

第二章 向量与矩阵的范数

本章将会介绍范数这一概念，主要内容包括介绍向量的范数、矩阵的范数和算子范数，同时在本章的最后会介绍范数的相关应用——借助范数来判断一个矩阵的逆矩阵或求解线性方程组因摄动带来的误差是否处于一个可接受的范围内。

为什么要引入范数？在本章中应时刻牢记范数是一个数而非其他之物，将矩阵、向量类的抽象概念转换成一个“大小”往往能够简化问题。

本章将会涉及到以下内容：

- 向量的范数
- 矩阵的范数
- 算子范数
- 范数的应用

注意：本章内的内容具有强连续性，需要按照顺序从头开始了解，同时，本章内容会涉及多个定理、性质的证明，需要静下心来仔细复盘证明方法，但本章内容与第一章的联系不大，因此如果觉得第一章理解起来有难度可以先跳过第一章的内容继续学习第二章。

2.1 向量的范数

本节我们首先介绍向量的范数，在开始详细介绍向量的范数前我们可以先简单“感受”一下何为范数。

2.1.1 范数的直观感受

在第一章，我们了解过向量的长度：

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \mathbf{x}^H \mathbf{x}$$

有了向量的长度，向量便在实际空间内有了意义，也给后续定义距离等概念打下了基础。

其实，向量的距离就是向量的 2-范数，范数的种类有很多，每一种范数的计算方法都不一样但大同小异，这里利用向量 2-范数作说明是因为向量的 2-范数便于理解（距离），同时也引出下面的对范数的意义的概括：这便是范数的具体意义——对“距离”概念的延伸，是一种广义的“长度”。

为什么说是广义的“长度”呢？继续阅读就会发现，向量范数并不止向量的 2-范数，还有其他类型的向量范数。

定义 2.1.1: 向量范数的定义

设映射 $\|\cdot\|: \mathbb{C}^n \rightarrow \mathbb{R}$ 满足：

1. 正定条件： $\|\mathbf{x}\| \geq 0$, 当且仅当 $\mathbf{x} = \mathbf{0}$ 时 $\|\mathbf{x}\| = 0$
2. 齐次条件： $\|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\|, \lambda \in \mathbb{C}, \mathbf{x} \in \mathbb{C}^n$
3. 三角不等式： $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|, \forall \mathbf{x}, \mathbf{y} \in \mathbb{C}^n$

则称映射 $\|\cdot\|$ 为 \mathbb{C}^n 上向量 \mathbf{x} 的范数，定义了范数的 \mathbb{C}^n 又叫做一个线性赋范空间。

定义1.1.1给出了能够成为向量范数的条件：是一个映射，并且能够满足三个条件，就可以将这种映射叫做向量的范数，这也是今后判断给出一种映射是否为向量范数的方法。

2.1.2 向量范数的性质

定义1.1.1中给出的能够成为向量范数的三个条件也是向量范数的三个性质，除此之外，根据上面的定义，我们还可以得出向量范数的下列性质：

1. 零向量的范数是 0

这一条性质不难理解，直接使用向量范数的正定条件即可得出。

2. 当 $\mathbf{x} \neq \mathbf{0}$ 时, $\|\frac{1}{\|\mathbf{x}\|}\mathbf{x}\| = 1$

这一条性质通过向量范数的齐次条件可以得出：向量范数是一个数，故 $\|\frac{1}{\|\mathbf{x}\|}\mathbf{x}\| = \frac{1}{\|\mathbf{x}\|}\|\mathbf{x}\| = 1$

3. 对任意 $\mathbf{x} \in \mathbb{C}^n$ ，有 $\|-\mathbf{x}\| = \|\mathbf{x}\|$

同样地，利用齐次性依旧可以得到这一条性质： $\|-\mathbf{x}\| = |-1| \cdot \|\mathbf{x}\| = \|\mathbf{x}\|$

4. 对任意 $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ ，有 $||\mathbf{x}\| - \|\mathbf{y}\|| \leq \|\mathbf{x} - \mathbf{y}\|$

上述不等式的证明过程如下：

证明.

$$\|\mathbf{x}\| = \|\mathbf{x} - \mathbf{y} + \mathbf{y}\| \leq \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y}\| \text{ (范数的三角不等式)}$$

从而

$$\|\mathbf{x} - \mathbf{y}\| = \|\mathbf{y} - \mathbf{x}\| \geq \|\mathbf{y}\| - \|\mathbf{x}\| \text{ (依旧是范数的三角不等式)}$$

那么

$$\|\mathbf{x}\| - \|\mathbf{y}\| \geq -\|\mathbf{x} - \mathbf{y}\| \text{ (不等式左右两边同时乘-1, 不等式方向改变)}$$

于是就有了

$$||\mathbf{x}\| - \|\mathbf{y}\|| \leq \|\mathbf{x} - \mathbf{y}\|$$

证毕

□

2.1.3 对向量范数三大性质的理解

在1.1.2节中列出的向量范数的性质，那如何理解向量的范数三大性质呢？

对于正定性的理解：可以从 2-范数来看，一般来说，我们认为 2-范数表示的是向量的距离，即

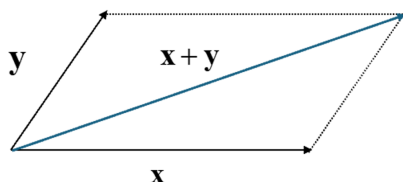
$$\|\mathbf{x}\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

(这里 x_1, x_2, \cdots, x_n 代表向量 \mathbf{x} 的各个分量), 在后面介绍到 p-范数的时候也可以从其式子来看出非负性。

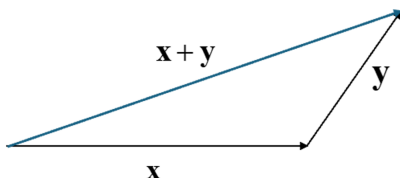
对于齐次性的理解：仍以二维空间上的 2-范数为例，对于向量 $\mathbf{x}, \lambda\mathbf{x}$ 我们可以将其理解为对该向量进行伸缩变换，当 λ 为正时，新的伸缩变换之后的向量与原向量方向相同，当 λ

为负时，新的伸缩变换之后的向量与原向量方向相反，由于 2-范数描述的为距离，所以 $|\lambda||x|$ 就代表将向量 x 的距离乘以缩放因子的绝对值（因为范数是个数，所以忽略掉正负带来的影响，因此加绝对值）。

对于三角不等式的理解：依旧以二维空间上的 2-范数为例，依据向量相加的平行四边形法则：我们可以画出 $x, y, x + y$ 的关系：



如果我们将向量 y 平移一下，就会得到下面这幅图：



这就构成了一个三角形，依据三角形两边之和大于第三边的性质，从而可以轻易理解范数性质 3——三角不等式的含义。

2.1.4 常见的向量范数

本节介绍整个向量范数的重点——常见的向量范数经常使用的向量范数有下面三个：

$$\begin{aligned}\|x\|_1 &= \sum_{i=1}^n |x_i| \\ \|x\|_2 &= \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2} \\ \|x\|_\infty &= \max_{1 \leq i \leq n} |x_i|\end{aligned}$$

他们分别为向量 1-范数，向量 2-范数和向量无穷范数，接下来将会一一给出这三种范数是向量范数的证明。

向量 1-范数的证明与 Minkovski 不等式

正定性证明:

若向量 $\mathbf{x} \neq \mathbf{0}$, 则表明 \mathbf{x} 至少有一个分量 $x_k \neq 0$, 从而 $|x_k| > 0$, 故 $\|\mathbf{x}\|_1 > 0$

当 $\mathbf{x} = \mathbf{0}$ 时, 显然向量 \mathbf{x} 的所有分量必为 0, 故 1-范数一定为 0, 正定性证毕。

齐次性证明:

$$\begin{aligned}\|\lambda \mathbf{x}\|_1 &= |\lambda x_1| + |\lambda x_2| + \cdots + |\lambda x_n| \\ &= |\lambda| |x_1| + |\lambda| |x_2| + \cdots + |\lambda| |x_n| \\ &= |\lambda| (|x_1| + |x_2| + \cdots + |x_n|) \\ &= |\lambda| \|\mathbf{x}\|_1\end{aligned}$$

齐次性证毕。

三角不等式证明:

$$\|\mathbf{x} + \mathbf{y}\|_1 = \sum_{i=1}^n |x_i + y_i| \quad (1)$$

由 $|a + b| \leq |a| + |b|$

式 1 可以变成

$$\sum_{i=1}^n |x_i + y_i| \leq \sum_{i=1}^n |x_i| + \sum_{i=1}^n |y_i| = \|\mathbf{x}\|_1 + \|\mathbf{y}\|_1$$

即

$$\|\mathbf{x} + \mathbf{y}\|_1 \leq \|\mathbf{x}\|_1 + \|\mathbf{y}\|_1$$

这个不等式也叫 Minkovski 不等式。

三角不等式证毕, 因此, 1-范数是向量范数。

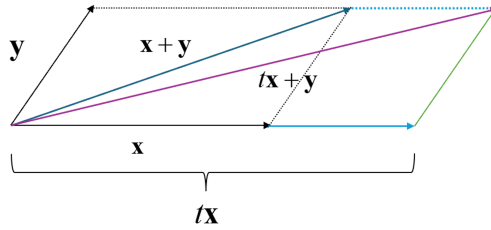
向量 2-范数的证明与 Cauchy 不等式 (柯西不等式)

在正式介绍向量 2-范数之前, 需要了解一下关于 Cauchy 不等式的有关内容。

Cauchy 不等式的内容为:

$$| \langle \mathbf{x}, \mathbf{y} \rangle | \leq \|\mathbf{x}\|_2 \cdot \|\mathbf{y}\|_2$$

即, 向量 \mathbf{x} 与向量 \mathbf{y} 的内积 ($\langle \cdot, \cdot \rangle$ 表示两个向量的内积) 的绝对值小于等于两个向量的 2-范数。



下面给出在二维欧式空间里 Cauchy 不等式的证明：

如图所示，假设 $t > 0$ ，向量 $x, y, x + y$ 和向量 $tx, y, tx + y$ 的情况如图，接下来针对向量 x 与向量 y 线性相关与线性无关时分类讨论：

当 x 与向量 y 线性无关时，两个向量的关系可以由上图表示，此时两向量不共线，可由反证法证明得 $tx + y \neq 0$ 。

所以

$$\begin{aligned} |tx + y|^2 &= \langle tx + y, tx + y \rangle \\ &= \langle x, x \rangle t^2 + 2 \langle x, y \rangle t + \langle y, y \rangle \end{aligned}$$

这是一个可以看做自变量为 t 的二次函数，令 $f(t) = \langle x, x \rangle t^2 + 2 \langle x, y \rangle t + \langle y, y \rangle$ 。

因为 tx 与 y 线性无关，所以 $tx + y \neq 0$ ，故 $|tx + y|^2$ 一定大于 0，所以 $f(t) > 0$ 。

由于 $f(t)$ 是一个二次函数，二次函数值恒大于零，代表二次函数与 x 轴没有交点，意味着 $\Delta = 4 \langle x, y \rangle^2 - 4 \langle x, x \rangle \langle y, y \rangle < 0$

即：

$$\langle x, y \rangle^2 \leq \|x\|_2^2 \|y\|_2^2$$

当 x 与向量 y 线性相关时，显而易见，两向量共线，即有关系

$$y = tx$$

此时

$$\begin{aligned} \langle x, y \rangle^2 &= \langle x, y \rangle \cdot \langle x, y \rangle \\ &= \langle x, tx \rangle \cdot \langle x, tx \rangle = k \|x\|_2^2 \cdot k \|x\|_2^2 = k^2 \|x\|_2^2 \cdot \|x\|_2^2 \\ &= \|y\|_2^2 \|x\|_2^2 \end{aligned}$$

故在欧式空间里，Cauchy 不等式成立。

需要特别注意的是，上面对于 Cauchy 不等式的证明是在欧式空间内的，如果放到酉空间中，因为酉空间内 $\langle x, y \rangle$ 的结果不一定与 $\langle y, x \rangle$ 的结果相同，故不能进行上面这样

的处理，如何证明 Cauchy 不等式在酉空间内依旧成立呢？请看后续对于 Hölder 不等式的证明。

接下来开始证明向量 2-范数是向量范数：

正定性证明：

若向量 $\mathbf{x} \neq \mathbf{0}$ ，则表明 \mathbf{x} 至少有一个分量 $x_k \neq 0$ ，从而 $|x_k| > 0$ ，故 $\|\mathbf{x}\|_2 > 0$

当 $\mathbf{x} = \mathbf{0}$ 时，显然向量 \mathbf{x} 的所有分量必为 0，故 2-范数一定为 0，正定性证毕。

齐次性证明：

$$\begin{aligned}\|\lambda \mathbf{x}\|_2 &= \sqrt{|\lambda x_1|^2 + |\lambda x_2|^2 + \cdots + |\lambda x_n|^2} \\ &= \sqrt{|\lambda|^2 |x_1|^2 + |\lambda|^2 |x_2|^2 + \cdots + |\lambda|^2 |x_n|^2} \\ &= |\lambda| \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2} \\ &= |\lambda| \|\mathbf{x}\|_2\end{aligned}$$

齐次性证毕。

三角不等式证明：

$$\begin{aligned}\|\mathbf{x} + \mathbf{y}\|_2^2 &= \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle \\ &= (\mathbf{x} + \mathbf{y})^H (\mathbf{x} + \mathbf{y}) \\ &= (\mathbf{x}^H + \mathbf{y}^H) (\mathbf{x} + \mathbf{y}) \\ &= \mathbf{x}^H \mathbf{x} + \mathbf{x}^H \mathbf{y} + \mathbf{y}^H \mathbf{x} + \mathbf{y}^H \mathbf{y} \\ &= |\mathbf{x}^H \mathbf{x} + \mathbf{x}^H \mathbf{y} + \mathbf{y}^H \mathbf{x} + \mathbf{y}^H \mathbf{y}| \tag{2}\end{aligned}$$

(2-范数一般可以理解为长度，长度一定非负，故加绝对值前后大小不变)

由 $|a + b| \leq |a| + |b|$ $|ab| = |a| \cdot |b|$

式 2 可以变为

$$\begin{aligned}|\mathbf{x}^H \mathbf{x} + \mathbf{x}^H \mathbf{y} + \mathbf{y}^H \mathbf{x} + \mathbf{y}^H \mathbf{y}| &\leq |\mathbf{x}^H \mathbf{x}| + |\mathbf{x}^H \mathbf{y}| + |\mathbf{y}^H \mathbf{x}| + |\mathbf{y}^H \mathbf{y}| \\ &= \|\mathbf{x}\|_2^2 + 2\|\mathbf{x}\|_2\|\mathbf{y}\|_2 + \|\mathbf{y}\|_2^2 \\ &= (\|\mathbf{x}\|_2 + \|\mathbf{y}\|_2)^2\end{aligned}$$

即

$$\|\mathbf{x} + \mathbf{y}\|_2^2 \leq (\|\mathbf{x}\|_2 + \|\mathbf{y}\|_2)^2$$

三角不等式证毕，因此，2-范数是向量范数。

向量无穷范数的证明

正定性证明：

若向量 $\mathbf{x} \neq \mathbf{0}$ ，则表明 \mathbf{x} 至少有一个分量 $x_k \neq 0$ ，从而 $|x_k| > 0$ ，故 $\|\mathbf{x}\|_\infty > 0$

当 $\mathbf{x} = \mathbf{0}$ 时，显然向量 \mathbf{x} 的所有分量必为 0，故无穷范数一定为 0，正定性证毕。

齐次性证明：

$$\begin{aligned}\|\lambda \mathbf{x}\|_\infty &= \max_{1 \leq i \leq n} |\lambda x_i| \\ &= |\lambda| \max_{1 \leq i \leq n} |x_i| \\ &= |\lambda| \|\mathbf{x}\|_\infty\end{aligned}$$

齐次性证毕。

三角不等式证明：

$$\|\mathbf{x} + \mathbf{y}\|_\infty = \max_{1 \leq i \leq n} |x_i + y_i| \leq \max_{1 \leq i \leq n} |x_i| + \max_{1 \leq i \leq n} |y_i| = \|\mathbf{x}\|_\infty + \|\mathbf{y}\|_\infty$$

三角不等式证毕，因此，无穷范数是向量范数。

2.1.5 向量 P-范数

在1.1.4小节中，着重介绍了常见的三种向量范数并给出了其能够成为向量范数的证明，不难发现这三种范数都有一定的相似的结构，能否有一种公式能够总括这些常见的范数呢？

实际上，1.1.4介绍的三种常见范数都是本小节要介绍的 P-范数的特殊情况，向量 P-范数又叫做 Hölder 范数，其形式如下

定义 2.1.2: 向量 P-范数的定义

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}, 1 \leq p < \infty$$

通过上面的定义，不难发现，当 $p = 1$ 的时候，其为向量 1-范数，当 $p = 2$ 的时候，其为向量 2-范数，当 $p \rightarrow \infty$ 的时候，其为向量无穷范数。

那如何证明向量 P 范数依旧是向量范数呢？接下来给出向量 P 范数是向量范数的证明，但在此之前，还需要介绍一些背景知识。

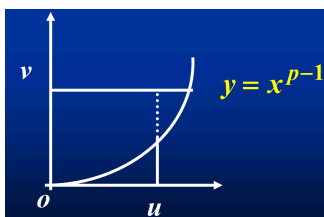
Young 不等式 (杨氏不等式)

接下来介绍 Young 不等式, Young 不等式的内容如下: 若 u 和 v 是非负实数, p, q 都是正实数, 且满足条件 $p > 1$ 和 $\frac{1}{p} + \frac{1}{q} = 1$, 则恒有不等式

$$uv \leq \frac{1}{p}u^p + \frac{1}{q}v^q$$

下面给出 Young 不等式的证明:

如下图



我们可以根据该图得出如下三部分:

- uv : 矩形的面积
- $\int_0^u x^{p-1} dx$: 以 (u, u^{p-1}) 与 x 轴包裹的小曲面三角形的面积
- $\int_0^v y^{\frac{1}{p-1}} dy$: 以 $(v^{\frac{1}{p-1}}, v)$ 与 y 轴包裹的小曲面三角形的面积

看图, 当 u 与 v 并不相交于函数同一点, 即 $v \neq u^{p-1}$ 的时候, 很明显能知道矩形的面积是小于两部分曲面三角形的面积的和的, 而当 $v = u^{p-1}$ 时, 矩形的面积恰好等于两部分曲面三角形的面积的和。

有了上面的认识, 接下来将通过严格的数学推受理性的证明 Young 不等式:

首先, 注意到

$$\frac{1}{q} + \frac{1}{p} = 1$$

对其进行变形, 有

$$pq = p + q$$

同时,

$$\begin{aligned} \int_0^u x^{p-1} dx &= \frac{1}{p} x^p \Big|_{x=0}^{x=u} \\ &= \frac{1}{p} u^p \end{aligned}$$

下面处理 $\int_0^v y^{\frac{1}{p-1}} dy$:

由 Young 不等式的前提, 得到了 $pq = p + q$, 从而有

$$p = (p - 1)q$$

同时有

$$\frac{q}{p} = \frac{1}{p - 1}$$

所以

$$\begin{aligned} \int_0^v y^{\frac{1}{p-1}} dy &= \int_0^v y^{\frac{q}{p}} dy \\ &= \frac{1}{\frac{q}{p} + 1} y^{\frac{q}{p} + 1} \Big|_{y=0}^{y=v} = \frac{1}{\frac{p+q}{p}} v^{\frac{p+q}{p}} = \frac{1}{\frac{pq}{p}} v^{\frac{pq}{p}} = \frac{1}{q} v^q \end{aligned}$$

故

$$uv \leq \frac{1}{p} u^p + \frac{1}{q} v^q$$

证毕。

Hölder 不等式

在1.1.4节介绍 Cauchy 不等式的最后提到, 用原本的在欧式空间中证明 Cauchy 不等式在西空间中依旧成立是不适用的, 那应该如何证明在西空间中, Cauchy 不等式依旧成立呢? 这就是本节要讨论的 Hölder 不等式, 其内容如下:

若 $p, q > 1$, 且 $\frac{1}{p} + \frac{1}{q} = 1$, 则对 \mathbf{C}^n 中任意向量 $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$, $\mathbf{y} = (y_1, y_2, \dots, y_n)^T$ 都有

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

这里看等号右边的这两个乘积项, 不难发现这与范数的定义很相似, 就是 $\|\mathbf{x}\|_p$ 与 $\|\mathbf{y}\|_q$

下面给出 Hölder 不等式的证明:

令

$$u = \frac{|x_i|}{\|\mathbf{x}\|_p}, v = \frac{|y_i|}{\|\mathbf{y}\|_q}$$

由 Young 不等式, 可以得到

$$uv = \frac{|x_i| |y_i|}{\|\mathbf{x}\|_p \|\mathbf{y}\|_q} \leq \frac{1}{p} \frac{|x_i|^p}{\|\mathbf{x}\|_p^p} + \frac{1}{q} \frac{|y_i|^q}{\|\mathbf{y}\|_q^q} \quad (i = 1, 2, \dots, n)$$

等式左右两侧分别进行累加, 从 1 加至 n , 从而有

$$\sum_{i=1}^n \frac{|x_i| |y_i|}{\|\mathbf{x}\|_p \|\mathbf{y}\|_q} \leq \frac{1}{p \|\mathbf{x}\|_p^p} \sum_{i=1}^n |x_i|^p + \frac{1}{q \|\mathbf{y}\|_q^q} \sum_{i=1}^n |y_i|^q \quad (1)$$

由 p -范数的式子

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

可以得出

$$\|\mathbf{x}\|_p^p = \sum_{i=1}^n |x_i|^p$$

从而有

$$\frac{\sum_{i=1}^n |x_i|^p}{\|\mathbf{x}\|_p^p} = 1 \quad (2)$$

故，不等式 1 的右侧不等号

$$\frac{1}{p\|\mathbf{x}\|_p^p} \sum_{i=1}^n |x_i|^p + \frac{1}{q\|\mathbf{y}\|_q^q} \sum_{i=1}^n |y_i|^q = \frac{1}{p} \sum_{i=1}^n \frac{|x_i|^p}{\|\mathbf{x}\|_p^p} + \frac{1}{q} \sum_{i=1}^n \frac{|y_i|^q}{\|\mathbf{y}\|_q^q} \quad (3)$$

由式 2 的结果，式 3 的结果最后为

$$\frac{1}{p} + \frac{1}{q} = 1 \text{ (根据 Young 不等式的前提)}$$

故式 1 化简为

$$\sum_{i=1}^n \frac{|x_i||y_i|}{\|\mathbf{x}\|_p\|\mathbf{y}\|_q} \leq 1$$

等式左右两侧同时乘 $\|\mathbf{x}\|_p\|\mathbf{y}\|_q$ ，有

$$\sum_{i=1}^n |x_i| \cdot |y_i| \leq \|\mathbf{x}\|_p\|\mathbf{y}\|_q$$

即

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

证毕。

上面的是 Hölder 不等式的全部证明过程，那 Hölder 不等式与 Cauchy 不等式的关系是什么呢？

仔细观察 Hölder 不等式的右侧，可以发现，当 $p = q = 2$ 时，这就是 Cauchy 不等式的右侧，那左侧的 $|\langle \mathbf{x}, \mathbf{y} \rangle|$ 要怎么得出来呢？

根据内积的定义，有

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^H \mathbf{y} = \bar{x}_1 y_1 + \bar{x}_2 y_2 + \cdots + \bar{x}_n y_n$$

故等式两侧都加绝对值，有

$$| \langle \mathbf{x}, \mathbf{y} \rangle | = | \mathbf{x}^H \mathbf{y} | = | \bar{x}_1 y_1 + \bar{x}_2 y_2 + \cdots + \bar{x}_n y_n | \quad (4)$$

根据绝对值三角不等式

$$|a + b| \leq |a| + |b|$$

式 4 的结果为

$$| \langle \mathbf{x}, \mathbf{y} \rangle | \leq | \bar{x}_1 y_1 | + | \bar{x}_2 y_2 | + \cdots + | \bar{x}_n y_n | \quad (5)$$

又根据性质

$$|ab| = |a| \cdot |b|$$

式 5 的结果为

$$| \langle \mathbf{x}, \mathbf{y} \rangle | \leq | \bar{x}_1 y_1 | + | \bar{x}_2 y_2 | + \cdots + | \bar{x}_n y_n | = |x_1| |y_1| + |x_2| |y_2| + \cdots + |x_n| |y_n| = \sum_{i=1}^n |x_i| |y_i| \quad (5)$$

所以我们就得出了如下的不等关系：

$$| \langle \mathbf{x}, \mathbf{y} \rangle | \leq \sum_{i=1}^n |x_i| |y_i| \leq \left(\sum_{i=1}^n |x_i|^p \right)^{1/p} \left(\sum_{i=1}^n |y_i|^q \right)^{1/q}$$

这样，Cauchy 不等式在酉空间下依旧成立，证毕。

P-范数是向量范数的证明

有了前面这些定理，接下来开始证明 P-范数是向量范数：

正定性证明：

若向量 $\mathbf{x} \neq \mathbf{0}$ ，则表明 \mathbf{x} 至少有一个分量 $x_k \neq 0$ ，从而 $|x_k| > 0$ ，故 $\|\mathbf{x}\|_p > 0$

当 $\mathbf{x} = \mathbf{0}$ 时，显然向量 \mathbf{x} 的所有分量必为 0，故 p -范数一定为 0，正定性证毕。

齐次性证明：

$$\begin{aligned} \|\lambda \mathbf{x}\|_p &= \sqrt[p]{|\lambda x_1|^p + |\lambda x_2|^p + \cdots + |\lambda x_n|^p} \\ &= \sqrt[p]{|\lambda|^p |x_1|^p + |\lambda|^p |x_2|^p + \cdots + |\lambda|^p |x_n|^p} \\ &= |\lambda| \sqrt[p]{|x_1|^p + |x_2|^p + \cdots + |x_n|^p} \\ &= |\lambda| \|\mathbf{x}\|_p \end{aligned}$$

齐次性证毕。

三角不等式证明：

当 $p = 1$ 的时候，很明显，这是 1-范数，1-范数的三角不等式证明过程请参见 1.1.4 节。

当 $p > 1$ 时

$$\begin{aligned} \sum_{i=1}^n (|x_i| + |y_i|)^p &= \sum_{i=1}^n (|x_i| + |y_i|) \cdot (|x_i| + |y_i|)^{p-1} \\ &= \sum_{i=1}^n |x_i| (|x_i| + |y_i|)^{p-1} + \sum_{i=1}^n |y_i| (|x_i| + |y_i|)^{p-1} \end{aligned} \quad (8)$$

由 Hölder 不等式，式 8 可以变为

$$\begin{aligned} \sum_{i=1}^n |x_i| (|x_i| + |y_i|)^{p-1} + \sum_{i=1}^n |y_i| (|x_i| + |y_i|)^{p-1} &\leq \sum_{i=1}^n [(|x_i| + |y_i|)^{(p-1)q}]^{\frac{1}{q}} \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} \\ &\quad + \sum_{i=1}^n [(|x_i| + |y_i|)^{(p-1)q}]^{\frac{1}{q}} \left(\sum_{i=1}^n |y_i|^p \right)^{\frac{1}{p}} \end{aligned}$$

注意到不等式右侧都有项 $\sum_{i=1}^n [(|x_i| + |y_i|)^{(p-1)q}]^{\frac{1}{q}}$ ，提取公因式，有

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

所以经过整理，我们得到了下面这个式子：

$$\sum_{i=1}^n (|x_i| + |y_i|)^p \leq \left(\left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |y_i|^p \right)^{\frac{1}{p}} \right) \left[\sum_{i=1}^n (|x_i| + |y_i|)^{(p-1)q} \right]^{\frac{1}{q}} \quad (9)$$

由 Young 不等式的前提条件（如果忘记了请看 1.4 节） $p = (p-1)q$ ，式 9 可以变为

$$\sum_{i=1}^n (|x_i| + |y_i|)^p \leq \left(\left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |y_i|^p \right)^{\frac{1}{p}} \right) \left[\sum_{i=1}^n (|x_i| + |y_i|)^p \right]^{\frac{1}{q}} \quad (10)$$

注意到不等号左右两侧都有项 $\sum_{i=1}^n (|x_i| + |y_i|)^p$ ，故提公因式，有

$$\left[\sum_{i=1}^n (|x_i| + |y_i|)^p \right]^{1-\frac{1}{q}} \leq \left(\left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |y_i|^p \right)^{\frac{1}{p}} \right) \quad (11)$$

又由 Young 不等式的前提条件： $\frac{1}{p} + \frac{1}{q} = 1$ ，故式 11 可以化简为：

$$\left[\sum_{i=1}^n (|x_i| + |y_i|)^p \right]^{\frac{1}{p}} \leq \left(\left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |y_i|^p \right)^{\frac{1}{p}} \right) \quad (12)$$

接下来, 只需要证明

$$\|\mathbf{x} + \mathbf{y}\|_p \leq \left[\sum_{i=1}^n (|x_i| + |y_i|)^p \right]^{\frac{1}{p}}$$

即可

由 p -范数的定义和绝对值三角不等式, 有

$$\|\mathbf{x} + \mathbf{y}\|_p = \left(\sum_{i=1}^n (|x_i + y_i|)^p \right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^n (|x_i| + |y_i|)^p \right)^{\frac{1}{p}} \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p$$

故三角不等式成立, 故 p 范数是向量范数, 证毕。

证明 p -范数的三角不等式比较复杂, 主要涉及下面五步:

1. 将起点 $\sum_{i=1}^n (|x_i| + |y_i|)^p$ 变换为 1 与 $p-1$ 两个幂相乘的形式, 并应用分配律展开。
2. 在第一步的基础上, 应用 Hölder 不等式进行放大。
3. 提公因式 $\sum_{i=1}^n [(|x_i| + |y_i|)^{(p-1)q}]^{\frac{1}{q}}$, 并应用 Young 不等式的前提条件进行化简
4. 不等式左右两侧同时除以 $\sum_{i=1}^n (|x_i| + |y_i|)^p$, 并继续应用 Young 不等式的前提条件进行化简。
5. 应用 $|x_i + y_i| \leq |x_i| + |y_i|$, 得到最后的结论。

由此, 通过 p -范数的定义, 我们知道了一个范数有无数种种类描述, 并非只有向量 1-范数、向量 2-范数或向量无穷范数, 同时经过 1.1.4 和 1.1.5 两个小节介绍, 对于向量的范数有了更直观的感受——范数的确是一种广义的“距离”。

2.1.6 向量范数的应用

在本节的最后一小节, 介绍一下向量范数的三大应用——利用高维空间的范数定义低维空间的向量范数、向量范数的等价和利用向量范数研究向量序列的收敛性

利用高维空间的向量范数定义低维空间的向量范数

定理 2.1.1

设 $\|\cdot\|$ 是 \mathbb{C}^n 上的范数, $A \in \mathbb{C}_n^{m \times n}$ (表示 A 矩阵的列数等于矩阵的秩, 即 A 矩阵列满秩), 则 $\|A \cdot\|$ 是 \mathbb{C}^n 上的范数。

定理 1.1.1 的作用是什么呢? 在很多情况下, 如果我们知道高维空间内的向量范数, 因为各种原因求得低维空间内的向量范数较为困难或者根本无法求得, 那么在这种情况下, 我们

可以构造一个列满秩的矩阵，与低维空间内的该向量（即定理中的“ \cdot ”）相乘，便可求出低维空间内的向量的范数。

为什么可以这样做呢？还记得在上一章中讲过的线性变换的概念吗？

假设有一个 m 维的向量 \mathbf{y} ，还有一个 n 维的向量 \mathbf{x} ($m > n$)，假设有一个矩阵 A ，便可以做如下线性变换：

$$\mathbf{y} = A\mathbf{x}$$

这样，通过一个矩阵 A ，便成功的将低维的向量 \mathbf{x} 升维至高维。

如果上面这样说还是觉得不理解，接下来通过一个具体的例子来展示一下何为“升维”

例 2.1.1. 假设有矩阵 $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$ (A 是一个三维矩阵)，向量 $\mathbf{x} = [x_1, x_2]^T$ ，(\mathbf{x} 是一个

二维矩阵)

$$\text{那么 } A\mathbf{x} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_2 \\ x_1 + x_2 \end{bmatrix}, \text{ 很显然, } A\mathbf{x} \text{ 变为了一个三维矩阵, 这就是通过}$$

一个矩阵实现将低维向量升维至高维空间。

那么在范数层面上，也可以实现一样的功能，公式如下：

$$\|\mathbf{x}\|_n \stackrel{\text{def}}{=} \|\mathbf{y}\|_m = \|A\mathbf{x}\|_m$$

下面给出使用高维空间向量范数来定义低维空间向量范数合理性的证明，依旧是从范数的三个性质出发

证明.

1. 正定性

由于 $\mathbf{x} \neq 0$ ，并且 A 是列满秩矩阵，由线性代数的知识可以得出齐次线性方程组 $A\mathbf{x} = \mathbf{0}$ 只有零解，故范数 $\|A\mathbf{x}\| > 0$

因此，由 $A\mathbf{x}$ 定义的 \mathbf{x} 的范数 $\|\mathbf{x}\|$ 也大于零，正定性证毕。

2. 齐次性

$$\|\lambda\mathbf{x}\|_n \stackrel{\text{def}}{=} \|A\lambda\mathbf{x}\|_m = \|\lambda A\mathbf{x}\|_m$$

由于 λ 是一个数，同时 $\|A\mathbf{x}\|_m$ 已经是一个范数，满足范数的齐次性，因此上式的结果为 $|\lambda|\|A\mathbf{x}\|$ ，齐次性证毕。

3. 三角不等式

$$\|\mathbf{x}_1 + \mathbf{x}_2\|_n = \|A(\mathbf{x}_1 + \mathbf{x}_2)\|_m = \|A\mathbf{x}_1 + A\mathbf{x}_2\|_m$$

显然, $\|A\mathbf{x}_1\|$ 与范数 $\|A\mathbf{x}_2\|$ 已经是范数, 由范数的三角不等式, 有

$$\|A\mathbf{x}_1 + A\mathbf{x}_2\|_m \leq \|A\mathbf{x}_1\|_m + \|A\mathbf{x}_2\|_m = \|\mathbf{x}_1\|_n + \|\mathbf{x}_2\|_n, \text{ 三角不等式证毕。}$$

□

下面用这个定理来看一道例题

例 2.1.2. 设 $\|\mathbf{x}\|$ 是 \mathbb{P}^n 中的向量范数, $A \in \mathbb{P}^{n \times n}$, 则 $\|A\mathbf{x}\|$ 也是 \mathbb{P}^n 中的向量范数的充要条件是 A 可逆。

解. 先证充分性 (右推左):

显然, 由定理 1.1.1 可知, 在 A 可逆的条件下, 其一定为列满秩矩阵, 因此 $\|A\mathbf{x}\|$ 一定为向量范数。

再证必要性 (左推右):

使用反证法, 假设 A 不可逆, 故齐次线性方程 $A\mathbf{x} = \mathbf{0}$ 存在非零解 \mathbf{x}_0

因此, $\|\mathbf{x}_0\| \stackrel{\text{def}}{=} \|A\mathbf{x}_0\| = 0$, 这与范数的正定性相矛盾。

证毕。

□

向量范数的等价

定义 2.1.3: 向量范数等价的定义

设在 $V_n(P)$ 上定义了 $\|\mathbf{x}\|_a, \|\mathbf{x}\|_b$ 两种向量范数, 若存在常数 $C_1 > 0, C_2 > 0$, 使得 $\forall \mathbf{x} \in V_n(P)$, 都有

$$\|\mathbf{x}\|_a \leq C_1 \|\mathbf{x}\|_b,$$

$$\|\mathbf{x}\|_b \leq C_2 \|\mathbf{x}\|_a$$

则称 $\|\mathbf{x}\|_a$ 与 $\|\mathbf{x}\|_b$ 等价

向量范数等价的概念是为后面利用向量范数判断向量序列的收敛性做服务的, 由定义 1.1.3 里的不等关系, 我们可以构造出下面的两组不等式:

$$\frac{1}{C_1} \|\mathbf{x}\|_a \leq \|\mathbf{x}\|_b \leq C_2 \|\mathbf{x}\|_a$$

和

$$\frac{1}{C_2} \|\mathbf{x}\|_b \leq \|\mathbf{x}\|_a \leq C_1 \|\mathbf{x}\|_b$$

由于范数是一个数, 故上式的不等式可以由夹逼定理确定向量范数不受范数的种类影响而能够唯一收敛到一个值。

利用向量范数判断向量序列的收敛性

何为向量序列? 在高等数学的课程中, 我们接触过数列这一概念, 如, 等差数列, 等比数列这些都是数列, 下面将数列这一概念延伸一下便能够得到向量序列的知识。

假设有如下的向量

$$\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \dots, \boldsymbol{x}^{(k)}, \dots$$

其中

$$\boldsymbol{x}^{(1)} = \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ \vdots \\ x_n^{(1)} \end{bmatrix}, \boldsymbol{x}^{(2)} = \begin{bmatrix} x_1^{(2)} \\ x_2^{(2)} \\ \vdots \\ x_n^{(2)} \end{bmatrix}, \dots, \boldsymbol{x}^{(k)} = \begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \\ \vdots \\ x_n^{(k)} \end{bmatrix}, \dots$$

这个时候, 序列 $x_1^{(1)}, x_1^{(2)}, \dots, x_1^{(k)}, \dots$, 序列 $x_2^{(1)}, x_2^{(2)}, \dots, x_2^{(k)}, \dots$ 等 (把这些向量的每一行组合在一起) 就是一个**数列**, 而把这些数列组合在一起, 就是这些向量, 这些向量组成了一个**向量序列**。

在高等数学的课程中, 我们了解过数列极限 (下面这段定义摘录自《张宇考研数学基础30讲——高等数学分册》):

设 x_n 为一数列, 若存在常数 a , 对于任意的 ε (不论它多么小), 总存在正整数 N , 使得当 $n > N$ 时, $|x_n - a| < \varepsilon$ 恒成立, 则称常数 a 是数列 x_n 的极限, 或者称数列 x_n 收敛于 a , 记为

$$\lim_{n \rightarrow \infty} x_n = a \text{ 或 } x_n \rightarrow a \quad (n \rightarrow \infty)$$

将数列极限这一概念延伸到向量序列的极限中, 上面介绍的每一个数列如果都能收敛, 那么就可以说整个向量序列也是收敛的, 趋近于某一个向量。

如果向量序列是收敛的, 会有如下等式:

$$\lim_{k \rightarrow \infty} \boldsymbol{x}^{(k)} = \boldsymbol{a}$$

上式还有下面这种利用范数来判断向量序列是否收敛的定义:

定理 2.1.2

设 $\|\cdot\|$ 是 \mathbb{C}^n 上的任一向量范数, 则

$$\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{a} \iff \lim_{k \rightarrow \infty} \|\mathbf{x}^{(k)} - \mathbf{a}\| = 0$$

上面这两个式子的等价证明如下:

证明.

因为

$$\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{a}$$

所以

$$\lim_{k \rightarrow \infty} \mathbf{x}_i^{(k)} = \mathbf{a}_i \quad (i = 1, 2, \dots, n)$$

所以

$$\lim_{k \rightarrow \infty} |\mathbf{x}_i^{(k)} - \mathbf{a}_i| = 0$$

这里注意到 $\mathbf{x}_i^{(k)}$ 和 \mathbf{a}_i 的任意性, $\lim_{k \rightarrow \infty} |\mathbf{x}_i^{(k)} - \mathbf{a}_i| = 0$ 这个等式表明的就是向量序列的每一个分量所构成的数列都是收敛的, 那自然存在下面的关系:

$$\lim_{k \rightarrow \infty} \max_{1 \leq i \leq n} \{|\mathbf{x}_i^{(k)} - \mathbf{a}_i|\} = 0$$

既然每一个分量构成的数列是收敛的, 那自然最大的那一行分量构成的数列依旧是收敛的, 仔细看这个式子, 可以与向量的无穷范数建立起了联系, 即

$$\lim_{k \rightarrow \infty} \|\mathbf{x}_i^{(k)} - \mathbf{a}_i\|_{\infty} = 0$$

证毕。 □

2.2 矩阵的范数

上一节介绍了向量范数的有关知识, 包括向量范数的性质、判断条件、常见的向量范数等, 本节将眼光放到矩阵范数中。

2.2.1 从向量范数到矩阵范数

上一节我们知道了任意的向量范数 (即向量 p -范数) 的形式, 那矩阵的范数又应该如何理解呢?

矩阵范数和向量范数的形式及其类似，甚至可以通过向量范数的形式引出矩阵范数，将 \mathbb{P}^n 上的范数推广到 $\mathbb{P}^{m \times n}$ 上。

为什么可以这样推广呢？这里需要做这样一步认定：我们可以将一个 $m \times n$ 的矩阵 A 看作是数域 \mathbb{P}^{mn} 上的一个向量，如此就可以将一个矩阵当做一个向量来处理，自然就可以用前一节的向量范数的相关概念来定义矩阵范数。

既然如此，那矩阵范数的定义就不难理解了：

定义 2.2.1: 矩阵范数的定义

设 $A \in \mathbb{P}^{m \times n}$ ，若映射 $\|\cdot\|: \mathbb{P}^{m \times n} \rightarrow \mathbb{R}$ 满足：

- 正定性： $\|A\| \geq 0$ ，当且仅当 $A = 0$ 时， $\|A\| = 0$
- 齐次性： $\|\lambda A\| = |\lambda| \|A\|, \forall \lambda \in \mathbb{R}, \forall A \in \mathbb{P}^{m \times n}$
- 三角不等式： $\|A + B\| \leq \|A\| + \|B\|, \forall A, B \in \mathbb{P}^{m \times n}$

则称映射 $\|\cdot\|$ 为 $\mathbb{P}^{m \times n}$ 上的**矩阵范数**。

2.2.2 常见的矩阵范数

与常见的向量范数相似，常见的矩阵范数有矩阵 1-范数，矩阵 2-范数和矩阵无穷范数：

$$\begin{aligned}\|A\|_{m_1} &= \sum_{j=1}^n \sum_{i=1}^m |a_{ij}| \\ \|A\|_{m_2} &= \left(\sum_{j=1}^n \sum_{i=1}^m |a_{ij}|^2 \right)^{\frac{1}{2}} \\ \|A\|_{m_\infty} &= \max_{i,j} \{|a_{ij}|\}, 1 \leq i \leq m, 1 \leq j \leq n\end{aligned}$$

由于矩阵范数和向量范数的相似性，故上面的三种常见矩阵范数的证明过程就不详细说了，矩阵范数更值得关注的是矩阵范数的相容性质。

2.2.3 矩阵范数的相容

矩阵范数相容的定义如下：

定义 2.2.2: 范数相容的定义

设 $\|\cdot\|_a: \mathbb{P}^{m \times l} \rightarrow \mathbb{R}$, $\|\cdot\|_b: \mathbb{P}^{l \times n} \rightarrow \mathbb{R}$, $\|\cdot\|_c: \mathbb{P}^{m \times n} \rightarrow \mathbb{R}$ 是矩阵范数, 如果

$$\|AB\|_c \leq \|A\|_a \cdot \|B\|_b$$

则称矩阵范数 $\|\cdot\|_a, \|\cdot\|_b$ 和 $\|\cdot\|_c$ 相容。

如果

$$\|AB\| \leq \|A\| \cdot \|B\|$$

则称 $\|\cdot\|$ 是子相容的矩阵范数

注. 今后若无特殊说明, 所有的“相容范数”均为自相容范数。

有矩阵范数相容的作用是什么呢? 我们知道, 矩阵乘法的法则是这样的 (下面这段话摘自同济大学《工程数学——线性代数》):

设 $A = (a_{ij})$ 是一个 $m \times s$ 矩阵, $B = (b_{ij})$ 是一个 $s \times n$ 矩阵, 那么规定矩阵 A 与矩阵 B 的乘积是一个 $m \times n$ 矩阵 $C = (c_{ij})$, 其中

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{is}b_{sj} = \sum_{k=1}^s a_{ik}b_{kj}, \quad (i = 1, 2, \cdots, m; j = 1, 2, \cdots, n)$$

通俗点说, 两个矩阵相乘, 得到的新矩阵的对应元素是第一个矩阵的行乘以第二个元素的列的结果求和得出的, 那么这个过程用计算机来表述的话, 代码是这样的 (以 C 语言为例):

```

code/mem/matmult/mm.c
1 for (i = 0; i < n; i++)
2   for (j = 0; j < n; j++) {
3     sum = 0.0;
4     for (k = 0; k < n; k++)
5       sum += A[i][k]*B[k][j];
6     C[i][j] += sum;
7   }

code/mem/matmult/mm.c
1 for (j = 0; j < n; j++)
2   for (i = 0; i < n; i++) {
3     sum = 0.0;
4     for (k = 0; k < n; k++)
5       sum += A[i][k]*B[k][j];
6     C[i][j] += sum;
7   }

code/mem/matmult/mm.c
1 for (j = 0; j < n; j++)
2   for (k = 0; k < n; k++)
3     for (i = 0; i < n; i++)
4       C[i][j] += A[i][k]*B[k][j];

code/mem/matmult/mm.c
1 for (k = 0; k < n; k++)
2   for (i = 0; i < n; i++)
3     for (j = 0; j < n; j++)
4       C[i][j] += A[i][k]*B[k][j];

code/mem/matmult/mm.c
1 for (i = 0; i < n; i++)
2   for (k = 0; k < n; k++)
3     for (j = 0; j < n; j++)
4       C[i][j] += A[i][k]*B[k][j];

code/mem/matmult/mm.c
1 for (j = 0; j < n; j++)
2   for (k = 0; k < n; k++)
3     for (i = 0; i < n; i++)
4       C[i][j] += A[i][k]*B[k][j];

```

图 6-44 矩阵乘法的六个版本。每个版本都以它循环的顺序来唯一地标识

图 2.1: 矩阵乘法代码示意

图1.1的代码节选自《深入理解计算机系统》，在书中对于矩阵乘法给出了六个版本的算

法，但算法中 i, j, k 变量的不同顺序仅仅是对访问内存的次数有影响¹，对于整体的时间复杂度，依旧为 $O(n^3)$ ，这个算法的时间复杂度在面对两个极为庞大的矩阵进行乘法运算是要消耗非常大的时间的，如果在这种情况下计算矩阵范数是一件十分耗时的事情。

但幸运的是，在许多工程中，我们其实没有必要知道两个矩阵相乘的范数的精确数值（在很多情况下都是这样，我们只需要一个近似的数值，或者说了解到一个上界或者下界就可以，比如上文中提到的时间复杂度大 O 记法，它仅仅描述了一个算法所耗费时间的最坏情况，即一个上界），在这种情况下，矩阵的相容性就会起到作用，计算两个单独的矩阵的范数再相乘一定比计算两个矩阵先相乘再求范数省时。²

2.2.4 常见矩阵范数的相容性

在了解了何为矩阵范数相容的概念后，接下来就要讨论常见矩阵范数（矩阵 1-范数，矩阵 2-范数和矩阵无穷范数）的相容性了。

矩阵 1-范数的相容性

前面已经介绍过矩阵 1-范数的结构

$$\|A\|_{m_1} = \sum_{j=1}^n \sum_{i=1}^m |a_{ij}|$$

下面来证明矩阵 1-范数是自相容范数：

证明。 若证明矩阵 1-范数是自相容范数，只需证

$$\|AB\|_{m_1} \leq \|A\|_{m_1} \|B\|_{m_1}$$

设 $A \in \mathbb{P}^{m \times l}$, $B \in \mathbb{P}^{l \times n}$ ，根据矩阵 1-范数的定义，有

$$\begin{aligned} \|AB\|_{m_1} &= \sum_{i=1}^m \sum_{j=1}^n |a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{il}b_{lj}| \quad (\text{矩阵乘法}) \\ &\leq \sum_{i=1}^m \sum_{j=1}^n (|a_{i1}||b_{1j}| + |a_{i2}||b_{2j}| + \cdots + |a_{il}||b_{lj}|) \quad (\text{绝对值三角不等式}) \\ &= \sum_{i=1}^m \sum_{j=1}^n \sum_{k=1}^l |a_{ik}||b_{kj}| \quad (\text{用一个连加号概括上式}) \end{aligned}$$

¹至于为什么交换变量顺序会对访问内存有影响，这里涉及到了程序的局部性原理，已超出了本门课程的范畴，具体请翻阅计算机组成原理相关教材。

²关于算法的时间复杂度，如何衡量一个计算机算法的好坏，具体请翻阅计算机算法方面的教材。

交换一下连加号的顺序，原式变为

$$\sum_{k=1}^l \sum_{i=1}^m \sum_{j=1}^n |a_{ik}| |b_{kj}|$$

由于连加号的计算是由内向外计算的，因此会首先计算 $\sum_{j=1}^n |a_{ik}| |b_{kj}|$ ，这个时候会发现 a_{ik} 并没有出现下标 j ，也就是说 $|a_{ik}|$ 并不会受 j 的影响，因此，在 $\sum_{j=1}^n |a_{ik}| |b_{kj}|$ 中， $|a_{ik}|$ 可以被认为是一个公因子提到公式外面去，则原式可以变为

$$\sum_{k=1}^l \sum_{i=1}^m |a_{ik}| \sum_{j=1}^n |b_{kj}|$$

此时观察 $\sum_{i=1}^m |a_{ik}| \sum_{j=1}^n |b_{kj}|$ ，会发现 $\sum_{j=1}^n |b_{kj}|$ 并不受 i 的影响，因此可以把这部分整体当做一个公因式提到外面，于是式子就会变成

$$\sum_{k=1}^l \left[\left(\sum_{i=1}^m |a_{ik}| \right) \left(\sum_{j=1}^n |b_{kj}| \right) \right]$$

这里， $(\sum_{i=1}^m |a_{ik}|)$ 叫做 A 的绝对列和， $(\sum_{j=1}^n |b_{kj}|)$ 叫做 B 的绝对行和，这两个概念在下一节算子范数中还会有应用。

如果想继续提公因式，需要对式子进行放缩，这里引入矩阵 A 的矩阵 1-范数，上式变为

$$\sum_{k=1}^l \left(\sum_{j=1}^n |b_{kj}| \right) \left(\sum_{k=1}^l \sum_{i=1}^m |a_{ik}| \right)$$

这相当于是在原本 A 矩阵的绝对列和的基础上把所有的 A 元素相加，显然结果更大，因此有如下不等关系

$$\sum_{k=1}^l \left[\left(\sum_{i=1}^m |a_{ik}| \right) \left(\sum_{j=1}^n |b_{kj}| \right) \right] \leq \sum_{k=1}^l \left(\sum_{j=1}^n |b_{kj}| \right) \left(\sum_{k=1}^l \sum_{i=1}^m |a_{ik}| \right)$$

这种情况下可以发现， $(\sum_{k=1}^l \sum_{i=1}^m |a_{ik}|)$ 并不受 $\sum_{k=1}^l (\sum_{j=1}^n |b_{kj}|)$ 中 $(\sum_{k=1}^l \sum_{i=1}^m |a_{ik}|)$ 中最外层的连加号 $\sum_{k=1}^l$ 的影响，因此 $(\sum_{k=1}^l \sum_{i=1}^m |a_{ik}|)$ 又是一个公因式，将其提出，最后式子变为

$$\sum_{k=1}^l \left[\left(\sum_{i=1}^m |a_{ik}| \right) \left(\sum_{j=1}^n |b_{kj}| \right) \right] \leq \left(\sum_{k=1}^l \sum_{i=1}^m |a_{ik}| \right) \left(\sum_{k=1}^l \sum_{j=1}^n |b_{kj}| \right)$$

注意到不等式右侧就是 A 矩阵的矩阵 1-范数和 B 矩阵的矩阵 1-范数，因此得到了

$$\|AB\|_{m_1} \leq \|A\|_{m_1} \|B\|_{m_1}$$

也就说明矩阵 1-范数是自相容范数，证毕。 □

矩阵 2-范数的相容性

前面已经介绍过矩阵 2-范数的结构

$$\|A\|_{m_2} = \left(\sum_{j=1}^n \sum_{i=1}^m |a_{ij}|^2 \right)^{\frac{1}{2}}$$

下面来证明矩阵 2-范数是自相容范数：

证明. 若想证矩阵 2-范数是自相容范数，只需证

$$\|AB\|_{m_2}^2 \leq \|A\|_{m_2}^2 \cdot \|B\|_{m_2}^2$$

根据矩阵 2-范数的定义，有

$$\begin{aligned} \|AB\|_{m_2}^2 &= \sum_{i=1}^m \sum_{j=1}^n |a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{il}b_{lj}|^2 \\ &\leq \sum_{i=1}^m \sum_{j=1}^n [|a_{i1}||b_{1j}| + \cdots + |a_{il}||b_{lj}|]^2 \\ &= \sum_{i=1}^m \sum_{j=1}^n \left(\sum_{k=1}^l |a_{ik}||b_{kj}| \right)^2 \end{aligned}$$

前面三步的方法跟证明矩阵 1-范数是自相容范数的流程是一样的，接下来是不一样的

根据 Cauchy 不等式，有

$$\|AB\|_{m_2}^2 \leq \sum_{i=1}^m \left[\sum_{j=1}^n \left(\sum_{k=1}^l |a_{ik}|^2 \right) \left(\sum_{k=1}^l |b_{kj}|^2 \right) \right]$$

请注意不等式右侧的 $\sum_{j=1}^n \left(\sum_{k=1}^l |a_{ik}|^2 \right) \left(\sum_{k=1}^l |b_{kj}|^2 \right)$ 这一部分，依旧是由于计算连加是从内向外计算的，因此项 $\sum_{k=1}^l |a_{ik}|^2$ 并不受 $\sum_{j=1}^n$ 中 j 的影响，故可以当作公因式提出，如此式子就变为了

$$\|AB\|_{m_2}^2 \leq \sum_{i=1}^m \left(\sum_{k=1}^l |a_{ik}|^2 \right) \left(\sum_{j=1}^n \sum_{k=1}^l |b_{kj}|^2 \right)$$

此时注意到在不等式的右侧，项 $\sum_{j=1}^n \sum_{k=1}^l |b_{kj}|^2$ 依旧不受 $\sum_{i=1}^m$ 中 i 的影响，故整体可以当作公因式提出，如此式子就变成了

$$\|AB\|_{m_2}^2 \leq \left(\sum_{i=1}^m \sum_{k=1}^l |a_{ik}|^2 \right) \left(\sum_{j=1}^n \sum_{k=1}^l |b_{kj}|^2 \right)$$

即

$$\|AB\|_{m_2}^2 \leq \|A\|_{m_2}^2 \|B\|_{m_2}^2$$

故矩阵 2-范数是自相容范数，证毕。 \square

矩阵 2-范数也叫 Frobenius 范数，其也是用途最多的一种范数，矩阵 2-范数还有下面的性质

定理 2.2.1: 矩阵 2-范数 (Frobenius 范数) 的性质

设 $A \in \mathbb{P}^{n \times n}$,

1. 若 $A = (a_1, a_2, \dots, a_n)$, 则

$$\|A\|_F^2 \stackrel{\textcircled{1}}{=} \|A\|_{m_2}^2 \stackrel{\textcircled{2}}{=} \sum_{i=1}^n \|a_i\|_2^2$$

其中, $\|a_i\|_2^2 = a_i^H a_i$ 是 \mathbb{P}^n 中的向量范数

2. $\|A\|_{m_2}^2 \stackrel{\textcircled{3}}{=} \text{tr}(A^H A) \stackrel{\textcircled{4}}{=} \sum_{i=1}^n \lambda_i(A^H A)$

3. 对任意的酉矩阵 $U, V \in \mathbb{P}^{n \times n}$, 有

$$\|A\|_{m_2} = \|U^H A V\|_{m_2} = \|U A V^H\|_{m_2}$$

这里只给出定理 2.2.1 中前两条的证明，第三条的证明在后面的酉不变范数中。

证明.

① 的证明省略，因为这就是定义——矩阵 2-范数也叫 Frobenius 范数，这无需证明。

④ 的证明省略，因为根据矩阵特征值与矩阵的迹的关系，矩阵特征值的和就是矩阵的迹，这也无需证明。

②, ③, ④ 的证明如下：设矩阵

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

则

$$\mathbf{A}^H = \begin{bmatrix} \bar{a}_{11} & \bar{a}_{21} & \cdots & \bar{a}_{n1} \\ \bar{a}_{21} & \bar{a}_{22} & \cdots & \bar{a}_{n2} \\ \vdots & \vdots & & \vdots \\ \bar{a}_{1n} & \bar{a}_{2n} & \cdots & \bar{a}_{nn} \end{bmatrix}$$

则

$$\mathbf{A}^H \cdot \mathbf{A} = \begin{bmatrix} \sum_{i=1}^n |a_{i1}|^2 & * & * & * \\ * & \sum_{i=1}^n |a_{i2}|^2 & * & * \\ \vdots & \vdots & & \vdots \\ * & * & * & \sum_{i=1}^n |a_{in}|^2 \end{bmatrix}$$

注. 注意, 这里的 * 代表的是计算出来的结果, 但这部分的结果对于证明过程没有意义, 故不进行计算用 * 代替。

仔细观察便能发现, 对角线上的元素分别为: 第一列元素的和的平方、第二列元素的和的平方, 一直到第 n 列元素和的平方, 这恰恰等于矩阵的矩阵 2-范数, 而根据矩阵的迹的计算方法, 矩阵的迹等于矩阵对角线元素的和, 因此, 就有了如下等式

$$\text{tr}(\mathbf{A}^H \mathbf{A}) = \sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2 = \|\mathbf{A}\|_{m2}$$

至此, 性质 1 与 2 证毕。 □

矩阵无穷范数的相容性

矩阵的无穷范数是不相容的, 证明不相容, 只需要给出一个反例即可, 即只需给出两个矩阵 \mathbf{A}, \mathbf{B} , 使得

$$\|\mathbf{AB}\|_{m\infty} \not\leq \|\mathbf{A}\|_{m\infty} \|\mathbf{B}\|_{m\infty}$$

即可

事实上, 若

$$\mathbf{A} = \mathbf{B} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

则

$$\mathbf{AB} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}^2 = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

故

$$\|AB\|_{m_\infty} = 2, \quad \|A\|_{m_\infty} \|B\|_{m_\infty} = 1, \quad \|AB\|_{m_\infty} > \|A\|_{m_\infty} \|B\|_{m_\infty}$$

因此，矩阵的无穷范数不是自相容范数。

2.2.5 酉不变范数

酉不变范数就是定理1.2.1中的性质 3，下面给出其证明：

证明.

$$\begin{aligned} \|UA\|_{m_2}^2 &= \text{tr} [(UA)^H UA] \\ &= \text{tr} [A^H U^H U A] \quad (\text{注意到 } U^H U = E) \\ &= \text{tr} [A^H A] \\ \|AV\|_{m_2}^2 &= \text{tr} [(AV)^H AV] \\ &= \text{tr} [V^H A^H AV] \quad (\text{注意到 } V^H = V^{-1}) \end{aligned}$$

根据矩阵相似的定义，注意到

$$V^{-1} A^H AV \sim A^H A$$

故

$$\text{原式} = \text{tr} [A^H A] \quad (\text{因为矩阵相似，故特征值相等，因此矩阵的迹也相等})$$

证毕。 □

在上面的过程中，分别利用酉矩阵与其共轭转置相乘为单位阵以及酉矩阵与其共轭转置互为逆矩阵从而导出矩阵相似的性质进行推导，将上面的两种情况合成在一起就是定理1.2.1中的第三条的所有。

2.3 算子范数

本章的最后一个知识点为将要讨论的算子范数，算子范数成功的将向量范数与矩阵范数联系在一起，如此便可通过向量范数来定义矩阵的算子范数。

2.3.1 向量范数与矩阵范数的相容关系

在前面的两个小节中，我们讨论了向量范数和矩阵范数，我们在引入矩阵范数的时候，是通过与向量范数相似的方式来定义矩阵范数，可如果将二者放在一起，他们的相容性关系是

如何的呢?

定义 2.3.1: 矩阵范数与向量范数相容

设 $\|\cdot\|_a$ 是 \mathbb{P}^n 上的向量范数, $\|\cdot\|_m$ 是 $\mathbb{P}^{n \times n}$ 上的矩阵范数, 且

$$\|\mathbf{A}\mathbf{x}\|_a \leq \|\mathbf{A}\|_m \|\mathbf{x}\|_a$$

则称 $\|\cdot\|_m$ 为与向量范数 $\|\cdot\|_a$ 相容的矩阵范数

定义1.3.1描述了矩阵范数与向量范数相容的条件, 接下来我们来看两个例子:

例 2.3.1. 设 $\mathbf{x} \in \mathbb{P}^n, \mathbf{A} \in \mathbb{P}^{n \times n}$, 则

$$\|\mathbf{A}\|_{m_1} = \sum_{j=1}^n \sum_{i=1}^n |a_{ij}|$$

是与向量范数 $\|\cdot\|_1$ 相容的矩阵范数。

解. 证明方法与前文证明矩阵 1-范数是自相容范数思路相似, 只需要将 \mathbf{B} 矩阵换为向量 \mathbf{x} 即可, 故不赘述。□

例 2.3.2. 设 $\mathbf{x} \in \mathbb{P}^n, \mathbf{A} \in \mathbb{P}^{n \times n}$, 则 $\|\mathbf{A}\|_{m_2}$ 是与 $\|\mathbf{x}\|_2$ 相容的矩阵范数

解. 证明方法与前文证明矩阵 2-范数是自相容范数思路相似, 只需要将 \mathbf{B} 矩阵换为向量 \mathbf{x} 即可, 故不赘述。□

可是我们知道, 向量范数有很多种 (参考介绍向量范数的 p -范数), 那对于任意一种向量范数, 是否也存在与该向量范数相容的矩阵范数呢? 这就是下面要介绍的算子范数:

2.3.2 算子范数

下面给出算子范数的定义:

定义 2.3.2: 算子范数的定义

设 $\|\mathbf{x}\|_a$ 是 \mathbb{P}^n 上的向量范数, $\mathbf{A} \in \mathbb{P}^{n \times n}$, 则

$$\|\mathbf{A}\|_a = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_a}{\|\mathbf{x}\|_a}$$

是与向量范数 $\|\mathbf{x}\|_a$ 相容的矩阵范数, 并且称该矩阵范数是从属于向量范数 $\|\mathbf{x}\|_a$ 的算子范数。

定义1.3.2说了两件事情：算子范数是与定义它的向量范数是相容的；同时，说了这个公式代表的事算子范数，为什么算子范数是与定义它的向量范数是相容的呢？证明如下：

证明. 由算子范数的定义，有

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

那就会有如下关系：

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

化简一下，就会有

$$\|Ax\|_a \leq \|A\|_a \cdot \|x\|_a$$

也就成功说明了，算子范数是与定义它的向量范数是相容的，证毕。

□

请注意，定义1.3.2还有另外一种形式，即

$$\|A\|_a = \max_{\|u\|_a=1} \|Au\|_a$$

两种形式的定义是怎么来的呢？推导如下：

$$\|Ax\|_a = \left\| A \cdot \frac{x}{\|x\|_a} \right\|_a$$

令 $u = \frac{x}{\|x\|_a}$ ，所以，原式变为

$$\max_{\|u\|_a=1} \|Au\|_a$$

接下来，我们给出算子范数是范数的证明：

1. 正定性证明

若 A 不为 0 ，则存在 $x_0 \neq 0$ ，使得 $Ax_0 \neq 0$

因此， $\|Ax_0\|_a > 0$ ，正定性证毕。

2. 齐次性证明

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

$$= \max_{x \neq 0} \frac{|\lambda| \|Ax\|_a}{\|x\|_a}$$

$$= |\lambda| \max_{x \neq 0} \frac{\|Ax\|_a}{\|x\|_a}$$

$$= |\lambda| \|A\|_a$$

齐次性证毕。

3. 三角不等式证明

$$\begin{aligned}\|A + B\|_a &= \max_{x \neq 0} \frac{\|(A + B)x\|_a}{\|x\|_a} = \max_{x \neq 0} \frac{\|Ax + Bx\|_a}{\|x\|_a} \\ &\leq \max_{x \neq 0} \frac{\|Ax\|_a + \|Bx\|_a}{\|x\|_a}\end{aligned}$$

令 $f(x) = \max_{x \neq 0} \frac{\|Ax\|_a}{\|x\|_a}$, $g(x) = \max_{x \neq 0} \frac{\|Bx\|_a}{\|x\|_a}$, 故

$$\max_{x \neq 0} \frac{\|Ax\|_a + \|Bx\|_a}{\|x\|_a} \leq f(x) + g(x) = \|A\|_a + \|B\|_a$$

至此, 三角不等式证毕。

算子范数的相关定理

定理 2.3.1: 算子范数一定是自相容范数

设 $\|x\|_a$ 是 \mathbb{P}^n 上的向量范数, $A, B \in \mathbb{P}^{n \times n}$, $\|A\|_a$ 是从属于 $\|x\|_a$ 的算子范数, 则它是相容的矩阵范数, 即

$$\|AB\|_a \leq \|A\|_a \cdot \|B\|_a$$

定理1.3.1表明, 任何一个算子范数都是自相容范数, 证明如下:

证明. 由算子范数定义, 有:

$$\begin{aligned}\|AB\|_a &= \max_{x \neq 0} \frac{\|ABx\|_a}{\|x\|_a} \\ &\leq \max_{x \neq 0} \frac{\|A\|_a \cdot \|Bx\|_a}{\|x\|_a} \\ &= \|A\|_a \cdot \max_{x \neq 0} \frac{\|Bx\|_a}{\|x\|_a} = \|A\|_a \cdot \|B\|_a\end{aligned}$$

证毕。 □

上面的定理与讨论当给定向量范数和矩阵时, 就能确定与给定向量范数相容的矩阵范数, 那如果给定矩阵范数时, 能否找到与其相容的向量范数呢?

定理 2.3.2

设 $\|\cdot\|_m$ 是相容的矩阵范数, 则存在向量范数 $\|x\|$, 使得

$$\|Ax\| \leq \|A\|_m \cdot \|x\|$$

定理1.3.2的证明如下：

证明. 定义下列向量范数：

$$\|\mathbf{x}\| = \|\mathbf{x}\mathbf{a}^H\|_m, \quad (\mathbf{a} \neq \mathbf{0})$$

正定性：由于 $\mathbf{x} \neq \mathbf{0}$ ，故 $\mathbf{x}\mathbf{a}^H \neq \mathbf{0}$ ，因此 $\|\mathbf{x}\| = \|\mathbf{x}\mathbf{a}^H\|_m > 0$ ，正定性证毕。

齐次性：

$$\|\lambda\mathbf{x}\| = \|\lambda\mathbf{x}\mathbf{a}^H\|_m$$

右侧的 $\|\lambda\mathbf{x}\mathbf{a}^H\|_m$ 是一个矩阵范数，由矩阵范数的齐次性，有

$$\|\lambda\mathbf{x}\mathbf{a}^H\|_m = |\lambda|\|\mathbf{x}\mathbf{a}^H\|_m$$

齐次性证毕。

三角不等式：

$$\begin{aligned} \|\mathbf{x}_1 + \mathbf{x}_2\| &= \|(\mathbf{x}_1 + \mathbf{x}_2)\mathbf{a}^H\|_m \\ &= \|\mathbf{x}_1\mathbf{a}^H + \mathbf{x}_2\mathbf{a}^H\|_m \end{aligned}$$

由矩阵范数的三角不等式，有

$$\|\mathbf{x}_1\mathbf{a}^H + \mathbf{x}_2\mathbf{a}^H\|_m \leq \|\mathbf{x}_1\mathbf{a}^H\|_m + \|\mathbf{x}_2\mathbf{a}^H\|_m = \|\mathbf{x}_1\| + \|\mathbf{x}_2\|$$

三角不等式证毕，因此，我们可以通过给定的矩阵范数找到其相容的向量范数 □

矩阵范数与特征值之间的关系

定理 2.3.3

如果 $\|\cdot\|_m : \mathbb{C}^{n \times n} \rightarrow \mathbb{R}$ 是一相容的矩阵范数，则对任一 $\mathbf{A} \in \mathbb{C}^{n \times n}$ ，有

$$|\lambda_i| \leq \|\mathbf{A}\|_m$$

其中， λ_i 是 \mathbf{A} 的特征值。

定理1.3.3指出了矩阵范数和矩阵特征值之间的关系，由此我们可以矩阵特征值来判断矩阵范数，同理也可以根据矩阵范数判断矩阵特征值。

定理1.3.3的证明如下：

证明.

根据特征值的定义，对于矩阵 \mathbf{A} 的任一特征值 λ_i ，有

$$\mathbf{A}\mathbf{x} = \lambda_i\mathbf{x}$$

因此，有

$$\|\mathbf{A}\mathbf{x}\| = \|\lambda_i\mathbf{x}\| = |\lambda_i|\|\mathbf{x}\|$$

由定理1.3.2, 若矩阵范数相容，则必定存在一个向量的向量范数与之相容，故，有

$$\|\mathbf{A}\mathbf{x}\| \leq \|\mathbf{A}\|_m \|\mathbf{x}\|$$

因此，就有

$$|\lambda_i|\|\mathbf{x}\| \leq \|\mathbf{A}\|_m \|\mathbf{x}\|$$

由于 $\|\mathbf{x}\| \neq 0$ ，因此，两侧可以约掉 \mathbf{x} ，于是式子就变成了

$$|\lambda_i| \leq \|\mathbf{A}\|_m$$

证毕。 □

2.3.3 算子范数的计算

本小节着重介绍算子范数的计算，由于特殊的定义关系，有时算子范数并不一定能够计算出来，本小节着重介绍算子 1-范数，算子 2-范数（也叫谱范数）和算子无穷范数的计算。

算子 1-范数的计算

从属于向量范数 $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$ 的算子范数为

$$\|\mathbf{A}\|_1 = \max_j \sum_{i=1}^n |a_{ij}|$$

又称为**极大列和范数**。

这个公式的意思是，首先分别计算出每一列的绝对列和出来，然后再找出来最大的那个绝对列和，这就是 \max_j 的含义，而不是找到数字最大的那一列。

如何证明这个等式是正确的呢？往常证明等式成立往往是通过从等式的左侧或右侧出发，经过一系列的变形之后推导到了等式的右侧/左侧，证明了等式的成立，但在本节证明等式成立中，我们并不会采用这种方法，而是首先假设等式的左侧小于等于右侧，证明其成立，而后假设等式的左侧大于等于右侧，证明其成立，两者一结合便只有相等这一种选项。

证明.

假设 $\|\mathbf{A}\|_1 \leq \max_j \sum_{i=1}^n |a_{ij}|$, 首先证明该不等式的成立。

根据矩阵与向量的乘法, 有

$$\|\mathbf{Ax}\|_1 = \sum_{i=1}^n |a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n|$$

根据绝对值三角不等式, 有如下关系

$$\sum_{i=1}^n |a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n| \leq \sum_{i=1}^n \sum_{j=1}^n |a_{ij}| |x_j|$$

由于多重累加可以交换顺序, 因此上面的式子可以变成

$$\sum_{j=1}^n \sum_{i=1}^n |a_{ij}| |x_j|$$

此时, 注意到 $\sum_{i=1}^n |a_{ij}| |x_j|$ 中的 $|x_j|$ 其实并不会受 j 的变化而变化, 因此可以当成公因式提出, 于是式子就成了

$$\sum_{j=1}^n (|x_j| \sum_{i=1}^n |a_{ij}|)$$

仔细观察 $\sum_{i=1}^n |a_{ij}|$, 不难发现, 这其实就是矩阵 \mathbf{A} 的绝对列和, 接下来, 对式子做类似证明矩阵范数相容性那样进行放缩, 有

$$\begin{aligned} \sum_{j=1}^n (|x_j| \sum_{i=1}^n |a_{ij}|) &\leq \sum_{j=1}^n (|x_j| \cdot \max_j \sum_{i=1}^n |a_{ij}|) \\ &= \max_j \sum_{i=1}^n |a_{ij}| \cdot \sum_{j=1}^n |x_j| \end{aligned}$$

注意到 $\sum_{j=1}^n |x_j|$ 其实就是向量 \mathbf{x} 的向量 1-范数, 因此有

$$\max_j \sum_{i=1}^n |a_{ij}| \cdot \sum_{j=1}^n |x_j| = \max_j \sum_{i=1}^n |a_{ij}| \cdot \|\mathbf{x}\|_1$$

整体的式子是这样的:

$$\|\mathbf{Ax}\|_1 \leq \max_j \sum_{i=1}^n |a_{ij}| \cdot \|\mathbf{x}\|_1$$

移项, 于是就有了

$$\frac{\|\mathbf{Ax}\|_1}{\|\mathbf{x}\|_1} \leq \max_j \sum_{i=1}^n |a_{ij}|$$

或许会有人会发现, 左侧 $\frac{\|\mathbf{Ax}\|_1}{\|\mathbf{x}\|_1}$ 缺少了 $\max_{\mathbf{x} \neq \mathbf{0}}$, 这是为什么呢?

因为在证明的过程中, 我们假设的是对任意的 $\mathbf{x} \neq \mathbf{0}$ 都成立, 因此不必加上 $\max_{\mathbf{x} \neq \mathbf{0}}$

很好, 我们顺利证明出了 $\|\mathbf{A}\|_1 \leq \max_j \sum_{i=1}^n |a_{ij}|$, 接下来继续证明 $\|\mathbf{A}\|_1 \geq \max_j \sum_{i=1}^n |a_{ij}|$

由于涉及到某一列的绝对列和最大, 所以不妨假设矩阵 \mathbf{A} 的第 s 列的绝对列和最大, 即矩阵 \mathbf{A} 的形式为

$$\mathbf{A} = [\alpha_1, \alpha_2, \dots, \alpha_s, \dots, \alpha_n]$$

此时, 取特殊向量 $\mathbf{x}_0 = \epsilon_s$, 这里的 ϵ_s 代表的是单位矩阵 \mathbf{E} 的第 s 列的列向量
由向量范数以及单位矩阵的性质, 不难得出下列等式关系

$$\|\mathbf{x}_0\|_1 = \|\epsilon_s\|_1 = 1$$

左乘一个矩阵 \mathbf{A} , 有

$$\|\mathbf{A}\mathbf{x}_0\|_1 = \|\mathbf{A}\epsilon_s\|_1 = \|\alpha_s\|_1 = \max_j \sum_{i=1}^n |a_{ij}|$$

由于 $\|\mathbf{x}_0\|_1 = 1$, 因此对于上面的等式同时除以 $\|\mathbf{x}_0\|_1$, 对式子结果不影响, 但我们得到了

$$\frac{\|\mathbf{A}\mathbf{x}_0\|_1}{\|\mathbf{x}_0\|_1} = \max_j \sum_{i=1}^n |a_{ij}|$$

此时请注意, 由算子范数最开始的定义, 算子范数其实是需要找出

$$\max_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_1}{\|\mathbf{x}\|_1}$$

这个值的范围是限定在整个矩阵中的, 而非仅仅某一列或某一行, 这是一种全局最大, 但是看上面的证明, 我们仅仅是找到了矩阵的某一行最大, 是一种局部的最大值, 全局最大值一定是大于等于局部最大值的, 因此, 我们有

$$\|\mathbf{A}\|_1 = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_1}{\|\mathbf{x}\|_1} \geq \frac{\|\mathbf{A}\mathbf{x}_0\|_1}{\|\mathbf{x}_0\|_1} = \max_j \sum_{i=1}^n |a_{ij}|$$

由此, 我们证明了 $\|\mathbf{A}\|_1 \geq \max_j \sum_{i=1}^n |a_{ij}|$, 结合上面已经证明出的 $\|\mathbf{A}\|_1 \leq \max_j \sum_{i=1}^n |a_{ij}|$, 我们就可以说, $\|\mathbf{A}\|_1 = \max_j \sum_{i=1}^n |a_{ij}|$, 证明完毕。

□

算子无穷范数的计算

从属于 $\|\mathbf{x}\|_\infty$ 的算子范数为

$$\|\mathbf{A}\|_\infty = \max_i \left(\sum_{j=1}^n |a_{ij}| \right)$$

也被称作**极大行和范数**.

与证明算子 1-范数的方式类似, 我们依旧需要构建出两个不等式: $\|\mathbf{A}\|_\infty = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{Ax}\|_\infty}{\|\mathbf{x}\|_\infty} \geq \max_i (\sum_{j=1}^n |a_{ij}|)$ 和 $\|\mathbf{A}\|_\infty = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{Ax}\|_\infty}{\|\mathbf{x}\|_\infty} \leq \max_i (\sum_{j=1}^n |a_{ij}|)$, 并依次证明成立, 最后推出不等关系, 证明过程如下:

证明. 首先证明

$$\|\mathbf{A}\|_\infty = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{Ax}\|_\infty}{\|\mathbf{x}\|_\infty} \leq \max_i \left(\sum_{j=1}^n |a_{ij}| \right)$$

与证明算子 1-范数的过程类似, 有

$$\|\mathbf{Ax}\|_\infty = \max_i |a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n|$$

根据绝对值三角不等式, 有如下关系

$$\max_i |a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n| \leq \max_i \left(\sum_{j=1}^n |a_{ij}| |x_j| \right)$$

仔细观察 $\sum_{j=1}^n |a_{ij}|$, 不难发现, 这其实就是矩阵 \mathbf{A} 的绝对行和, 接下来, 对式子做类似证明矩阵范数相容性那样进行放缩, 有

$$\begin{aligned} \max_i \left(\sum_{j=1}^n |a_{ij}| |x_j| \right) &\leq \max_i \left(\sum_{j=1}^n |a_{ij}| \cdot \max_j |x_j| \right) \\ &= \max_i \sum_{j=1}^n |a_{ij}| \cdot \max_j |x_j| \\ &= \left(\max_i \sum_{j=1}^n |a_{ij}| \right) \cdot \|\mathbf{x}\|_\infty \end{aligned}$$

整体的式子是这样的:

$$\|\mathbf{Ax}\|_\infty \leq \max_i \sum_{j=1}^n |a_{ij}| \cdot \|\mathbf{x}\|_\infty$$

移项, 于是就有了

$$\frac{\|\mathbf{Ax}\|_\infty}{\|\mathbf{x}\|_\infty} \leq \max_i \sum_{j=1}^n |a_{ij}|$$

故, 只要有 $\mathbf{x} \neq \mathbf{0}$, 就会有上式成立

接下来, 证明

$$\max_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{Ax}\|_\infty}{\|\mathbf{x}\|_\infty} \geq \max_i \left(\sum_{j=1}^n |a_{ij}| \right)$$

但与假设最大绝对列和不同的是, 我们其实并没有办法通过简单的一个矩阵乘一个行的形式来构造绝对行和, 为此, 我们采取下面的方式:

令 $\mu = \sum_{j=1}^n |a_{sj}| = \max_i \sum_{j=1}^n |a_{ij}|$, 其中 $a_{sj} = |a_{sj}| \cdot e^{i\theta_j}$ ($j = 1, 2, \dots, n$)¹

由于复数表示的性质, 可以知道 $e^{i\theta_j}$ 的值是永远小于等于 1 的, 构造一个特殊的向量 \mathbf{z} , 其形式为

$$\mathbf{z} = (e^{-i\theta_1}, e^{-i\theta_2}, \dots, e^{-i\theta_n})$$

由欧拉公式可知, $\|e^{i\theta}\| = 1$, 故 $\|\mathbf{z}\|_\infty = 1$

故 $\|\mathbf{Az}\|_\infty = \max_i |a_{i1}e^{-i\theta_1} + \dots + a_{in}e^{-i\theta_n}|$

假设矩阵 \mathbf{A} 的第 s 行的绝对行和最大, 即

$$\sum_{j=1}^n |a_{sj}| = \max_i \sum_{j=1}^n |a_{ij}|$$

于是我们有

$$\|\mathbf{Az}\|_\infty = \max_i |a_{i1}e^{-i\theta_1} + \dots + a_{in}e^{-i\theta_n}| \geq |a_{s1}e^{-i\theta_1} + \dots + a_{sn}e^{-i\theta_n}|$$

为什么这个不等关系会成立呢, 因为 $\max_i |a_{i1}e^{-i\theta_1} + \dots + a_{in}e^{-i\theta_n}|$ 是所有分量的模的最大值, 而 $|a_{s1}e^{-i\theta_1} + \dots + a_{sn}e^{-i\theta_n}|$ 仅仅只是其中一个分量。

注意到

$$a_{si}e^{-i\theta_i} = |a_{si}|e^{i\theta_i} \cdot e^{-i\theta_i} = |a_{si}|$$

所以

$$|a_{s1}e^{-i\theta_1} + \dots + a_{sn}e^{-i\theta_n}| = |a_{s1}| + |a_{s2}| + \dots + |a_{sn}| = \max_i \sum_{j=1}^n |a_{ij}|$$

故

$$\max_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{Ax}\|_\infty}{\|\mathbf{x}\|_\infty} \geq \frac{\|\mathbf{Az}\|_\infty}{\|\mathbf{z}\|_\infty} = \max_i \sum_{j=1}^n |a_{ij}|$$

证毕。 □

算子 2-范数 (谱范数) 的计算

从属于 $\|\mathbf{x}\|_2$ 的算子范数 (又称为谱范数) 为

$$\|\mathbf{A}\|_2 = \sqrt{r(\mathbf{A}^H \mathbf{A})}$$

¹这里涉及到了复数的表示形式和欧拉公式的内容, 欧拉公式的表示为 $e^{i\theta} = \cos \theta + i \sin \theta$, 根据该公式, 复数还可以使用 $z = \rho(\cos \theta + i \sin \theta)$ 的形式表示。

这里出现了一个新的东西： $r(\mathbf{A}^H \mathbf{A})$ ，这是什么东西呢，这个东西表示 $\mathbf{A}^H \mathbf{A}$ 的谱半径，谱半径的计算方法为：

$$r(\mathbf{A}) = \max_i |\lambda_i|$$

即，一个矩阵的谱半径的值就是其特征向量绝对值的最大值。

利用算子范数的定义，即算子 2-范数的计算公式为

$$\|\mathbf{A}\|_2 = \max_{\|\mathbf{u}\|_2=1} \|\mathbf{A}\mathbf{u}\|_2 = \sqrt{r(\mathbf{A}^H \mathbf{A})}$$

这里，依旧只需要证明

$$\max_{\|\mathbf{u}\|_2=1} \|\mathbf{A}\mathbf{u}\|_2 \geq \sqrt{r(\mathbf{A}^H \mathbf{A})}$$

和

$$\max_{\|\mathbf{u}\|_2=1} \|\mathbf{A}\mathbf{u}\|_2 \leq \sqrt{r(\mathbf{A}^H \mathbf{A})}$$

即可

这里要简单介绍一下 Hermite 矩阵的概念

定义 2.3.3: Hermite 矩阵

设 $\mathbf{A} \in \mathbb{C}^{n \times n}$ ，若 $\mathbf{A}^H = \mathbf{A}$ ，则称 \mathbf{A} 是 Hermite 矩阵；若 $\mathbf{A}^H = -\mathbf{A}$ ，则称 \mathbf{A} 是反 Hermite 矩阵。

Hermite 矩阵还拥有如下性质 (这里只写后面需要用到的，具体更多的性质请看后面的相关内容)：

- \mathbf{A} 是正定矩阵
- \mathbf{A} 的特征值全为正实数

因此，对于这里的谱范数，我们不妨设矩阵 $\mathbf{A}^H \mathbf{A}$ 的特征值有如下关系：

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$$

这里 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 分别为单位特征向量 x_1, x_2, \cdots, x_n 对应的特征值，又由于矩阵 $\mathbf{A}^H \mathbf{A}$ 是 Hermite 矩阵，故其所有的特征向量均正交。

故

$$r(\mathbf{A}^H \mathbf{A}) = \lambda_1$$

任取 $\mathbf{u} = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$ ，并且令 $\|\mathbf{u}\|_2^2 = 1$

因此, 有

$$\|\mathbf{u}\|_2^2 = 1$$

证明. 先证

$$\max_{\|\mathbf{u}\|_2=1} \|\mathbf{A}\mathbf{u}\|_2 \leq \sqrt{r(\mathbf{A}^H \mathbf{A})}$$

由定义, 有

$$\begin{aligned} \|\mathbf{A}\mathbf{u}\|_2^2 &= \mathbf{u}^H \mathbf{A}^H \mathbf{A} \mathbf{u} \\ &= \mathbf{u}^H \mathbf{A}^H \mathbf{A} (a_1 x_1 + a_2 x_2 + \cdots + a_n x_n) \\ &= \mathbf{u}^H (a_1 \mathbf{A}^H \mathbf{A} x_1 + a_2 \mathbf{A}^H \mathbf{A} x_2 + \cdots + a_n \mathbf{A}^H \mathbf{A} x_n) \\ &= \mathbf{u}^H (a_1 \lambda_1 x_1 + a_2 \lambda_2 x_2 + \cdots + a_n \lambda_n x_n) \\ &= (\bar{a}_1 x_1^H + \cdots + \bar{a}_n x_n^H) \cdot (a_1 \lambda_1 x_1 + a_2 \lambda_2 x_2 + \cdots + a_n \lambda_n x_n) \\ &= \lambda_1 |a_1|^2 + \lambda_2 |a_2|^2 + \cdots + \lambda_n |a_n|^2 \end{aligned}$$

由于最开始的特征值之间的大小关系, 我们有:

$$\begin{aligned} \lambda_1 |a_1|^2 + \lambda_2 |a_2|^2 + \cdots + \lambda_n |a_n|^2 &\leq \lambda_1 (|a_1|^2 |a_2|^2 + \cdots + |a_n|^2) \\ &= \lambda_1 \|\mathbf{u}\|_2^2 = \lambda_1 \end{aligned}$$

即

$$\|\mathbf{A}\mathbf{u}\|_2^2 \leq \lambda_1 = r(\mathbf{A}^H \mathbf{A})$$

接下来证明

$$\max_{\|\mathbf{u}\|_2=1} \|\mathbf{A}\mathbf{u}\|_2 \geq \sqrt{r(\mathbf{A}^H \mathbf{A})}$$

由特征值和特征向量的定义, 有

$$\mathbf{A}^H \mathbf{A} \mathbf{x}_1 = \lambda_1 \mathbf{x}_1$$

等式左右两侧同时乘 \mathbf{x}_1^H , 有

$$\mathbf{x}_1^H \mathbf{A}^H \mathbf{A} \mathbf{x}_1 = \lambda_1 \mathbf{x}_1^H \mathbf{x}_1 = \lambda_1$$

即

$$\max_{\|\mathbf{u}\|_2=1} \|\mathbf{A}\mathbf{u}\|_2^2 \geq \|\mathbf{A}\mathbf{x}_1\|_2^2 = \mathbf{x}_1^H \mathbf{A}^H \mathbf{A} \mathbf{x}_1 = \mathbf{x}_1^H \lambda_1 \mathbf{x}_1 = \lambda_1$$

证毕。

(数学真好玩, 证明真有趣, 我要天天学数学数学)

□

谱范数的其他性质

首先介绍谱范数的酉不变性质

定理 2.3.4: 酉不变性质

设 $\mathbf{A} \in \mathbb{C}^{n \times n}$, 则

1. $\|\mathbf{A}\|_2 = \|\mathbf{A}^H\|_2 = \|\mathbf{A}^T\|_2 = \|\bar{\mathbf{A}}\|_2$
2. $\|\mathbf{A}^H \mathbf{A}\|_2 = \|\mathbf{A} \mathbf{A}^H\|_2 = \|\mathbf{A}\|_2^2$
3. 对任何 n 阶酉矩阵 \mathbf{U} 和 \mathbf{V} , 都有

$$\|\mathbf{U} \mathbf{A}\|_2 = \|\mathbf{A} \mathbf{V}\|_2 = \|\mathbf{U} \mathbf{A} \mathbf{V}\|_2 = \|\mathbf{A}\|_2$$

上面三个性质的证明如下:

证明.

性质 1:

$$\|\mathbf{A}\|_2 = \sqrt{r(\mathbf{A}^H \mathbf{A})} \text{ 因此, 有 } \|\mathbf{A}^H\|_2 = \sqrt{r[(\mathbf{A}^H)^H \mathbf{A}^H]} = \sqrt{r(\mathbf{A} \mathbf{A}^H)} = \sqrt{r(\mathbf{A}^H \mathbf{A})} = \|\mathbf{A}\|_2$$

$$\|\mathbf{A}^T\|_2 = \sqrt{r[(\mathbf{A}^T)^H \mathbf{A}^T]} = \sqrt{r(\mathbf{A} \mathbf{A}^H)^T} = \sqrt{r(\mathbf{A} \mathbf{A}^H)} = \|\mathbf{A}^H\|_2$$

由于 $\bar{\mathbf{A}} = (\mathbf{A}^T)^H$, 矩阵的转置与求共轭均不会影响矩阵的范数, 因此显然相等, 性质 1 证明完毕。

性质 2:

$$\|\mathbf{A}^H \mathbf{A}\|_2 = \sqrt{r[(\mathbf{A}^H \mathbf{A})^H \mathbf{A}^H \mathbf{A}]} = \sqrt{r(\mathbf{A}^H \mathbf{A})^2} = r(\mathbf{A}^H \mathbf{A}) = \|\mathbf{A}\|_2^2$$

性质 3:

性质 3 的证明方式与矩阵范数中的酉不变范数的证明方式类似, 在这里不再赘述

□

接下来介绍谱范数的下一条性质:

定理 2.3.5

设 $\mathbf{A} \in \mathbb{C}^{n \times n}$, 则

- $\|\mathbf{A}\|_2 = \max_{\|\mathbf{x}\|_2=\|\mathbf{y}\|_2=1} |\mathbf{y}^H \mathbf{A} \mathbf{x}|$
- $\|\mathbf{A}\|_2^2 \leq \|\mathbf{A}\|_1 \|\mathbf{A}\|_\infty$

定理1.3.5的证明如下：

对于第一条性质，采用的证明思路是这样的：

1. 首先证明右侧的值是一个上界
2. 其次证明这个上界是可以达到的

首先证明第一项，由 Cauchy 不等式，有

$$|\mathbf{y}^H(\mathbf{Ax})| = |\langle \mathbf{y}, \mathbf{Ax} \rangle| \leq \|\mathbf{y}\|_2 \cdot \|\mathbf{Ax}\|_2$$

由范数相容性可知

$$\|\mathbf{Ax}\|_2 \leq \|\mathbf{A}\|_2 \|\mathbf{x}\|_2$$

因此，不等式左右两侧同乘 $\|\mathbf{y}\|_2$ ，有

$$\|\mathbf{y}\|_2 \cdot \|\mathbf{Ax}\|_2 \leq \|\mathbf{y}\|_2 \cdot \|\mathbf{A}\|_2 \|\mathbf{x}\|_2$$

由于 $\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 = 1$ 因此

$$|\langle \mathbf{y}, \mathbf{Ax} \rangle| \leq \|\mathbf{A}\|_2$$

证明了右侧的确是一个上界，接下来证明该上界可以达到

由算子范数的定义

$$\|\mathbf{A}\|_2 = \max_{\|\mathbf{x}\|_2=1} \|\mathbf{Ax}\|_2$$

因此， $\exists x_0$ ，有 $\|\mathbf{A}\|_2 = \|\mathbf{Ax}_0\|_2$

对 \mathbf{Ax}_0 进行单位化，有 $\mathbf{y}_0 = \frac{\mathbf{Ax}_0}{\|\mathbf{Ax}_0\|_2}$

因此，

$$|\mathbf{y}_0^H \cdot \mathbf{Ax}_0| = \left| \frac{(\mathbf{Ax}_0)^H \mathbf{Ax}_0}{\|\mathbf{Ax}_0\|_2} \right|$$

分子恰好是 \mathbf{Ax}_0 的内积，内积同时也是向量 2-范数，故

$$\left| \frac{(\mathbf{Ax}_0)^H \mathbf{Ax}_0}{\|\mathbf{Ax}_0\|_2} \right| = \frac{\|\mathbf{Ax}_0\|_2^2}{\|\mathbf{Ax}_0\|_2} = \|\mathbf{Ax}_0\|_2 = \|\mathbf{A}\|_2$$

如此，证明可以达到这个上界

因此

$$\max_{\|\mathbf{x}\|_2=\|\mathbf{y}\|_2=1} |\mathbf{y}^H \mathbf{Ax}| = \|\mathbf{A}\|_2$$

第二条性质的证明：

$$\|\mathbf{A}\|_2^2 = r(\mathbf{A}^H \mathbf{A})$$

由于 $r(\mathbf{A}^H \mathbf{A}) = \max_i |\lambda_i|$ ，故根据定理1.3.3，有

$$r(\mathbf{A} \mathbf{A}^H) \leq \|\mathbf{A}^H\|_\infty \|\mathbf{A}\|_\infty$$

这里请回忆一下计算算子无穷范数的时候给出的计算方式，不难发现这里与标准算子无穷范数相比，多了一个转置共轭，算子无穷范数又叫行和范数，转置一下就变成了列和范数，这恰好就是算子 1-范数的形式，因此，上式可以变为

$$\|\mathbf{A}\|_1 \|\mathbf{A}\|_\infty$$

即

$$\|\mathbf{A}\|_2^2 = r(\mathbf{A}^H \mathbf{A}) \leq \|\mathbf{A}\|_1 \|\mathbf{A}\|_\infty$$

证毕。

2.4 酉不变范数

考试不考，上课没讲，略。

2.5 矩阵的测度

考试不考，上课没讲，略。

2.6 范数的应用

这一章虽然老师讲了，但是考试不考，主要就是希望各位能够培养一个严谨的精神，一个小小的误差可能会导致求出的结果与理想结果差距甚远，学习方法只需要看老师这一节 PPT 的几个例子即可。