MATH10003.01 高等线性代数 Fall 2025

Lecture 3 范数（二） — 2025.09.23

教授：邵美悦

Scribe: 卤肉卷

一整节课只记住了 *Gelfand* 是苏联人。

未经校对，如有错误，欢迎指正。

1 课程纲要

1. l_p 向量范数
2. 矩阵范数的定义
3. 算子范数
4. l_p 矩阵范数
5. Frobenius 范数
6. 赋范线性空间

2 l_p 向量范数

上节课我们介绍了 l_p 向量范数的定义：

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}, \quad 1 \le p < \infty,$$

但我们并没有解释为什么它是一个范数。今天我们对其进行补充。

首先，我们要先介绍几个常见的不等式，它对我们的证明是至关重要的。

引理 2.1. （*Jensen* 不等式）若 f 在区间 $[a,b]$ 上是一个上凸函数，则对于任意 $x_1, x_2, \ldots, x_n \in [a,b]$ 和非负实数 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 满足 $\sum_{i=1}^{n} \alpha_i = 1$，有

$$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \le \sum_{i=1}^{n} \alpha_i f(x_i),$$

当且仅当 $x_1 = x_2 = \cdots = x_n$ 时等号成立。

这个不等式大家在数学分析中都学习过，我们就不加以证明了。

引理 2.2. （*Young 不等式*）$\forall a, b \geq 0, p, q > 1$，有

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}, \quad 其中\ \frac{1}{p} + \frac{1}{q} = 1.$$

证明. 当 $a = 0$ 或 $b = 0$ 时，显然成立。下面我们假设 $a, b > 0$。幂函数的形式不够美观，我们对两边取对数，只需证明

$$\ln a + \ln b \leq \ln(\frac{a^p}{p} + \frac{b^q}{q}).$$

观察不等号右边的式子结构：

$$\frac{a^p}{p} + \frac{b^q}{q} = \frac{1}{p}a^p + \frac{1}{q}b^q.$$

这正好是一个加权平均数的形式。我们能否把左边也变成一个加权平均数的形式呢？事实上，

$$\mathrm{LHS} = \frac{1}{p}(p \ln a) + \frac{1}{q}(q \ln b) = \frac{1}{p}\ln(a^p) + \frac{1}{q}\ln(b^q).$$

这样，不等号两边形式就得到了统一。接下来，我们只需证明

$$\frac{1}{p}\ln(a^p) + \frac{1}{q}\ln(b^q) \leq \ln(\frac{1}{p}a^p + \frac{1}{q}b^q).$$

这正好是 Jensen 不等式的形式。注意到 $\ln x$ 在 $(0, +\infty)$ 上是一个上凸函数，取 $x_1 = a^p, x_2 = b^q, \alpha_1 = \frac{1}{p}, \alpha_2 = \frac{1}{q}$ 即可得到上面的不等式。当且仅当 $a^p = b^q$ 时等号成立。 □

引理 2.3. （*Cauchy-Schwarz 不等式*）对于任意的 $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$，有

$$|\boldsymbol{x}^\star \boldsymbol{y}| \leq \|\boldsymbol{x}\|_2 \|\boldsymbol{y}\|_2.$$

证明. 当 $\boldsymbol{x} = \boldsymbol{0}$ 或 $\boldsymbol{y} = \boldsymbol{0}$ 时，显然成立。下面我们假设 $\boldsymbol{x} \neq \boldsymbol{0}$ 且 $\boldsymbol{y} \neq \boldsymbol{0}$. 我们先证明 $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ 的情形。考虑函数

$$\begin{aligned} f(\lambda) &= \|\lambda \boldsymbol{x} - \boldsymbol{y}\|_2^2 \\ &= (\lambda \boldsymbol{x} - \boldsymbol{y})^{\mathrm{T}}(\lambda \boldsymbol{x} - \boldsymbol{y}) \\ &= \lambda^2 \|\boldsymbol{x}\|_2^2 - 2\lambda \boldsymbol{x}^\star \boldsymbol{y} + \|\boldsymbol{y}\|_2^2. \end{aligned}$$

显然 $f(\lambda) \geq 0$，因此它的判别式一定小于等于零，即

$$\Delta = 4(\boldsymbol{x}^\star \boldsymbol{y})^2 - 4\|\boldsymbol{x}\|_2^2 \|\boldsymbol{y}\|_2^2 \leq 0,$$

化简后即有

$$(\boldsymbol{x}^\star \boldsymbol{y})^2 \leq \|\boldsymbol{x}\|_2^2 \|\boldsymbol{y}\|_2^2,$$

当且仅当 $\boldsymbol{x}, \boldsymbol{y}$ 线性相关或有一方为零时等号成立。取正平方根即为 Cauchy-Schwarz 不等式。

下面我们证明 $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ 的情形。我们要如何将实数域上的结论推广到复数域上呢？事实上，我们可以对 $\boldsymbol{x}, \boldsymbol{y}$ 进行旋转变换，若设 $\boldsymbol{z}, \boldsymbol{w}$ 分别是旋转后的复数，我们的目标就是让 $\boldsymbol{z}, \boldsymbol{w} \in \mathbb{R}^n$. 我们定义

$$\begin{cases} z_i = e^{\mathrm{i}\theta_i} x_i \in [0, +\infty), \\ w_i = e^{\mathrm{i}\psi_i} y_i \in [0, +\infty), \end{cases} \qquad i = 1, 2, \ldots, n,$$

这样变换之后，我们可以发现：

$$|x_i y_i| = |z_i w_i|, \quad \|x\|_2 = \|z\|_2, \quad \|y\|_2 = \|w\|_2.$$

因此，根据实数域上的结论，我们有

$$\begin{aligned} |\boldsymbol{x}^\star \boldsymbol{y}| &\le \sum_{i=1}^n |x_i y_i| \\ &= \sum_{i=1}^n |z_i w_i| \\ &= \boldsymbol{z}^{\mathrm{T}} \boldsymbol{w} \\ &\le \|\boldsymbol{z}\|_2 \|\boldsymbol{w}\|_2 \\ &= \|\boldsymbol{x}\|_2 \|\boldsymbol{y}\|_2. \end{aligned}$$

□

根据 Cauchy-Schwarz 不等式，我们可以很容易地证明 2-范数的三角不等式。

定理 2.4. （*2-范数的三角不等式*）*对于任意的* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$，*有*

$$\|\boldsymbol{x} + \boldsymbol{y}\|_2 \le \|\boldsymbol{x}\|_2 + \|\boldsymbol{y}\|_2.$$

证明. 2-范数带着根号，不方便我们操作，因此我们先对其平方。注意到

$$\begin{aligned} \|\boldsymbol{x} + \boldsymbol{y}\|_2^2 &= (\boldsymbol{x} + \boldsymbol{y})^\star (\boldsymbol{x} + \boldsymbol{y}) \\ &= \|\boldsymbol{x}\|_2^2 + \|\boldsymbol{y}\|_2^2 + 2\mathrm{Re}(\boldsymbol{x}^\star \boldsymbol{y}) \\ &\le \|\boldsymbol{x}\|_2^2 + \|\boldsymbol{y}\|_2^2 + 2|\boldsymbol{x}^\star \boldsymbol{y}| \\ &\le \|\boldsymbol{x}\|_2^2 + \|\boldsymbol{y}\|_2^2 + 2\|\boldsymbol{x}\|_2 \|\boldsymbol{y}\|_2 \\ &= (\|\boldsymbol{x}\|_2 + \|\boldsymbol{y}\|_2)^2. \end{aligned}$$

第一个不等号我们利用了 $\mathrm{Re}(z) \le |z|$，第二个不等号我们利用了 Cauchy-Schwarz 不等式。最后两边取正平方根即得所需结论。当且仅当 $\boldsymbol{x}, \boldsymbol{y}$ 线性相关或有一方为零时等号成立。 □

Cauchy-Schwarz 不等式有一个推广形式，即 Hölder 不等式。

定理 2.5. （*Hölder* 不等式）对于任意的 $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ 和 $p, q > 1$ 满足 $\frac{1}{p} + \frac{1}{q} = 1$，有

$$|\boldsymbol{x}^\star \boldsymbol{y}| \le \|\boldsymbol{x}\|_p \|\boldsymbol{y}\|_q.$$

证明. 我们可以注意到，在待证的不等式中，如果对向量 $\boldsymbol{x}$ 或 $\boldsymbol{y}$ 进行缩放，得到的不等式与原不等式是完全等价的。因此，我们可以只证明 $\|\boldsymbol{x}\|_p = 1, \|\boldsymbol{y}\|_q = 1$ 的情形。我们还可以注意到，我们可以对 $\boldsymbol{x}, \boldsymbol{y}$ 进行旋转变换，使得旋转后的向量 $\boldsymbol{z}, \boldsymbol{w}$ 满足 $z_i, w_i \in [0, +\infty)$，得到的不等式与原不等式是等价的。因此，结合以上两个观察，我们只需要证明满足 $\|\boldsymbol{x}\|_p = 1, \|\boldsymbol{y}\|_q = 1$ 的实向量 $\boldsymbol{x}, \boldsymbol{y}$ 有如下不等式成立即可：

$$\boldsymbol{x}^{\mathrm{T}} \boldsymbol{y} \le 1.$$

然后我们可以对这两个向量进行旋转和伸缩变换，便可将结论扩展到整个 $\mathbb{C}^n$ 空间。

我们将待证不等式写成分量形式，有

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

现在，我们的目标就是让左边的 x_i, y_i 彼此分离，且让 x_i 带上 p 次幂，y_i 带上 q 次幂。回想 Young 不等式，我们发现它正好可以满足我们的需求。对每一项 $x_i y_i$ 应用 Young 不等式，有

$$x_i y_i \le \frac{x_i^p}{p} + \frac{y_i^q}{q}.$$

两边求和，有

$$\begin{aligned}
\sum_{i=1}^n x_i y_i &\le \sum_{i=1}^n \left(\frac{x_i^p}{p} + \frac{y_i^q}{q}\right) \\
&= \frac{1}{p}\sum_{i=1}^n x_i^p + \frac{1}{q}\sum_{i=1}^n y_i^q \\
&= \frac{1}{p}\|\boldsymbol{x}\|_p^p + \frac{1}{q}\|\boldsymbol{y}\|_q^q \\
&= \frac{1}{p} + \frac{1}{q} \\
&= 1.
\end{aligned}$$

当且仅当 $\exists \lambda$ 使得 $x_i^p = \lambda y_i^q$ 对任意 i 成立时等号成立。

当我们将结论扩展到整个 $\mathbb{C}^n$ 空间时，等号成立的条件变为 $\exists \lambda$ 使得 $|x_i|^p = \lambda |y_i|^q$ 对任意 i 成立。

□

根据 Hölder 不等式，我们可以证明 l_p 范数的三角不等式。

定理 2.6. （*Lp* 范数的三角不等式，*Minkowski* 不等式）对于任意的 $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ 和 $p \ge 1$，有

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

证明. 我们将左边写成分量形式，有

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

上面的不等号利用了复数的三角不等式。这里出现了求和，还出现了 p 次方，一个自然的想法就是利用 Hölder 不等式对其进行放缩。但 Hölder 不等式的左边是一个内积，我们要怎样把上面的求和式变成内积呢？事实上，我们可以将上面的式子拆分成 $|x_i|(|x_i|+y_i)^{p-1}$ 和 $|y_i|(|x_i|+|y_i|)^{p-1}$，即

$$\|\boldsymbol{x}+\boldsymbol{y}\|_p^p\leq\sum_{i=1}^{n}|x_i|(|x_i|+|y_i|)^{p-1}+\sum_{i=1}^{n}|y_i|(|x_i|+|y_i|)^{p-1}.$$

而这实际上就是内积项。如果我们令

$$\boldsymbol{z}=\begin{bmatrix}(|x_1|+|y_1|)^{p-1}\\(|x_2|+|y_2|)^{p-1}\\\vdots\\(|x_n|+|y_n|)^{p-1}\end{bmatrix},$$

则上面的式子可以写成

$$\|\boldsymbol{x}+\boldsymbol{y}\|_p^p\leq\boldsymbol{x}^\star\boldsymbol{z}+\boldsymbol{y}^\star\boldsymbol{z}.$$

现在，我们可以对 $\boldsymbol{x},\boldsymbol{z}$ 和 $\boldsymbol{y},\boldsymbol{z}$ 分别应用 Hölder 不等式，有

$$\begin{aligned}\|\boldsymbol{x}+\boldsymbol{y}\|_p^p&\leq\boldsymbol{x}^\star\boldsymbol{z}+\boldsymbol{y}^\star\boldsymbol{z}\\&\leq\|\boldsymbol{x}\|_p\|\boldsymbol{z}\|_q+\|\boldsymbol{y}\|_p\|\boldsymbol{z}\|_q\\&=(\|\boldsymbol{x}\|_p+\|\boldsymbol{y}\|_p)\|\boldsymbol{z}\|_q,\end{aligned}$$

现在，我们需要计算 $\|\boldsymbol{z}\|_q$，其中 q 满足 $\frac{1}{p}+\frac{1}{q}=1$.

$$\begin{aligned}\|\boldsymbol{z}\|_q&=\left(\sum_{i=1}^{n}(|x_i|+|y_i|)^{(p-1)q}\right)^{\frac{1}{q}}\\&=\left(\sum_{i=1}^{n}(|x_i|+|y_i|)^{p}\right)^{\frac{p-1}{p}}\\&=\|\boldsymbol{x}+\boldsymbol{y}\|_p^{p-1}.\end{aligned}$$

代入上面的不等式，有

$$\|\boldsymbol{x}+\boldsymbol{y}\|_p^p\leq(\|\boldsymbol{x}\|_p+\|\boldsymbol{y}\|_p)\|\boldsymbol{x}+\boldsymbol{y}\|_p^{p-1}.$$

两边除以 $\|\boldsymbol{x}+\boldsymbol{y}\|_p^{p-1}$ 即得所需结论。当且仅当 $\boldsymbol{x},\boldsymbol{y}$ 线性相关或有一方为零时等号成立。

于是我们证明了 l_p 范数在 $p\geq 1$ 时是一个范数。这个不等式也被称为 Minkowski 不等式。 □

上节课我们说范数是关于坐标 $\boldsymbol{x}$ 的连续函数。而对于 l_p 范数来说，它实际上是关于 $p, \boldsymbol{x}$ 的多元函数，我们自然会好奇它在 p 轴上的连续性。

定理 2.7. 对于给定的 $\boldsymbol{x} \in \mathbb{C}^n$，$l_p$ 范数 $\|\boldsymbol{x}\|_p$ 是 p 的连续函数。

证明.

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

$$\|\boldsymbol{x}\|_{p+\Delta p} \triangleq \left(\sum_{i=1}^n |x_i|^{p+\Delta p}\right)^{\frac{1}{p+\Delta p}} = \exp\left(\frac{1}{p+\Delta p}\ln\sum_{i=1}^n |x_i|^{p+\Delta p}\right)$$

由于指数函数是 $\mathbb{R}$ 上的连续函数，我们只需证明

$$\lim_{\Delta p\to 0}\frac{1}{p+\Delta p}\ln\sum_{i=1}^n |x_i|^{p+\Delta p} = \frac{1}{p}\ln\sum_{i=1}^n |x_i|^p.$$

由于 $\frac{1}{p+\Delta p}$ 在 $\Delta p \to 0$ 时连续，我们只需证明

$$\lim_{\Delta p\to 0}\ln\sum_{i=1}^n |x_i|^{p+\Delta p} = \ln\sum_{i=1}^n |x_i|^p.$$

由于对数函数在 $(0, +\infty)$ 上连续，我们只需证明

$$\lim_{\Delta p\to 0}\sum_{i=1}^n |x_i|^{p+\Delta p} = \sum_{i=1}^n |x_i|^p.$$

我们将求和与求极限交换次序，只需证明

$$\lim_{\Delta p\to 0}|x_i|^{p+\Delta p} = |x_i|^p, \quad i = 1, 2, \ldots, n.$$

根据指数函数的连续性，这显然是成立的。

最后，我们还需要证明

$$\lim_{p\to\infty}\|\boldsymbol{x}\|_p = \|\boldsymbol{x}\|_\infty.$$

注意到

$$\begin{aligned}
\|\boldsymbol{x}\|_p &= \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}} \\
&\leq \left(n\max_{1\leq i\leq n}|x_i|^p\right)^{\frac{1}{p}} \\
&= n^{\frac{1}{p}}\|\boldsymbol{x}\|_\infty \\
&\to \|\boldsymbol{x}\|_\infty, \quad p \to +\infty,
\end{aligned}$$

以及

$$\begin{aligned}\|\boldsymbol{x}\|_p &= \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}} \\ &\geq \left(\max_{1\leq i\leq n} |x_i|^p\right)^{\frac{1}{p}} \\ &= \|\boldsymbol{x}\|_\infty,\end{aligned}$$

由夹逼性，我们有

$$\lim_{p\to\infty} \|\boldsymbol{x}\|_p = \|\boldsymbol{x}\|_\infty.$$

这就证明了 l_p 范数在 $p > 0$ 时的连续性。 □

事实上，我们还可以证明 l_p 范数在 p 轴上是单调递减的。

定理 2.8. 对于给定的 $\boldsymbol{x} \in \mathbb{C}^n$，$l_p$ 范数 $\|\boldsymbol{x}\|_p$ 是 p 的单调递减函数。

证明. 设有 $1 \leq p < q \leq +\infty$，我们要证明 $\|\boldsymbol{x}\|_p \geq \|\boldsymbol{x}\|_q$.

当 $q = +\infty$ 时，我们已经在上面的连续性证明中给出了证明。下面我们假设 $q < +\infty$.

当 $\boldsymbol{x} = \boldsymbol{0}$ 时，显然成立。下面我们假设 $\boldsymbol{x} \neq \boldsymbol{0}$.

我们对 $\boldsymbol{x}$ 进行缩放变换，使得 $\|\boldsymbol{x}\|_q = 1$，则我们只需证明 $\|\boldsymbol{x}\|_p \geq 1$.

由于 $q > 1$，显然 $|x_i| \in [0, 1], \forall i \in \{1, 2, \ldots, n\}$. 又由于 $p < q$，因此对每个分量，都有 $|x_i|^p > |x_i|^q$. 两边求和，有

$$\sum_{i=1}^n |x_i|^p > \sum_{i=1}^n |x_i|^q = 1.$$

两边取 p 次方根，即得所需结论：

$$\|\boldsymbol{x}\|_p > 1 = \|\boldsymbol{x}\|_q.$$

如果我们对 $\|\boldsymbol{x}\|_q$ 进行缩放变换，结果仍然是成立的：设 $\|\boldsymbol{y}\|_q = 1$，我们刚才已经证明了 $\|\boldsymbol{y}\|_p \geq 1$，现在我们令 $\boldsymbol{x} = \|\boldsymbol{x}\|_q \boldsymbol{y}$，则有

$$\|\boldsymbol{x}\|_p = \|\boldsymbol{x}\|_q \|\boldsymbol{y}\|_p \geq \|\boldsymbol{x}\|_q.$$

因此对于给定的 $\boldsymbol{x} \in \mathbb{C}^n$，$l_p$ 范数 $\|\boldsymbol{x}\|_p$ 是 p 的单调递减函数。

这种缩放变换的方法实质上利用了范数的齐次性，在许多和范数相关的证明中都非常有用。 □

下图给出了 l_1, l_2, l_∞ 范数的单位球。可以发现，随着 p 的增大，单位球在向外扩张。这是因为 l_p 范数是 p 的单调递减函数，同一个向量在 l_1 范数意义下具有范数 1，但在其他 p 值下范数都小于 1，因而其他 p 值的单位球要比 l_1 范数的单位球来得大；同理，l_∞ 范数的单位球比 l_1, l_2 的都要大——并且 l_∞ 范数的单位球是所有范数中最大的。

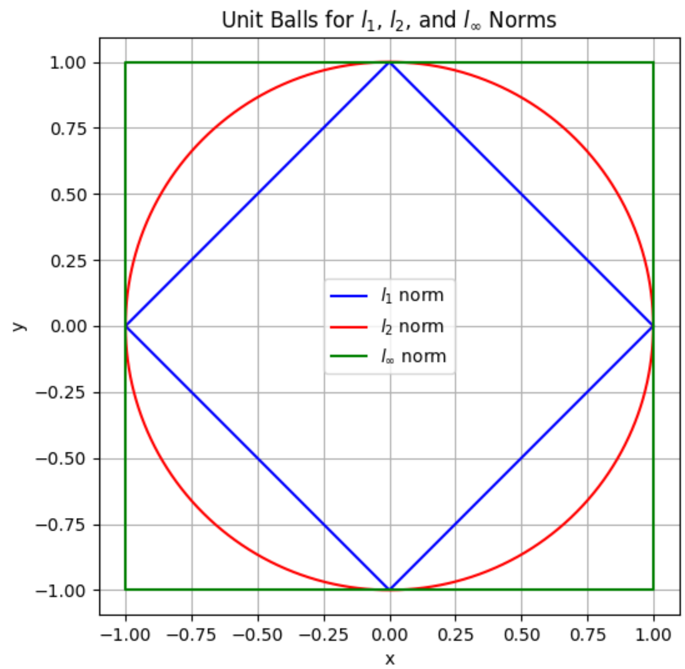

图 1: l_1, l_2, l_∞ 范数的单位球

3 矩阵范数的定义

对于向量，我们可以利用向量范数来衡量其大小。那么对于矩阵呢？我们自然也希望有一个类似的工具来衡量矩阵的大小。我们先来看几个例子，借以说明引入矩阵范数的必要性。

例 3.1. 若 $x \in \mathbb{C}$，且有 $|x| < 1$，我们知道下面这个幂级数收敛：

$$(1-x)^{-1} = 1 + x + x^2 + x^3 + \cdots = \sum_{i=0}^{\infty} x^i.$$

那么对于矩阵 $\boldsymbol{E} \in \mathbb{C}^{n\times n}$，我们猜测可以用

$$(\boldsymbol{I}-\boldsymbol{E})^{-1} = \boldsymbol{I} + \boldsymbol{E} + \boldsymbol{E}^2 + \boldsymbol{E}^3 + \cdots = \sum_{i=0}^{\infty} \boldsymbol{E}^i$$

来表示 $(\boldsymbol{I}-\boldsymbol{E})$ 的逆矩阵。但是，这个级数什么时候收敛呢？根据复数域上的幂级数收敛，我们猜测 $\boldsymbol{E}$ 的某种“大小”要小于 1 才行。这就需要我们有一个工具来衡量矩阵的大小。

同样地，对于矩阵函数

$$\mathrm{e}^{\boldsymbol{A}} = \boldsymbol{I} + \boldsymbol{A} + \frac{\boldsymbol{A}^2}{2!} + \frac{\boldsymbol{A}^3}{3!} + \cdots = \sum_{i=0}^{\infty} \frac{\boldsymbol{A}^i}{i!},$$

我们似乎同样需要一个工具来衡量矩阵 $\boldsymbol{A}$ 的大小，以保证上面的级数收敛。这就是矩阵范数的作用。

例 3.2. 在数值计算中，我们获得的输入往往有着一定的误差。例如，假如我们想要计算矩阵 $\boldsymbol{A}$ 的逆矩阵 $\boldsymbol{A}^{-1}$，但我们无法获得 $\boldsymbol{A}$ 的精确值——我们只有其近似值 $\tilde{\boldsymbol{A}} = \boldsymbol{A} + \boldsymbol{E}$，其中 $\boldsymbol{E}$ 是一个小扰动矩阵。于是，我们只能计算 $\tilde{\boldsymbol{A}}$ 的逆矩阵 $\tilde{\boldsymbol{A}}^{-1}$，而无法计算 $\boldsymbol{A}^{-1}$. 这时，我们希望 $\tilde{\boldsymbol{A}}^{-1}$ 能够尽可能地接近 $\boldsymbol{A}^{-1}$，即希望误差 $\tilde{\boldsymbol{A}}^{-1} - \boldsymbol{A}^{-1}$ 尽可能地小。但我们如何衡量误差的大小呢？这就需要我们有一个工具来衡量矩阵的大小。这也是矩阵范数的作用。

看来我们确乎需要一个量来衡量矩阵的大小。我们称其为矩阵范数。由于矩阵可以看成是一个向量空间，我们自然希望其范数继承向量范数的性质。

定义 3.3. （矩阵范数）对于 $\|\cdot\| : M_n \to \mathbb{R}$，如果它满足以下性质：

1. 非负性：$\forall \boldsymbol{A} \in M_n$，有 $\|\boldsymbol{A}\| \geq 0$，且当且仅当 $\boldsymbol{A} = \boldsymbol{0}$ 时等号成立；
2. 齐次性：$\forall \boldsymbol{A} \in M_n, \alpha \in \mathbb{C}$，有 $\|\alpha \boldsymbol{A}\| = |\alpha| \|\boldsymbol{A}\|$；
3. 三角不等式：$\forall \boldsymbol{A}, \boldsymbol{B} \in M_n$，有 $\|\boldsymbol{A} + \boldsymbol{B}\| \leq \|\boldsymbol{A}\| + \|\boldsymbol{B}\|$；

则称 $\|\cdot\|$ 是 M_n 上的一个矩阵范数。

其中一些矩阵范数还满足**相容性**：$\forall \boldsymbol{A}, \boldsymbol{B} \in M_n$，有 $\|\boldsymbol{A}\boldsymbol{B}\|_C \leq \|\boldsymbol{A}\|_A \|\boldsymbol{B}\|_B$；这里 $\|\cdot\|_A, \|\cdot\|_B, \|\cdot\|_C$ 分别作用于 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{A}\boldsymbol{B}$，它们可以是同一类范数，也可以是不同类范数。

可以看到，矩阵范数的前三个性质和向量范数是一样的，也是所有矩阵范数共有的，而相容性是一些矩阵范数所特有的。相容性本质上是“大小上的控制关系”，只要乘积的上界能用两个因子的范数乘积控制住即可，不要求范数类型相同。只要我们能控制住，我们就称涉及到的这些范数相容。那么我们为什么要引入相容性呢？理由有非常多，这里我们只给出其中的一条：如果我们把矩阵看作一个线性变换，那么 $\|\boldsymbol{A}\|$ 就是这个线性变换对空间的缩放因子。如果 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 分别对应两个线性变换，那么 $\boldsymbol{A}\boldsymbol{B}$ 就是先做 $\boldsymbol{B}$ 变换再做 $\boldsymbol{A}$ 变换的复合变换。从直观上来说，$\boldsymbol{A}\boldsymbol{B}$ 的缩放因子不可能大于 $\boldsymbol{A}, \boldsymbol{B}$ 的缩放因子的乘积，除非 $\boldsymbol{A}, \boldsymbol{B}$ 都是各向同性的缩放变换。而相容性正好保证了这一点。

具有相容性的矩阵范数有一些特殊的性质。

命题 3.4. 设 $\|\cdot\|$ 是一个相容的矩阵范数，则 $\|\boldsymbol{I}\| \geq 1$.

证明. 由于 $\boldsymbol{I} \cdot \boldsymbol{I} = \boldsymbol{I}$，由相容性，有

$$\|\boldsymbol{I}\| = \|\boldsymbol{I} \cdot \boldsymbol{I}\| \leq \|\boldsymbol{I}\| \|\boldsymbol{I}\|.$$

因为 $\|\boldsymbol{I}\| > 0$，所以两边除以 $\|\boldsymbol{I}\|$ 即得所需结论。 □

这条性质说明单位阵的范数不会小于 1. 事实上，我们可以对这条性质进行进一步推广。

命题 3.5. 设有 $\|\cdot\|$ 是一个相容的矩阵范数，则对任意的 $\boldsymbol{A}\in\mathbb{C}^{n\times n}$，

$$\|\boldsymbol{A}\| \geq \max_{1\leq i\leq n} \lambda_i.$$

其中 λ_i 是 $\boldsymbol{A}$ 的第 i 个特征值。

证明. 我们先给一个错误的证法。

由特征方程，$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{x}\lambda$，两边取范数，有

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

两边除以 $\|\boldsymbol{x}\|$ 即得 $\|\boldsymbol{A}\| \geq |\lambda|$.

这种证明方法的错误在于，我们在题中所说的“相容”是与自身相容，但上面的错误证法错误地认为矩阵范数与向量范数相容了，这个条件我们并没有给出。如果我们假设矩阵范数是 $\|\cdot\|_m$，向量范数是 $\|\cdot\|_v$，那么上面的式子可以重写为

$$\|\boldsymbol{A}\boldsymbol{x}\|_v = \|\boldsymbol{x}\lambda\|_v = |\lambda|\|\boldsymbol{x}\|_v \leq \|\boldsymbol{A}\|_m\|\boldsymbol{x}\|_v.$$

但是 $\|\cdot\|_m$ 和 $\|\cdot\|_v$ 并不一定相容，因此最后一个不等号是错误的。

下面我们给出正确的证明。上面的证明虽然错误，但给我们提供了一个思路。我们只需要在上面证法的基础上做一些修改，规避向量范数，即可得到正确的证明。问题的症结在于，代入特征方程后，必然会出现 $\|\boldsymbol{x}\|$ 这个向量范数。那么，我们能不能让 $\boldsymbol{x}$ 右乘上某个因子，让其变为 n 阶方阵呢？这样我们就可以规避向量范数了。事实上，我们可以让 $\boldsymbol{x}$ 右乘上 $\boldsymbol{x}^\star$，即

$$\begin{aligned}
\|\boldsymbol{x}\lambda\boldsymbol{x}^\star\| &= |\lambda|\|\boldsymbol{x}\boldsymbol{x}^\star\| \\
&= \|\boldsymbol{A}\boldsymbol{x}\boldsymbol{x}^\star\| \\
&= \|\boldsymbol{A}(\boldsymbol{x}\boldsymbol{x}^\star)\| \\
&\leq \|\boldsymbol{A}\|\|\boldsymbol{x}\boldsymbol{x}^\star\|.
\end{aligned}$$

由于 $\boldsymbol{x}\boldsymbol{x}^\star$ 是半正定矩阵，范数一定大于零。两边除以 $\|\boldsymbol{x}\boldsymbol{x}^\star\|$ 即得 $\|\boldsymbol{A}\| \geq |\lambda|$. 由于 λ_i 是 $\boldsymbol{A}$ 的任意一个特征值，因此 $\|\boldsymbol{A}\| \geq \max_{1\leq i\leq n}|\lambda_i|$.

我们定义谱半径 $\rho(\boldsymbol{A}) = \max\{|\lambda| : \det(\lambda\boldsymbol{I}-\boldsymbol{A})=0\}$，则上面的结论可以改写为

$$\|\boldsymbol{A}\| \geq \rho(\boldsymbol{A}).$$

之所以称之为谱半径，是因为如果我们把特征值看成复平面上的点，那么这些点一定落在以原点为圆心，$\rho(\boldsymbol{A})$ 为半径的圆盘内。 □

这条性质说明，任何矩阵的范数都不能过分地小——至少要大于谱半径。

下面我们给出几种常见的满足相容性的矩阵范数。

4 算子范数

上面我们提到矩阵可以看作一个线性变换，矩阵范数就是线性变换的缩放因子。下面这个定义正是基于这个思想。它被称为算子范数。

定义 4.1. （算子范数，*operator norm*）设 $\|\cdot\|_\alpha, \|\cdot\|_\beta$ 分别是 $\mathbb{C}^n$ 上的两个向量范数，则对于 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$，定义

$$\|\boldsymbol{A}\|_{\alpha,\beta} = \sup_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_\beta}{\|\boldsymbol{x}\|_\alpha} = \sup_{\|\boldsymbol{x}\|_\alpha=1} \|\boldsymbol{A}\boldsymbol{x}\|_\beta.$$

则称 $\|\cdot\|_{\alpha,\beta}$ 是 $\mathbb{C}^{m\times n}$ 上的一个算子范数。

其中 $\|\cdot\|_\alpha$ 作用于 $\boldsymbol{x}$，$\|\cdot\|_\beta$ 作用于 $\boldsymbol{A}\boldsymbol{x}$，即 α 衡量输入大小，β 衡量输出大小。

特别地，当 $\alpha = \beta$ 时，我们称这个算子范数为标准算子范数，记作 $\|\cdot\|_\alpha$.

实际上，算子范数还有一种等价形式。由齐次性，我们可以将 $\|\boldsymbol{x}\|_\alpha$ 塞入分子的范数中，得到一个 $\|\cdot\|_\beta$ 范数为 1 的新向量。因此，我们可以把算子范数写成

$$\|\boldsymbol{A}\|_{\alpha,\beta} = \sup_{\|\boldsymbol{y}\|_\alpha=1} \|\boldsymbol{A}\boldsymbol{y}\|_\beta,$$

根据这个等价形式，如果我们把 $\boldsymbol{A}$ 看成一个线性变换，那么算子范数就是这个线性变换最多能把 α 范数下的单位球映射到 β 范数下多远的位置。对于标准算子范数，就是把 α 范数下的单位球映射到同范数下多远的位置。

由算子范数的定义，我们很容易能够证明算子范数是一个矩阵范数，并且与向量范数相容，标准算子范数还与自身相容。

命题 4.2. 算子范数与向量范数相容。

证明. 由定义，有

$$\|\boldsymbol{A}\|_{\alpha,\beta} = \sup_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_\beta}{\|\boldsymbol{x}\|_\alpha} \geq \frac{\|\boldsymbol{A}\boldsymbol{x}\|_\beta}{\|\boldsymbol{x}\|_\alpha}, \quad \forall \boldsymbol{x} \neq \boldsymbol{0}.$$

因此，有

$$\|\boldsymbol{A}\boldsymbol{x}\|_\beta \leq \|\boldsymbol{A}\|_{\alpha,\beta}\|\boldsymbol{x}\|_\alpha, \quad \forall \boldsymbol{x} \neq \boldsymbol{0}.$$

因此算子范数与向量范数相容。 □

命题 4.3. 标准算子范数与自身相容。

证明. 由算子范数与向量范数的相容性，有

$$\|\boldsymbol{A}\boldsymbol{x}\| \leq \|\boldsymbol{A}\|\|\boldsymbol{x}\|, \quad \forall \boldsymbol{x} \neq \boldsymbol{0}.$$

令 $\boldsymbol{y}=\boldsymbol{B}\boldsymbol{x}$，有

$$\|\boldsymbol{A}\boldsymbol{B}\boldsymbol{y}\| \leq \|\boldsymbol{A}\|\|\boldsymbol{B}\boldsymbol{y}\| \leq \|\boldsymbol{A}\|\|\boldsymbol{B}\|\|\boldsymbol{y}\|, \quad \forall \boldsymbol{y} \neq \mathbf{0}.$$

取 $\|\boldsymbol{y}\|=1$，由 $\boldsymbol{y}$ 的任意性，左边可放大到上确界，因此有

$$\|\boldsymbol{A}\boldsymbol{B}\| \leq \|\boldsymbol{A}\|\|\boldsymbol{B}\|.$$

因此标准算子范数与自身相容。 □

请注意，非标准算子范数不与自身相容，但可以与不同类型的非标准算子范数相容，这要求范数类型满足某些特定条件。这种特殊条件指的是：$\|\cdot\|_{q,r}, \|\cdot\|_{p,q}, \|\cdot\|_{p,r}$ 分别作用于 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{A}\boldsymbol{B}$. 如果我们用线性变换的角度来看待矩阵，我们先作用 $\boldsymbol{B}$ 线性变换，范数从 p 变为 q；再作用 $\boldsymbol{A}$ 线性变换，范数从 q 变为 r；复合变换 $\boldsymbol{A}\boldsymbol{B}$ 的范数就是从 p 变为 r. 这样，我们就能够理解为什么三个范数一定要满足这样的系数关系才能相容了。可以用上面命题中的证明方法证明在这种情况下三个不同的非标准算子范数相容。

标准算子范数又称为 l_p 范数。

5 l_p 矩阵范数

定义 5.1. （l_p 矩阵范数）对于 $\boldsymbol{A} \in \mathbb{C}^{m\times n}, \boldsymbol{x} \in \mathbb{C}^n$，定义

$$\|\boldsymbol{A}\|_p = \sup_{\|\boldsymbol{x}\|_p=1} \|\boldsymbol{A}\boldsymbol{x}\|_p.$$

由于 l_p 矩阵范数是标准算子范数，因此它与自身相容，也与任何向量范数相容。

下面我们给出几种常见的 l_p 矩阵范数。

1. l_1 范数（列和范数）：$\|\boldsymbol{A}\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}|$.
2. l_2 范数（谱范数）：$\|\boldsymbol{A}\|_2 = \sqrt{\rho(\boldsymbol{A}^\star \boldsymbol{A})}$.
3. l_∞ 范数（行和范数）：$\|\boldsymbol{A}\|_\infty = \max_{1\leq i\leq m} \sum_{j=1}^{n} |a_{ij}|$.

我们将对上面的三种范数一一进行说明。

命题 5.2. （l_1 范数的形式）对矩阵 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$，其 l_1 范数可表示为

$$\|\boldsymbol{A}\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}|.$$

证明. 由定义

$$\|\boldsymbol{A}\|_1 = \sup_{\|\boldsymbol{x}\|_1=1} \|\boldsymbol{A}\boldsymbol{x}\|_1$$

如果我们将 $\boldsymbol{A}\boldsymbol{x}$ 看成是 $\boldsymbol{A}$ 的列向量的线性组合，则有

$$\|\boldsymbol{A}\|_1 = \max_{|x_1|+|x_2|+\cdots+|x_n|=1} \|a_1x_1 + a_2x_2 + \cdots + a_nx_n\|_1,$$

由向量范数的三角不等式，有

$$\begin{aligned}
\|\boldsymbol{A}\|_1 &\leq \max_{|x_1|+|x_2|+\cdots+|x_n|=1} (\|a_1\|_1|x_1| + \|a_2\|_1|x_2| + \cdots + \|a_n\|_1|x_n|) \\
&\leq \max\{\|a_1\|_1, \ldots, \|a_n\|_1\} \\
&= \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}| \\
&= \|\boldsymbol{a}_i\|_1.
\end{aligned}$$

然后取 $\boldsymbol{x} = \boldsymbol{e}_i$，则 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{a}_i$，则有

$$\|\boldsymbol{A}\|_1 = \max_{\|\boldsymbol{x}\|_1=1} \|\boldsymbol{A}\boldsymbol{x}\|_1 \geq \|\boldsymbol{A}\boldsymbol{e}_i\|_1 = \|\boldsymbol{a}_i\|_1.$$

于是

$$\|\boldsymbol{A}\|_1 = \|\boldsymbol{a}_i\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}|.$$

由于 l_1 范数是矩阵列向量的最大 l_1 范数，因此它也被称为列和范数。 □

命题 5.3. （l_∞ 范数的形式）对于 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$，其 l_∞ 范数可表示为

$$\|\boldsymbol{A}\|_\infty = \max_{1\leq i\leq m} \sum_{j=1}^{n} |a_{ij}|.$$

证明. 由定义

$$\|\boldsymbol{A}\|_\infty = \max_{\|\boldsymbol{x}\|_\infty=1} \|\boldsymbol{A}\boldsymbol{x}\|_\infty.$$

其中 $\boldsymbol{A}\boldsymbol{x}$ 的第 i 个分量是 $\boldsymbol{A}$ 的第 i 行 $A_i^\star$ 与 $\boldsymbol{x}$ 的内积。因此，上式可化简为

$$\begin{aligned}
\|\boldsymbol{A}\|_\infty &= \max_{1\leq i\leq m} |A_i^\star \boldsymbol{x}| \\
&\leq \max_{1\leq i\leq m} \sum_{j=1}^{n} |a_{ij}||x_j| \\
&\leq \max_{1\leq i\leq m} \sum_{j=1}^{n} |a_{ij}|.
\end{aligned}$$

第一个不等号利用了复数的三角不等式，第二个不等号利用了 $\|\boldsymbol{x}\|_\infty = 1$，因而 $|x_j| \le 1$，且等号均能取到。取 $\boldsymbol{x} = [1, 1, \cdots, 1]^{\mathrm{T}}$，则有

$$\|\boldsymbol{A}\|_\infty = \max_{\|\boldsymbol{x}\|_\infty = 1} \|\boldsymbol{A}\boldsymbol{x}\|_\infty \ge \|\boldsymbol{A}[1, 1, \cdots, 1]^{\mathrm{T}}\|_\infty = \max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}|.$$

由于 l_∞ 范数是矩阵行向量的最大 l_∞ 范数，因此它也被称为行和范数。 □

命题 5.4. （l_2 范数的形式）对于 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$，其 l_2 范数可表示为

$$\|\boldsymbol{A}\|_2 = \sqrt{\rho(\boldsymbol{A}^\star \boldsymbol{A})}.$$

证明. 由定义

$$\|\boldsymbol{A}\|_2^2 = \sup_{\|\boldsymbol{x}\|_2 = 1} \|\boldsymbol{A}\boldsymbol{x}\|_2^2 = \sup_{\|\boldsymbol{x}\|_2 = 1} \boldsymbol{x}^\star \boldsymbol{A}^\star \boldsymbol{A}\boldsymbol{x}.$$

由于 $\boldsymbol{A}^\star \boldsymbol{A}$ 是半正定矩阵，因此它有 n 个非负的特征值 $\lambda_1, \lambda_2, \ldots, \lambda_n$，且存在酉矩阵 $\boldsymbol{Q}$，使得

$$\boldsymbol{A}^\star \boldsymbol{A} = \boldsymbol{Q}\Lambda\boldsymbol{Q}^\star,$$

其中 $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$.

令 $\boldsymbol{y} = \boldsymbol{Q}^\star \boldsymbol{x}$，则 $\|\boldsymbol{y}\|_2 = \|\boldsymbol{x}\|_2 = 1$，因此

$$\|\boldsymbol{A}\|_2^2 = \sup_{\|\boldsymbol{y}\|_2 = 1} \boldsymbol{y}^\star \Lambda \boldsymbol{y} = \sup_{\|\boldsymbol{y}\|_2 = 1} \sum_{i=1}^{n} \lambda_i |y_i|^2.$$

由于 $\sum_{i=1}^{n} |y_i|^2 = 1$，有

$$\|\boldsymbol{A}\|_2^2 \le \rho(\boldsymbol{A}^\star \boldsymbol{A}) \sum_{i=1}^{n} |y_i|^2 = \rho(\boldsymbol{A}^\star \boldsymbol{A}).$$

取 $\boldsymbol{y}$ 为 Λ 的最大特征值对应的单位特征向量，则有

$$\|\boldsymbol{A}\|_2^2 = \sup_{\|\boldsymbol{y}\|_2 = 1} \|\boldsymbol{A}\boldsymbol{x}\|_2^2 \ge \|\rho(\boldsymbol{A}^\star \boldsymbol{A})\boldsymbol{x}\|_2^2 = \rho(\boldsymbol{A}^\star \boldsymbol{A}).$$

因此，$\|\boldsymbol{A}\|_2^2 = \rho(\boldsymbol{A}^\star \boldsymbol{A})$.

亦即

$$\|\boldsymbol{A}\|_2 = \sqrt{\rho(\boldsymbol{A}^\star \boldsymbol{A})}.$$

由于 l_2 范数带有谱半径，因此它也被称为谱范数。 □

l_2 范数虽然计算复杂度比较高，但它有着非常好的数学性质，因此在很多理论分析中都使用 l_2 范数。下面我们给出几个 l_2 范数的性质。

命题 5.5. （l_2 范数的酉不变性）对于任意的复方阵 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$，以及任意的酉矩阵 $\boldsymbol{Q}_1, \boldsymbol{Q}_2 \in \mathbb{C}^{n\times n}$，有

$$\|\boldsymbol{Q}_1\boldsymbol{A}\boldsymbol{Q}_2\|_2 = \|\boldsymbol{A}\|_2.$$

证明. 由 l_2 范数的公式，有

$$\|\boldsymbol{Q}_1\boldsymbol{A}\boldsymbol{Q}_2\|_2 = \sqrt{\rho((\boldsymbol{Q}_1\boldsymbol{A}\boldsymbol{Q}_2)^\star(\boldsymbol{Q}_1\boldsymbol{A}\boldsymbol{Q}_2))} = \sqrt{\rho(\boldsymbol{Q}_2^\star\boldsymbol{A}^\star\boldsymbol{A}\boldsymbol{Q}_2)}.$$

由于 $\boldsymbol{Q}_2$ 是酉矩阵，因此 $\boldsymbol{Q}_2^\star = \boldsymbol{Q}_2^{-1}$，所以 $\boldsymbol{Q}_2^\star\boldsymbol{A}^\star\boldsymbol{A}\boldsymbol{Q}_2$ 和 $\boldsymbol{A}^\star\boldsymbol{A}$ 相似，因此它们有相同的特征值，因而有相同的谱半径。于是，有

$$\|\boldsymbol{Q}_1\boldsymbol{A}\boldsymbol{Q}_2\|_2 = \sqrt{\rho(\boldsymbol{A}^\star\boldsymbol{A})} = \|\boldsymbol{A}\|_2.$$

□

命题 5.6. （l_2 范数与 l_1, l_∞ 范数的关系）对矩阵 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$，有

$$\|\boldsymbol{A}\|_2^2 \le \|\boldsymbol{A}\|_1\|\boldsymbol{A}\|_\infty.$$

证明. 由 l_2 范数的公式，有

$$\|\boldsymbol{A}\|_2^2 = \rho(\boldsymbol{A}^\star\boldsymbol{A}) \le \|\boldsymbol{A}^\star\boldsymbol{A}\|_1.$$

这里不等号利用了 $\rho(\boldsymbol{A}) \le \|\boldsymbol{A}\|$，其中 $\|\cdot\|$ 是任意满足相容性的矩阵范数。

由 l_1 范数的公式，有

$$\|\boldsymbol{A}^\star\boldsymbol{A}\|_1 = \max_{1\le j\le n}\sum_{i=1}^n |(\boldsymbol{A}^\star\boldsymbol{A})_{ij}| = \max_{1\le j\le n}\sum_{i=1}^n |\sum_{k=1}^m \bar{a}_{ki}a_{kj}|.$$

由复数的三角不等式，有

$$|\sum_{k=1}^m \bar{a}_{ki}a_{kj}| \le \sum_{k=1}^m |\bar{a}_{ki}||a_{kj}| = \sum_{k=1}^m |a_{ki}||a_{kj}|.$$

因此，有

$$\begin{aligned}
\|\boldsymbol{A}^\star\boldsymbol{A}\|_1 &\le \max_{1\le j\le n}\sum_{i=1}^n(\sum_{k=1}^m |a_{ki}||a_{kj}|) \\
&= \max_{1\le j\le n}(\sum_{k=1}^m |a_{kj}|(\sum_{i=1}^n |a_{ki}|)) \\
&\le (\max_{1\le j\le n}\sum_{k=1}^m |a_{kj}|)(\max_{1\le k\le m}\sum_{i=1}^n |a_{ki}|) \\
&= \|\boldsymbol{A}\|_1\|\boldsymbol{A}\|_\infty.
\end{aligned}$$

综上所述，有

$$\|\boldsymbol{A}\|_2^2 \le \|\boldsymbol{A}\|_1\|\boldsymbol{A}\|_\infty.$$

□

前面我们提到，有相容性的矩阵范数都有 $\|\boldsymbol{A}\| \geq \rho(\boldsymbol{A})$. 那么，对于一个给定的矩阵 $\boldsymbol{A}$，虽然它的范数一定大于等于谱半径，但它的范数能不能无限接近谱半径呢？答案是肯定的。下面这个定理就教我们如何为矩阵 $\boldsymbol{A}$ “定制”一个范数，使得这个范数能够无限接近 $\rho(\boldsymbol{A})$.

定理 5.7. *对于任意的 $\varepsilon > 0$，以及任意的 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$，存在一个相容的矩阵范数 $\|\cdot\|_\varepsilon$，使得*

$$\rho(\boldsymbol{A}) \leq \|\boldsymbol{A}\|_\varepsilon \leq \rho(\boldsymbol{A}) + \varepsilon.$$

证明. 第一个不等号我们在前面已经证明过了。

由于相似变换不改变矩阵的特征值，也就是不改变矩阵的谱半径，因此我们可以对 $\boldsymbol{A}$ 做相似变换。由 Schur 定理，存在酉矩阵 $\boldsymbol{Q}$，使得 $\boldsymbol{QAQ}^\star = \boldsymbol{T}$ 是上三角矩阵。如果我们变换后的矩阵是对角阵，我们对其取 l_∞ 范数，那么这个范数就应该等于其谱半径。但由于 $\boldsymbol{Q}^\star\boldsymbol{AQ}$ 只是上三角矩阵，上三角的因子对范数有着不小的影响。但我们可以通过进一步的相似变换，将上三角的非对角线元素缩小到任意小。方法如下：

设 $\boldsymbol{D} = \operatorname{diag}\{1, \delta, \delta^2, \ldots, \delta^{n-1}\}$，则有

$$\boldsymbol{D}^{-1}\boldsymbol{TD} = \begin{bmatrix} \lambda_1 & \delta t_{12} & \delta^2 t_{13} & \cdots & \delta^{n-1} t_{1n} \\ 0 & \lambda_2 & \delta t_{23} & \cdots & \delta^{n-2} t_{2n} \\ 0 & 0 & \lambda_3 & \cdots & \delta^{n-3} t_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

这里 λ_i 是 $\boldsymbol{A}$ 的特征值，t_{ij} 是 $\boldsymbol{T}$ 的非对角线元素。由于 $\boldsymbol{D}$ 是非奇异矩阵，因此 $\boldsymbol{D}^{-1}\boldsymbol{TD}$ 和 $\boldsymbol{T}$ 相似，因而和 $\boldsymbol{A}$ 相似，因此它们有相同的特征值，也就是相同的谱半径。由于 $\boldsymbol{D}^{-1}\boldsymbol{TD}$ 的非对角线元素都带有 δ 的幂次，因此我们可以通过让 δ 足够小，使得这些非对角线元素足够小。取 δ 使得 $\max_{i<j} |\delta^{j-i} t_{ij}| < \varepsilon/n$，则有

$$\|\boldsymbol{D}^{-1}\boldsymbol{TD}\|_\infty = \max_{1\leq i\leq n} (|\lambda_i| + \sum_{j=i+1}^{n} |\delta^{j-i} t_{ij}|) < \rho(\boldsymbol{A}) + (n-1)\frac{\varepsilon}{n} < \rho(\boldsymbol{A}) + \varepsilon.$$

这里不等号利用了 $\rho(\boldsymbol{A}) \geq \max_i |\lambda_i|$.

现在我们定义 $\|\boldsymbol{A}\|_\varepsilon = \|\boldsymbol{D}^{-1}\boldsymbol{QAQ}^\star\boldsymbol{D}\|_\infty$，则 $\|\cdot\|_\varepsilon$ 是一个相容的矩阵范数，且

$$\rho(\boldsymbol{A}) \leq \|\boldsymbol{A}\|_\varepsilon < \rho(\boldsymbol{A}) + \varepsilon.$$

□

6 Frobenius 范数

前面我们提到，矩阵可以看成是一个 $n \times n$ 维的向量。那么我们能否用向量范数直接定义矩阵范数呢？答案是肯定的，这就是 Frobenius 范数。

定义 6.1. （*Frobenius* 范数）对于任意的 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$，定义其 *Frobenius* 范数为

$$\|\boldsymbol{A}\|_{\mathrm{F}} = \|\mathrm{vec}(\boldsymbol{A})\|_2 = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2}.$$

如果诸位对 tr 运算符比较熟悉的话，也可以把 *Frobenius* 范数写成

$$\|\boldsymbol{A}\|_{\mathrm{F}} = \sqrt{\mathrm{tr}(\boldsymbol{A}^\star\boldsymbol{A})}.$$

Frobenius 范数是一个非常常用的矩阵范数，它非常符合我们对“长度”的直观认识，计算复杂度也比较低。它有着非常优良的数学性质。

Frobenius 范数还有一个非常重要的性质，就是它具有酉不变性。

引理 6.2. （*Frobenius* 范数的酉不变性）对于任意的 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$，以及任意的酉矩阵 $\boldsymbol{Q}_1 \in \mathbb{C}^{m\times m}, \boldsymbol{Q}_2 \in \mathbb{C}^{n\times n}$，有

$$\|\boldsymbol{Q}_1\boldsymbol{A}\boldsymbol{Q}_2\|_{\mathrm{F}} = \|\boldsymbol{A}\|_{\mathrm{F}}.$$

证明. 由 Frobenius 范数的定义，有

$$\begin{aligned}
\|\boldsymbol{Q}_1\boldsymbol{A}\boldsymbol{Q}_2\|_{\mathrm{F}} &= \mathrm{tr}((\boldsymbol{Q}_1\boldsymbol{A}\boldsymbol{Q}_2)^\star(\boldsymbol{Q}_1\boldsymbol{A}\boldsymbol{Q}_2)) \\
&= \mathrm{tr}(\boldsymbol{Q}_2^\star\boldsymbol{A}^\star\boldsymbol{A}\boldsymbol{Q}_2) \\
&= \mathrm{tr}(\boldsymbol{A}^\star\boldsymbol{A}\boldsymbol{Q}_2\boldsymbol{Q}_2^\star) \\
&= \mathrm{tr}(\boldsymbol{A}^\star\boldsymbol{A}) \\
&= \|\boldsymbol{A}\|_{\mathrm{F}}.
\end{aligned}$$

这里第二个等号利用了 $\boldsymbol{Q}_1$ 是酉矩阵，因此 $\boldsymbol{Q}_1^\star\boldsymbol{Q}_1 = \boldsymbol{I}$，第三个等号利用了 $\mathrm{tr}(\boldsymbol{A}\boldsymbol{B}) = \mathrm{tr}(\boldsymbol{B}\boldsymbol{A})$，第四个等号利用了 $\boldsymbol{Q}_2$ 是酉矩阵，因此 $\boldsymbol{Q}_2\boldsymbol{Q}_2^\star = \boldsymbol{I}$.

所以，Frobenius 范数具有酉不变性。 □

对于一个矩阵范数，我们最关心的是它有没有相容性。接下来我们就来研究 Frobenius 范数的相容性。

定理 6.3. *Frobenius* 范数是相容范数，即若 $\boldsymbol{A} \in \mathbb{C}^{m\times n}, \boldsymbol{B} \in \mathbb{C}^{n\times p}$，则有

$$\|\boldsymbol{A}\boldsymbol{B}\|_{\mathrm{F}} \le \|\boldsymbol{A}\|_{\mathrm{F}}\|\boldsymbol{B}\|_{\mathrm{F}}.$$

证明. 由 Frobenius 范数的定义，有

$$\|\boldsymbol{A}\boldsymbol{B}\|_{\mathrm{F}}^2 = \sum_{i,j}|\boldsymbol{A}\boldsymbol{B}_{i,j}|^2 = \sum_{i,j}|\boldsymbol{A}_i\boldsymbol{b}_j|^2,$$

其中 $\boldsymbol{A}_i$ 是 $\boldsymbol{A}$ 的第 i 行，$\boldsymbol{b}_j$ 是 $\boldsymbol{B}$ 的第 j 列。

由 Cauchy-Schwarz 不等式，有

$$|\boldsymbol{A}_i\boldsymbol{b}_j|^2 \le \|\boldsymbol{A}_i\|_2^2\|\boldsymbol{b}_j\|_2^2.$$

因此，有

$$\begin{aligned}\|\boldsymbol{AB}\|_{\mathrm{F}}^2 &\le \sum_{i,j}\|\boldsymbol{A}_i\|_2^2\|\boldsymbol{b}_j\|_2^2\\ &= (\sum_i\|\boldsymbol{A}_i\|_2^2)(\sum_j\|\boldsymbol{b}_j\|_2^2)\\ &= \|\boldsymbol{A}\|_{\mathrm{F}}^2\|\boldsymbol{B}\|_{\mathrm{F}}^2.\end{aligned}$$

□

Frobenius 范数还与 l_2 向量范数相容。

定理 6.4. *Frobenius 范数与 l_2 向量范数相容，即对于任意的 $\boldsymbol{A}\in\mathbb{C}^{m\times n},\boldsymbol{x}\in\mathbb{C}^n$，有*

$$\|\boldsymbol{Ax}\|_2 \le \|\boldsymbol{A}\|_{\mathrm{F}}\|\boldsymbol{x}\|_2.$$

证明. 我们将 $\boldsymbol{Ax}$ 视为 $\boldsymbol{A}$ 的列向量的线性组合，即

$$\boldsymbol{Ax} = \boldsymbol{a}_1x_1 + \boldsymbol{a}_2x_2 + \cdots + \boldsymbol{a}_nx_n,$$

其中 $\boldsymbol{a}_i$ 是 $\boldsymbol{A}$ 的第 i 列，x_i 是 $\boldsymbol{x}$ 的第 i 个分量。

由 l_2 向量范数的三角不等式，有

$$\|\boldsymbol{Ax}\|_2 = \|\boldsymbol{a}_1x_1 + \boldsymbol{a}_2x_2 + \cdots + \boldsymbol{a}_nx_n\|_2 \le \sum_{i=1}^n|x_i|\|\boldsymbol{a}_i\|_2.$$

这里出现了内积结构。由 Cauchy-Schwarz 不等式，有

$$(\sum_{i=1}^n|x_i|\|\boldsymbol{a}_i\|_2)^2 \le (\sum_{i=1}^n|x_i|^2)(\sum_{i=1}^n\|\boldsymbol{a}_i\|_2^2) = \|\boldsymbol{x}\|_2^2\|\boldsymbol{A}\|_{\mathrm{F}}^2.$$

因此，有

$$\|\boldsymbol{Ax}\|_2^2 \le \|\boldsymbol{x}\|_2^2\|\boldsymbol{A}\|_{\mathrm{F}}^2.$$

亦即

$$\|\boldsymbol{Ax}\|_2 \le \|\boldsymbol{A}\|_{\mathrm{F}}\|\boldsymbol{x}\|_2.$$

□

由上面这个引理我们可以得到一个推论：l_2 矩阵范数不大于 Frobenius 范数。

推论 6.5. 对于任意的 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$，有

$$\|\boldsymbol{A}\|_2 \le \|\boldsymbol{A}\|_{\mathrm{F}}.$$

证明. 由 l_2 矩阵范数的定义，有

$$\|\boldsymbol{A}\|_2 = \sup_{\|\boldsymbol{x}\|_2=1} \|\boldsymbol{A}\boldsymbol{x}\|_2.$$

由 Frobenius 范数与 l_2 向量范数的相容性，有

$$\|\boldsymbol{A}\boldsymbol{x}\|_2 \le \|\boldsymbol{A}\|_{\mathrm{F}}\|\boldsymbol{x}\|_2.$$

取 $\|\boldsymbol{x}\|_2 = 1$，则有

$$\|\boldsymbol{A}\boldsymbol{x}\|_2 \le \|\boldsymbol{A}\|_{\mathrm{F}}.$$

由于 $\boldsymbol{x}$ 的任意性，因此有

$$\|\boldsymbol{A}\|_2 = \sup_{\|\boldsymbol{x}\|_2=1} \|\boldsymbol{A}\boldsymbol{x}\|_2 \le \|\boldsymbol{A}\|_{\mathrm{F}}.$$

□

我们还可以发现，Frobenius 范数与 l_2 矩阵范数也是相容的。

定理 6.6. *Frobenius* 范数与 l_2 矩阵范数相容，即对于任意的 $\boldsymbol{A} \in \mathbb{C}^{m\times n}, \boldsymbol{B} \in \mathbb{C}^{n\times p}$，有

$$\|\boldsymbol{A}\boldsymbol{B}\|_{\mathrm{F}} \le \|\boldsymbol{A}\|_2\|\boldsymbol{B}\|_{\mathrm{F}}$$

以及

$$\|\boldsymbol{A}\boldsymbol{B}\|_{\mathrm{F}} \le \|\boldsymbol{A}\|_{\mathrm{F}}\|\boldsymbol{B}\|_2.$$

证明. 由于 Frobenius 范数带有根号，我们考虑其平方。由 Frobenius 范数的定义，有

$$\|\boldsymbol{A}\boldsymbol{B}\|_{\mathrm{F}}^2 = \sum_{i=1}^{p} \|\boldsymbol{A}\boldsymbol{b}_i\|_2^2,$$

上面的等式利用了矩阵乘法的列分块形式，即 $\boldsymbol{A}\boldsymbol{B}$ 的第 i 列是 $\boldsymbol{A}$ 乘上 $\boldsymbol{B}$ 的第 i 列。

由 Frobenius 范数与 l_2 向量范数的相容性，有

$$\|\boldsymbol{A}\boldsymbol{b}_i\|_2 \le \|\boldsymbol{A}\|_2\|\boldsymbol{b}_i\|_2.$$

因此，有

$$\|\boldsymbol{A}\boldsymbol{B}\|_{\mathrm{F}}^2 \le \sum_{i=1}^{p} \|\boldsymbol{A}\|_2^2\|\boldsymbol{b}_i\|_2^2 = \|\boldsymbol{A}\|_2^2\|\boldsymbol{B}\|_{\mathrm{F}}^2.$$

亦即

$$\|\boldsymbol{A}\boldsymbol{B}\|_{\mathrm{F}} \le \|\boldsymbol{A}\|_2\|\boldsymbol{B}\|_{\mathrm{F}}.$$

同理，将 $\boldsymbol{A}$ 进行行分块，也可以得到

$$\|\boldsymbol{AB}\|_{\mathrm{F}} \leq \|\boldsymbol{A}\|_{\mathrm{F}}\|\boldsymbol{B}\|_2.$$

但更简洁的做法是利用矩阵转置的性质：

$$\begin{aligned}\|\boldsymbol{AB}\|_{\mathrm{F}} &= \|\boldsymbol{B}^{\boldsymbol{H}}\boldsymbol{A}^{\boldsymbol{H}}\|_{\mathrm{F}} \\ &\leq \|\boldsymbol{B}^{\star}\|_{\mathrm{F}}\|\boldsymbol{A}^{\star}\|_2 \\ &= \|\boldsymbol{B}\|_{\mathrm{F}}\|\boldsymbol{A}\|_2.\end{aligned}$$

综上所述，我们得到了 Frobenius 范数与 l_2 矩阵范数之间的相容性关系：

$$\|\boldsymbol{AB}\|_{\mathrm{F}} \leq \|\boldsymbol{A}\|_2\|\boldsymbol{B}\|_{\mathrm{F}}$$

以及

$$\|\boldsymbol{AB}\|_{\mathrm{F}} \leq \|\boldsymbol{A}\|_{\mathrm{F}}\|\boldsymbol{B}\|_2.$$

□

综上，Frobenius 范数是一个非常好的矩阵范数，它具有酉不变性，又与 l_2 向量范数和 l_2 矩阵范数相容，还与其自身相容，因此在很多理论分析中都非常有用。

7 赋范线性空间

讲了这么多矩阵范数，现在让我们回到最开始的问题：为什么要研究矩阵范数？引入矩阵范数，我们可以建立起赋范线性空间的理论，在赋范线性空间中我们可以引入微积分这一强大的工具，从而可以对矩阵进行更深入的分析。

定义 7.1. （赋范线性空间）设 $\mathcal{V}$ 是一个线性空间，$\|\cdot\|$ 是 $\mathcal{V}$ 上的一个范数，则称 $(\mathcal{V}, \|\cdot\|)$ 是一个赋范线性空间。

赋范线性空间是一个非常重要的数学概念，它将线性空间与范数结合起来，从而使得我们可以在这个空间中进行距离、收敛等分析。

例 7.2. （*Neumann* 级数）已知 $\det(\boldsymbol{A}) \neq 0$，请问最多加多大的扰动，可以使其逆矩阵仍然存在？

解. 记 $\hat{\boldsymbol{A}} = \boldsymbol{A} + \Delta\boldsymbol{A} = \boldsymbol{A}(\boldsymbol{I} + \boldsymbol{A}^{-1}\Delta\boldsymbol{A}) = \boldsymbol{A}(\boldsymbol{I} - \boldsymbol{E})$，其中 $\boldsymbol{E} = -\boldsymbol{A}^{-1}\Delta\boldsymbol{A}$. 由于 $\det(\boldsymbol{A}) \neq 0$，因此 $\boldsymbol{A}$ 可逆，因此 $\hat{\boldsymbol{A}}$ 可逆的充分必要条件是 $\boldsymbol{I} - \boldsymbol{E}$ 可逆。现在我们就来研究 $\boldsymbol{I} - \boldsymbol{E}$ 的可逆性。

注意到有如下恒等式：

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

我们将 $\boldsymbol{I}$ 移到左边，并对两边同时取范数，有

$$\|(\boldsymbol{I}-\boldsymbol{E})(\boldsymbol{I}+\boldsymbol{E}+\boldsymbol{E}^2+\cdots+\boldsymbol{E}^{k-1})-\boldsymbol{I}\|=\|\boldsymbol{E}^k\|\leq\|\boldsymbol{E}\|^k.$$

因此，下面这个级数（称为 *Neumann* 级数）收敛的充要条件是 $\|\boldsymbol{E}^k\|\to 0$：

$$(\boldsymbol{I}-\boldsymbol{E})^{-1}=\boldsymbol{I}+\boldsymbol{E}+\boldsymbol{E}^2+\cdots$$

而 $\|\boldsymbol{E}^k\|\to 0$ 的充要条件是 $\rho(\boldsymbol{E})<1$，因此只需 $\rho(\boldsymbol{A}^{-1}\Delta\boldsymbol{A})<1$ 即可让上面这个 *Neumann* 级数收敛，从而 $\hat{\boldsymbol{A}}$ 可逆。

如果我们引入矩阵范数的相容性，则条件可进一步简化为

$$\|\boldsymbol{E}\|<1.$$

下面这个结论是 Bolzano-Weierstrass 定理在赋范线性空间中的推论。

推论 7.3. 酉矩阵序列 $\{\boldsymbol{Q}_n\}$ 必有收敛子列。

证明. 酉矩阵 $\boldsymbol{Q}$ 满足

$$\boldsymbol{Q}^\star\boldsymbol{Q}=\boldsymbol{I},$$

因此任何列向量的 2-范数均为 1，则有

$$\|\boldsymbol{Q}\|_{\mathrm{F}}=\sqrt{n}.$$

由范数的等价性，酉矩阵序列 $\{\boldsymbol{Q}_n\}$ 是有界的，因此必有收敛子列。 □

下面我们给出一个著名的谱半径公式，并给出大致的证明思路，以结束本讲的内容。

定理 7.4. （*Gelfand* 公式）设 $\boldsymbol{A}\in\mathbb{C}^{n\times n}$，则有

$$\rho(\boldsymbol{A})=\lim_{k\to\infty}\|\boldsymbol{A}^k\|^{1/k}.$$

证明. 要证明这个定理，我们只需证明两个不等式：

$$\rho(\boldsymbol{A})\leq\varliminf_{k\to\infty}\|\boldsymbol{A}^k\|^{1/k},$$
$$\rho(\boldsymbol{A})\geq\varlimsup_{k\to\infty}\|\boldsymbol{A}^k\|^{1/k}.$$

如果这两个不等式成立，根据上极限大于等于下极限，我们即可证明

$$\rho(\boldsymbol{A})=\varliminf_{k\to\infty}\|\boldsymbol{A}^k\|^{1/k}=\varlimsup_{k\to\infty}\|\boldsymbol{A}^k\|^{1/k}=\lim_{k\to\infty}\|\boldsymbol{A}^k\|^{1/k}.$$

现在，我们就来证明这两个不等式。

首先，我们需要说明 $\|\boldsymbol{A}^k\|^{1/k}$ 的有界性，这样它才会有上下极限。它有一个下界为 0，因此我们只需说明它有上界。由矩阵范数的等价性，我们只需说明在相容范数情况下 $\|\boldsymbol{A}^k\|^{1/k}$ 有上界即可，非相容范数的情况可以通过范数的等价性转化为相容范数的情况。由矩阵范数的相容性，有

$$\|\boldsymbol{A}^k\| \leq \|\boldsymbol{A}\|^k,$$

两边开 k 次根号，则有

$$\|\boldsymbol{A}^k\|^{1/k} \leq \|\boldsymbol{A}\|.$$

这里 $\|\boldsymbol{A}\|$ 是一个常数，因此 $\|\boldsymbol{A}^k\|^{1/k}$ 有上界。因此 $\|\boldsymbol{A}^k\|^{1/k}$ 有上下极限。

下面我们来证明第一个不等式。

这里我们同样可以只考虑相容范数的情况。因为非相容范数和相容范数是等价的，常数开 k 次根号后会趋于 1，不会改变极限。由矩阵范数的相容性，有

$$\|\boldsymbol{A}^k\| \geq \rho(\boldsymbol{A}^k) = \rho(\boldsymbol{A})^k.$$

这里的等号利用了一个基本事实：若 $\boldsymbol{A}$ 有特征值 λ，则 $\boldsymbol{A}^k$ 有特征值 λ^k.

两边开 k 次根号，则有

$$\|\boldsymbol{A}^k\|^{1/k} \geq \rho(\boldsymbol{A}).$$

由于 $\|\boldsymbol{A}^k\|^{1/k}$ 的任意子列都满足上面的不等式，因此它的下极限也满足上面的不等式，即

$$\varliminf_{k\to\infty} \|\boldsymbol{A}^k\|^{1/k} \geq \rho(\boldsymbol{A}).$$

下面我们来证明第二个不等式。

我们只需证明 $\forall \varepsilon > 0$，有

$$\rho(\boldsymbol{A}) + \varepsilon \geq \varlimsup_{k\to\infty} \|\boldsymbol{A}^k\|^{1/k}.$$

由上面的定理，我们可以为 $\boldsymbol{A}$ 定义一个相容范数 $\|\cdot\|_\varepsilon$，使得

$$\rho(\boldsymbol{A}) \leq \|\boldsymbol{A}\|_\varepsilon \leq \rho(\boldsymbol{A}) + \varepsilon.$$

由矩阵范数的相容性，有

$$\|\boldsymbol{A}^k\|_\varepsilon \leq \|\boldsymbol{A}\|_\varepsilon^k \leq (\rho(\boldsymbol{A}) + \varepsilon)^k.$$

两边开 k 次根号，则有

$$\|\boldsymbol{A}^k\|_\varepsilon^{1/k} \leq \rho(\boldsymbol{A}) + \varepsilon.$$

由于 $\|\boldsymbol{A}^k\|_\varepsilon^{1/k}$ 的任意子列都满足上面的不等式，因此它的上极限也满足上面的不等式，即

$$\varlimsup_{k\to\infty} \|\boldsymbol{A}^k\|_\varepsilon^{1/k} \leq \rho(\boldsymbol{A}) + \varepsilon.$$

由于非相容范数和相容范数是等价的，常数开 k 次根号后会趋于 1，因此

$$\overline{\lim_{k\to\infty}}\|\boldsymbol{A}^k\|^{1/k}\le\rho(\boldsymbol{A})+\varepsilon.$$

ε 具有任意性，因此有

$$\overline{\lim_{k\to\infty}}\|\boldsymbol{A}^k\|^{1/k}\le\rho(\boldsymbol{A}).$$

综上，结合两个不等式，我们得到了 Gelfand 公式：

$$\rho(\boldsymbol{A})=\lim_{k\to\infty}\|\boldsymbol{A}^k\|^{1/k}.$$

$\square$