权赫尔德不等式：

$$\int f(x)g(x)\rho(x)\mathrm{d}x\le(\int f^p(x)\rho(x)\mathrm{d}x)^{1/p}(\int g^q(x)\rho(x)\mathrm{d}x)^{1/q}$$

6. 证明加权闵可夫斯基不等式

$$\left(\sum_{i=1}^{n}|a_i+b_i|^p\gamma_i\right)^{1/p}\le\left(\sum_{i=1}^{n}\gamma_i|a_i|^p\right)^{1/p}+\left(\sum_{i=1}^{n}\gamma_i|b_i|^p\right)^{1/p}$$

其积分形式为

$$[\int|f(x)+g(x)|^p\ \rho(x)\,\mathrm{d}x\ ]^{1/p}\le[\int|f(x)|^p\ \rho(x)\,\mathrm{d}x\ ]^{1/p}+[\int|g(x)|^p\ \rho(x)\,\mathrm{d}x\ ]^{1/p}$$

7. 证明：(1) $\|\boldsymbol{x}-\boldsymbol{y}\|\ge\|\boldsymbol{x}\|-\|\boldsymbol{y}\|$

(2) $\|A-B\| \geq \|A\|-\|B\|$

8. 设 $A=(a_{ij})_{n\times n}$ 是正定对称矩阵，证明：不论 $x_1,x_2,\cdots,x_n$ 和 $y_1,y_2,\cdots,y_n$ 是何实数，恒有不等式

$$\left(\sum a_{ij}x_i y_j\right)^2 \leq \left(\sum a_{ij}x_i x_j\right)\left(\sum a_{ij}y_i y_j\right)$$

9. 对于由 n 个分量组成的向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$，证明如下等价关系.

(1) $\|\boldsymbol{x}\|_\infty \leq \|\boldsymbol{x}\|_1 \leq n\|\boldsymbol{x}\|_\infty$

(2) $\dfrac{1}{\sqrt{n}}\|\boldsymbol{x}\|_2 \leq \|\boldsymbol{x}\|_\infty \leq n\|\boldsymbol{x}\|_2$

(3) $\|\boldsymbol{x}\|_2 \leq \|\boldsymbol{x}\|_1 \leq \sqrt{n}\|\boldsymbol{x}\|_2$

10. 设 $f(x)=(x-1/2)^3$， $x\in[0,1]$，求 $\|f\|_1,\|f\|_2,\|f\|_\infty$.

11. 设 $A=(a_{ij})_{n\times n}$， $\|A\|$ 是向量范数诱导的矩阵范数.证明 $\|A\| \geq |a_{ij}|$.

12. 设 $\|A\|$ 是诱导范数， $\det A \neq 0$.证明

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

(2) $\|A^{-1}\|^{-1} = \min\limits_{\boldsymbol{x}\neq 0}\dfrac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|}$

13. 验证，$\|A\|_{\mathrm{F}}$ 与 $\|\boldsymbol{x}\|_2$ 相容.

14. 证明 $\|A\|_2 \leq \|A\|_{\mathrm{F}}$

15. 计算下列矩阵的行范数，列范数，谱范数和 F 范数.

(1) $A=\begin{pmatrix} 3 & -1 & 1 \\ 1 & 1 & 1 \\ 2 & 1 & -1 \end{pmatrix}$

(2) $A=\begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix}$，$a$ 是实数.

16. 设 $A=\begin{pmatrix} 1 & 2 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$，求 $\rho(A)$.

21. 若 k 是负实数，则 7.3.3 小节中的命题 4 应做如何修改？

22. 证明巴拿赫空间上有界线性变换的连续性与线性性质.

(1) 设F是巴拿赫空间B上的有界线性泛函.假设在B内有一个序列$\{f_n\}$,它按范数趋于$f\in B$.

证明$F(f_n)\to F(f)$，即F是连续的.

(2) 设$f\in L_p$和$1/p+1/q=1$，那么在L_q上定义为

$$F(g)=(f,g)=\int f^*(x)g(x)\mathrm{d}x, g\in L_q$$

的泛函是有界线性泛函.

23. 在 7.4.1 小节中已经定义了有界线性泛函的范数.

(1) 证明，对于在L_q上按照$F(g)=(f,g)=\int f^*(x)g(x)\mathrm{d}x$定义的泛函$F$，其中$f\in L_p$，

$1/p+1/q=1$，有$\|F\|\le\|f\|$，这里f的范数是相应于L_p的，即

$$\|f\|=[\int|f(x)|^p\mathrm{d}x]^{1/p}$$

(2) 证明，假如$f\in L_p$，那么$g=\dfrac{1}{f^*}|f|^{p/q+1}$属于$L_q(1/p+1/q=1)$，证明 i)

$|F(g)|=|(f,g)|=\|f\|^p$，ii) $\|f\|\|g\|=\|f\|^p$.因此$|F(g)|=\|f\|\|g\|$，因此$\|F\|\ge\|f\|$.

将(1)和(2)结合起来，得到$\|F\|=\|f\|$.

24. 由投影算子的定义式(7.4.23)证明投影算子是自伴的和等幂的.

25. 证明 7.4 节中的定理 10.并 证明：若$P_{X_2}P_{X_1}=P_{X_1}P_{X_2}$，那么$P_{X_1}+P_{X_2}-P_{X_1}P_{X_2}$是投影算子，

这个投影算子将X中的向量投影到X的一个什么子空间？