矩阵论

第六章　范数与极限

本章概要

- 矩阵的代数运算 Vs.矩阵的分析运算
- 范数理论：矩阵分析、数值分析、数据处理
- 本章内容：
  - 向量范数：在线性空间中定义向量的范数，介绍赋范线性空间的概念；
  - 矩阵范数：讨论矩阵的范数及其性质；
  - 矩阵序列与矩阵级数：介绍矩阵序列及矩阵幂级数的收敛性概念

§6.1 向量范数

- 对于同一线性空间中的向量“大小”的度量、两个向量之间的“接近”程度的比较——引入非负实数
- 对于线性空间 V ，在 V 上定义内积 (a, b) ，由内积可导出 $a \in V$ 的“长度”或范数 $\|a\| = \sqrt{(a, a)}$ 来进行度量
- 实际应用中，仅使用由内积导出的范数往往会带来不便，而且，范数不一定从内积导出，它只要满足一些基本的公理条件

基本概念

定义 6.1.1 设 V 是数域 P 上的线性空间，$\|a\|$是以 V 中的向量a 为自变量的非负实值函数，如果它满足以下三个条件：

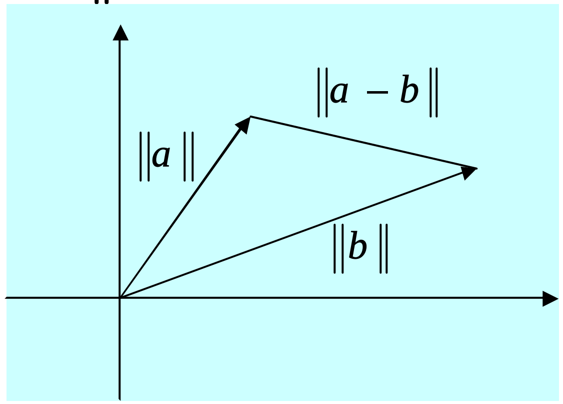
（1） **非负性**：当$a \neq 0$时，$\|a\|>0$；当$a=0$时，$\|a\|=0;$

（2） **齐次性**：对任意$k \in \mathrm{P}$，$a \in V$，有$\|ka\|=|k|\|a\|$；

（3） **三角不等式**：对任意$a, b \in V$，有$\|a+b\| \leq \|a\|+\|b\|$，

则称$\|a\|$为向量a 的**范数**，并称定义了范数的线性空间为**赋范线性空间**。

- 对任意的$a,b\in V$，$\|a\|=\|-a\|$，$\big|\|a\|-\|b\|\big|\le\|a-b\|$



证明： $\|a\|=\|a-b+b\|\le\|a-b\|+\|b\|$，

$\|b\|=\|b-a+a\|\le\|b-a\|+\|a\|$，

可得：

$\|a\|-\|b\|\le\|a-b\|$，$\|b\|-\|a\|\le\|b-a\|=\|a-b\|$

所以：$\big|\|a\|-\|b\|\big|\le\|a-b\|$

- 在赋范线性空间 V 中，可以由范数定义两点间的距离。对$a,b\in V$，定义a与b之间的距离为$d(a,b)=\|a-b\|$。

- 由上式规定的距离$d(a,b)$称为由范数$\|\cdot\|$决定的距离。

例 6.1.1 在 n 维向量空间C^n中，对任意向量$x=\left(x_1,\cdots,x_n\right)^{\mathrm{T}}\in C^n$，定义

$$\|x\|_1=\sum_{i=1}^{n}|x_i| \text{，} \quad \|x\|_2=\left(\sum_{i=1}^{n}|x_i|^2\right)^{\frac{1}{2}}\text{，} \quad \|x\|_\infty=\max_{1\le i\le n}|x_i|$$

它们都满足定义 6.1.1 中的三个条件，因此都是C^n上的范数，分别称为 1 范数，2 范数（或 Euclid 范数）和∞范数。

证明：仅给出∞范数满足定义 6.1.1 中条件的证明

(1)当$x\neq 0$时，$\|x\|_\infty=\max\limits_{1\le i\le n}|x_i|\neq 0$；当$x=0$时，$\|x\|_\infty=\max\limits_{1\le i\le n}|x_i|=0$;

(2)$\|kx\|_\infty=\max\limits_{1\le i\le n}|kx_i|=|k|\max\limits_{1\le i\le n}|x_i|=|k|\|x\|_\infty$；

(3)$\|x+y\|_\infty=\max\limits_{1\le i\le n}|x_i+y_i|\le\max\limits_{1\le i\le n}|x_i|+\max\limits_{1\le i\le n}|y_i|=\|x\|_\infty+\|y\|_\infty$

更一般的，对$1\le p<+\infty$，在C^n上定义$\|x\|_p=\left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}}$， $1\le p<+\infty$

则当$p=1$时，$\|x\|_p=\sum_{i=1}^{n}|x_i|=\|x\|_1$；当$p=2$时，$\|x\|_p=\left(\sum_{i=1}^{n}|x_i|^2\right)^{\frac{1}{2}}=\|x\|_2$。

可以证明以上定义的$\|x\|_p$是C^n上的一种向量范数，称为 p 范数。在给出该结论（定理 6.1.3）前，首先看以下的引理和定理

- ■引理 6.1.1
- ■定理 6.1.1（Hölder 不等式）
- ■定理 6.1.2（Minkowski 不等式）

引理 6.1.1 如果实数$p>1$，$q>1$且$\frac{1}{p}+\frac{1}{q}=1$，则对任意非负实数$a$，$b$有$ab\le\frac{a^p}{p}+\frac{b^q}{q}$

证明：当$ab=0$时，显然成立。当a,b均为正数时，定义$\varphi(t)=\frac{t^p}{p}+\frac{t^{-q}}{q}$，$0<t<+\infty$。因$\varphi'(t)=t^{p-1}+t^{-(q+1)}=\frac{t^{p+q}-1}{t^{q+1}}$，可见：当$0<t<1$时，$\varphi'(t)<0$；当$1\le t<+\infty$，$\varphi'(t)\ge 0$，故有$\varphi(t)\ge\varphi(1)=1$，$0<t<+\infty$。

令$t=a^{\frac{1}{q}}b^{-\frac{1}{p}}$，则有$1\le\frac{1}{p}\left(a^{\frac{1}{q}}b^{-\frac{1}{p}}\right)^p+\frac{1}{q}\left(a^{\frac{1}{q}}b^{-\frac{1}{p}}\right)^{-q}=\frac{1}{ab}\left(\frac{a^p}{p}+\frac{b^q}{q}\right)$

定理 6.1.1（Hölder 不等式） 设 $x=(x_1,\cdots,x_n)^{\mathrm{T}}$，$y=(y_1,\cdots,y_n)^{\mathrm{T}}\in C^n$，则

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

其中实数 $p>1$，$q>1$ 且 $\frac{1}{p}+\frac{1}{q}=1$

定理 6.1.2（Minkowski 不等式） 设 $x=(x_1,\cdots,x_n)^{\mathrm{T}}$，$y=(y_1,\cdots,y_n)^{\mathrm{T}}\in C^n$，则

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

其中实数 $p>1$.

定理 6.1.3 对任意向量$x=(x_1,\cdots,x_n)^{\mathrm{T}}\in C^n$，$1\le p<\infty$，$\|x\|_p=\left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}}$

是C^n上的向量范数，并且$\lim\limits_{p\to\infty}\|x\|_p=\|x\|_\infty$。

证明：（只要证明它满足定义 6.1.1 的三个条件）

$\|x\|_p$显然满足定义 6.1.1 中的条件（1）非负性、（2）齐次性，由定理

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

可知$\|x\|_p$满足定义 6.1.1 中的条件（3）三角不等式，

因此$\|x\|_p$是C^n上的向量范数。

$$\left(\|x\|_\infty\right)^p = \left(\max_i |x_i|\right)^p \le \sum_{i=1}^{n} |x_i|^p = \left(\|x\|_p\right)^p$$

$$\le n \max_i |x_i|^p = n\left(\|x\|_\infty\right)^p$$

$$\|x\|_\infty \le \|x\|_p \le n^{\frac{1}{p}} \|x\|_\infty,\ p \ge 1$$

$$\lim_{p\to\infty} \|x\|_p \le \|x\|_\infty$$

由已知范数构造新的向量范数

定理 6.1.4 设$\|\cdot\|_b$是C^m上的向量范数，$A\in C^{m\times n}$且$rank(A)=n$，则由

$$\|x\|_a=\|Ax\|_b,x\in C^n$$

所定义的$\|\cdot\|_a$是C^n上的向量范数。

定理 6.1.5设 V 是数域 P 上的 n 维线性空间，$e_1,\cdots e_n$为 V 的一组基，则 V 中的任一向量a 可以唯一地表示为$a=\sum_{i=1}^{n}x_ie_i$，$x=(x_1,\cdots x_n)^{\mathrm{T}}\in \mathrm{P}^n$，又设$\|\cdot\|$是$\mathrm{P}^n$上的向量范数，令$\|a\|_v=\|x\|$，则$\|a\|_v$是 V 上的向量范数。

证明：只需验证$\|\cdot\|_a$或$\|a\|_v$满足定义 6.1.1 中的三个条件。

下面只证明定理 6.1.5。

（1）对 V 中的任一向量a，如果$a \neq 0$，则其坐标向量$x \neq 0$，从而$\|a\|_v = \|x\| > 0$；如果$a = 0$，则其坐标向量$x = 0$，从而$\|a\|_v = \|x\| = 0$.

（2）对$k \in \mathrm{P}$，$ka = \sum_{i=1}^{n} kx_i e_i$，即$ka$的坐标向量为$kx = (kx_1, \cdots kx_n)^{\mathrm{T}}$，故

$$\|ka\|_v = \|kx\| = |k|\|x\| = |k|\|a\|_v$$

（3）对$b \in V$，则$b = \sum_{i=1}^{n} y_i e_i$，$y = (y_1, \cdots y_n)^{\mathrm{T}}$，$a + b$的坐标向量为$x + y$.

于是

$$\|a + b\|_v = \|x + y\| \leq \|x\| + \|y\| = \|b\|_v + \|b\|_v$$

因此$\|a\|_v$是 V 上的向量范数.

范数的等价性

定义 6.1.2 设$\|a\|_a$，$\|a\|_b$是 n 维线性空间 V 上定义的两种向量范数，如果存在两个与a 无关的正常数d_1，d_2使得

$$d_1\|\boldsymbol{a}\|_b \le \|\boldsymbol{a}\|_a \le d_2\|\boldsymbol{a}\|_b, \forall \boldsymbol{a} \in V$$

则称范数$\|a\|_a$与$\|a\|_b$是等价的。

容易证明：向量范数的等价具有**自反性**（$\|a\|_a$与$\|a\|_a$等价），**对称性**（若$\|a\|_a$与$\|a\|_b$等价，则$\|a\|_b$与$\|a\|_a$等价），**传递性**（若$\|a\|_a$与$\|a\|_b$等价，$\|a\|_b$与$\|a\|_c$等价，则$\|a\|_a$与$\|a\|_c$等价）

定理 6.1.6 设$\|\cdot\|$是数域 P 上 n 维线性空间 V 上的任一向量范数，$e_1,\cdots e_n$为 V 的一组基，V 中的任一向量a 可以唯一地表示为 $a=\sum_{i=1}^{n}x_i e_i$，$x=(x_1,\cdots x_n)^{\mathrm{T}}\in \mathrm{P}^n$，则$\|a\|$是$x_1,\cdots,x_n$的连续函数。

证明：对任意$e>0$及任意向量$a=\sum_{i=1}^{n}x_i e_i$，$b=\sum_{i=1}^{n}y_i e_i\in V$，令

$$d=e\Big/\left(\sum_{i=1}^{n}\|e_i\|^2\right)^{\frac{1}{2}}$$

，则当$\|x-y\|_2<d$时，有

$$\big|\|a\|-\|b\|\big|\le\|a-b\|=\left\|\sum_{i=1}^{n}(x_i-y_i)e_i\right\|\le\sum_{i=1}^{n}|x_i-y_i|\|e_i\|\le\sqrt{\sum_{i=1}^{n}|x_i-y_i|^2}\sqrt{\sum_{i=1}^{n}\|e_i\|^2}<d$$

于是$\|a\|$是$x_1,\cdots,x_n$的连续函数。

定理 6.1.7 有限维线性空间 V 上的任意两个向量范数都是等价的。

证明：设 V 是数域 P 上的 n 维线性空间，$e_1, \cdots e_n$ 为 V 的一组基，则 V 中的任一向量 a 可以唯一地表示为 $a = \sum_{i=1}^{n} x_i e_i$，$x = (x_1, \cdots x_n)^{\mathrm{T}} \in \mathrm{P}^n$. 令

$$\|a\|_2 = \|x\|_2 = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{\frac{1}{2}}$$

由定理 6.1.5 知 $\|a\|_2$ 是 V 上的一个向量范数。

设 $\|a\|_a$ 是 V 上的任一向量范数。因为向量范数的等价具有对称性和传递性，所以我们只需证明 $\|a\|_a$ 与 $\|a\|_2$ 等价即可。

如果 $a = 0$，则 $\|a\|_a$ 与 $\|a\|_2$ 显然等价。因此我们不妨假定 $a \neq 0$。

由定理 6.1.6 **知**$\|a\|_a$**是**$x_1,\cdots,x_n$**的连续函数，所以**$\|a\|_a$**在有界闭集**

$$j=\left\{x=\left(x_1,\cdots,x_n\right)^{\mathrm{T}}\in \mathrm{P}^n \left| \left|x_1\right|^2+\left|x_2\right|^2+\cdots\left|x_n\right|^2=1\right.\right\}$$

上可取得最大值 M **和最小值** m**。因为如果**$x\in j$ **，则**$x\neq 0$**，所以**$m>0$**。**

记
$$b=\sum_{i=1}^{n}\frac{x_i}{\|x\|_2}e_i \text{ ，}$$

则b **在基**$e_1,\cdots e_n$**下的坐标向量为**$y=\left(\frac{x_1}{\|x\|_2},\frac{x_2}{\|x\|_2},\cdots,\frac{x_n}{\|x\|_2}\right)^{\mathrm{T}}\in j$ **,从而有**

$$0<m\leq\|b\|_a\leq M$$

但$b=\dfrac{a}{\|x\|_2}$**，故**$m\|x\|_2\leq\|a\|_a\leq M\|x\|_2$**，即**$m\|a\|_2\leq\|a\|_a\leq M\|a\|_2$**。得证。**

向量序列的收敛性

定义 6.1.3 设$\left\{x^{(k)}\right\}$是C^n中的向量序列，其中$x^{(k)}=\left(x_1^{(k)},x_2^k,\cdots,x_n^k\right)^{\mathrm{T}}$。如果当$k\to\infty$时，$x^{(k)}$的每一个分量$x_i^{(k)}$都有极限$x_i\left(i=1,2,\cdots n\right)$，则称向量序列$\left\{x^{(k)}\right\}$是收敛的，并且向量$x=\left(x_1,x_2,\cdots x_n\right)^{\mathrm{T}}$称为$\left\{x^{(k)}\right\}$的极限，记为

$$\lim_{k\to\infty}x^{(k)}=x$$

不收敛的向量序列称为发散的。

向量序列的收敛性具有如下性质：

设$\{x^{(k)}\}$，$\{y^{(k)}\}$**是**C^n**中两个向量序列，**a,b **是复常数，**$A\in C^{m\times n}$。

如果$\lim\limits_{k\to\infty} x^{(k)}=x$，$\lim\limits_{k\to\infty} y^{(k)}=y$，**则**

（1）$\lim\limits_{k\to\infty}\left(ax^{(k)}+by^{(k)}\right)=ax+by$

（2）$\lim\limits_{k\to\infty} Ax^{(k)}=Ax$

定理 6.1.8 C^n**中向量序列**$\{x^{(k)}\}$**收敛于向量** x **的充分必要条件是对任一向量范数**$\|\cdot\|$，**数列**$\{\|x^{(k)}-x\|\}$**收敛于** 0。

证明：由范数的等价性，只要对某一种范数证明即可。为此取$\|\cdot\|=\|\cdot\|_\infty$。

必要性：

如果$\lim_{k\to\infty}x^{(k)}=x$，则对$i=1,2,\cdots,n$，有$x_i^{(k)}-x_i\to 0\left(k\to 0\right)$，

故 $\max_i\left|x_i^{(k)}-x_i\right|\to 0\left(k\to\infty\right)$，即$\left\{\left\|x^{(k)}-x\right\|_\infty\right\}$收敛于 0.

充分性：如果$\lim_{k\to\infty}\left\|x^{(k)}-x\right\|=0$，即$\max_i\left|x_i^{(k)}-x_i\right|\to 0\left(k\to\infty\right)$，

因为$\left|x_j^{(k)}-x_j\right|\le\max_i\left|x_i^{(k)}-x_i\right|\left(j=1,2,\cdots,n\right)$，所以对$j=1,2,\cdots,n$，有

$$x_j^{(k)}-x_j\to 0\left(k\to\infty\right)$$

因此$\lim_{k\to\infty}x^{(k)}=x$。

例　设$a=(1,1,\cdots,1)^T\in C^n$，且　$x^{(k)}=\left(1+\dfrac{1}{2^k},1+\dfrac{1}{3^k},\cdots,1+\dfrac{1}{(n+1)^k}\right)^T$

则　$\lim\limits_{k\to\infty}x^{(k)}=a$

证明：由于

$$\lim_{k\to\infty}\left\|x^{(k)}-a\right\|_\infty=\lim_{k\to\infty}\max_{1\le i\le n}\left\{\left|x_i^{(k)}-a_i\right|\right\}$$

$$=\lim_{k\to\infty}\max_{1\le i\le n}\left\{\frac{1}{(i+1)^k}\right\}=\lim_{k\to\infty}\frac{1}{2^k}=0$$

由向量范数的等价关系可知，在$\|\cdot\|_\infty$意义下收敛，在其他任意向量范数意义下也一定收敛，所以有$\lim\limits_{k\to\infty}x^{(k)}=a$

§6.2 矩阵范数

- 基本概念

  矩阵范数的定义 6.2.1 及其等价性定理 6.2.1

- 相容矩阵范数

  相容矩阵范数的定义 6.2.2 及相关定理 6.2.2，6.2.3

- 算子范数（从属范数、诱导范数：Induced Norm）

  算子范数的定义 6.2.3

  有关定理

基本概念

定义 6.2.1 设$\|A\|$是以$C^{m\times n}$中的矩阵A为自变量的非负实值函数，如果它满足以下三个条件：

（1） **非负性**：当$A\neq 0$时，$\|A\|>0$；当$A=0$时，$\|A\|=0$；

（2） **齐次性**：对任意$k\in C$，$A\in C^{m\times n}$，有$\|kA\|=|k|\|A\|$

（3） **三角不等式**：对任意$A,B\in C^{m\times n}$，有$\|A+B\|\leq\|A\|+\|B\|$，

则称$\|A\|$为$m\times n$矩阵A的范数。

容易导出：$\big|\|A\|-\|B\|\big|\leq\|A-B\|$

例 6.2.1 对于 $A=\left(a_{ij}\right)\in C^{m\times n}$，令

$$\|A\|_1' \equiv \sum_{i=1}^{m}\sum_{j=1}^{n}\left|a_{ij}\right|$$

$$\|A\|_\infty' \equiv \max_{i,j}\left|a_{ij}\right|$$

$$\|A\|_F \equiv \left(\sum_{i=1}^{m}\sum_{j=1}^{n}\left|a_{ij}\right|^2\right)^{\frac{1}{2}} = \left(tr\left(A^H A\right)\right)^{\frac{1}{2}}$$

容易证明，$\|\cdot\|_1'$，$\|\cdot\|_\infty'$ 和 $\|\cdot\|_F$ 都是 $C^{m\times n}$ 上的矩阵范数。$\|A\|_F$ 称为 A 的 Frobenius 范数，是 $C^{m\times n}$ 中的内积 $(A,B)=tr\left(B^H A\right)$ 导出的矩阵范数，因此是向量 Euclid 范数的自然推广。

因为$C^{m\times n}$是复数域上 mn 维线性空间，则由定理 6.1.7 得如下矩阵范数等价性定理

定理 6.2.1 设$\|\cdot\|_a$，$\|\cdot\|_b$是$C^{m\times n}$上的矩阵范数，则存在两个仅与$\|\cdot\|_a$，$\|\cdot\|_b$有关的正常数d_1，d_2使得

$$d_1\|A\|_b \le \|A\|_a \le d_2\|A\|_b, \forall A\in C^{m\times n}$$

则称范数$\|\underset{A}{\cancel{a}}\|_a$与$\|\underset{A}{\cancel{a}}\|_b$是等价的。

相容矩阵范数

我们经常遇到矩阵之间或矩阵与向量之间的乘法运算，对 $A\in C^{m\times n}$，$B\in C^{n\times k}$，希望对定义在 $C^{m\times n}$ 与 $C^{n\times k}$ 上的同一矩阵范数 $\|\cdot\|$ 有

$$\|AB\|\le\|A\|\|B\|$$

满足该条件的矩阵范数 $\|\cdot\|$ 成为具备相容性条件。

称为

- Frobenius 范数 $\|\cdot\|_F$ 具备相容性条件
- 矩阵范数 $\|\cdot\|_1'$ 也具备相容性条件
- 并非所有的矩阵范数都具备相容性条件：如矩阵范数 $\|\cdot\|_\infty'$

■矩阵范数$\|\cdot\|_1^{'}$具备相容性条件：对$A\in C^{m\times n}$，$B\in C^{n\times k}$

$$
\begin{aligned}
\|AB\|_1^{'} &= \sum_{i=1}^{m}\sum_{j=1}^{k}\left|\sum_{l=1}^{n}a_{il}b_{lj}\right| \le \sum_{i=1}^{m}\sum_{j=1}^{k}\left(\sum_{l=1}^{n}|a_{il}||b_{lj}|\right) \\
&\le \sum_{i=1}^{m}\sum_{j=1}^{k}\left\{\left(\sum_{l=1}^{n}|a_{il}|\right)\left(\sum_{l=1}^{n}|b_{lj}|\right)\right\} = \left(\sum_{i=1}^{m}\sum_{l=1}^{n}|a_{il}|\right)\left(\sum_{j=1}^{k}\sum_{l=1}^{n}|b_{lj}|\right) \\
&= \|A\|_1^{'}\|B\|_1^{'}
\end{aligned}
$$

■对矩阵

$$A=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}, B=\begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}$$

有$\|A\|_\infty^{'}=1$，$\|B\|_\infty^{'}=1$，而$\|AB\|_\infty^{'}=2>\|A\|_\infty^{'}\|B\|_\infty^{'}$

定义 6.2.2 设$\|\cdot\|_a$，$\|\cdot\|_b$和$\|\cdot\|_g$分别是$C^{m\times n}$，$C^{n\times k}$和$C^{m\times k}$上的矩阵范数，如果

$$\|AB\|_g \le \|A\|_a \|B\|_b, \forall A \in C^{m\times n}, \forall B \in C^{n\times k}$$

则称$\|\cdot\|_a$，$\|\cdot\|_b$和$\|\cdot\|_g$是相容的。特别地，如果$C^{n\times n}$上的矩阵范数$\|\cdot\|$满足

$$\|AB\| \le \|A\| \|B\|, \forall A, B \in C^{n\times n}$$

则称$\|\cdot\|$是自相容的矩阵范数，或简称为相容范数。

定义 6.2.2 包括了矩阵范数与向量范数的相容性定义。例如，矩阵 Frobenius 范数与向量 Euclid 范数是相容的，

$$\|Ax\|_2 \le \|A\|_F \|x\|_2, \forall A \in C^{n\times n}, \forall x \in C^n$$

定理 6.2.2 设$\|\cdot\|$是$C^{n\times n}$上的相容矩阵范数，则在C^n上存在与$\|\cdot\|$相容的向量范数。

证明：任取一非零向量$a\in C^n$，定义

$$\|x\|_a=\|xa^T\|,x\in C^n$$

容易验证$\|x\|_a$是C^n上的向量范数，并且

$$\|Ax\|_a=\|Axa^T\|\leq\|A\|\|xa^T\|=\|A\|\|x\|_a$$

即矩阵范数$\|\cdot\|$与向量范数$\|\cdot\|_a$相容。

定理 6.2.3 设$\|\cdot\|$是$C^{n\times n}$上的任一相容矩阵范数，则对任意$A\in C^{n\times n}$有

$$|l_i|\le\|A\|,\forall l_i\in l(A)$$

证明：由定理 6.2.2 知在C^n上存在与$\|\cdot\|$相容的向量范数$\|\cdot\|_a$。设$x_i\in C^n$是对应于特征值l_i的特征向量，则

$$Ax_i=l_i x_i,x_i\ne 0$$

从而由

$$|l_i|\|x_i\|_a=\|l_i x_i\|_a=\|Ax_i\|_a\le\|A\|\|x_i\|_a$$

和$\|x_i\|_a>0$即可得证。

$\rho(A)=\max\limits_i|\lambda_i| \le \|A\|$

定理 6.2.3 可用于谱半径的估计：$\cancel{r(A)=l_{\max}(A)\le\|A\|}$

算子范数

问题：对于给定的向量范数，能否定义与之相容的矩阵范数？

引理 6.2.1 设$\|\cdot\|_u$是C^n上的向量范数，则点集

$$j_u = \left\{ x \in C^n \,\middle|\, \|x\|_u = 1 \right\}$$

是C^n中的有界闭集。

引理 6.2.2 设$\|\cdot\|_m$是C^m上的向量范数，$A \in C^{m \times n}$，则$\|Ax\|_m$是$x \in C^n$的连续函数。

如果$\|\cdot\|_m$和$\|\cdot\|_u$分别是C^m和C^n上~~的向量范数上~~的两个向量范数，则由引理 6.2.1 和引理 6.2.2 知，$\|Ax\|_m$在j_u上达到最大值。

定理 6.2.4 设$\|\cdot\|_m$和$\|\cdot\|_u$分别是C^m和C^n上的两个向量范数，对$A\in C^{m\times n}$，令

$$\|A\|_{m,u}=\max_{\|x\|_u=1}\|Ax\|_m$$

则$\|\cdot\|_{m,u}$是$C^{m\times n}$上的矩阵范数，并且$\|\cdot\|_m$，$\|\cdot\|_u$和$\|\cdot\|_{m,u}$相容。

证明：对任意$x\in C^n$且$x\neq 0$，则$\dfrac{x}{\|x\|_u}\in j_u$，从而有$\left\|A\left(\dfrac{x}{\|x\|_u}\right)\right\|_m\leq\|A\|_{m,u}$

于是 $\|Ax\|_m\leq\|A\|_{m,u}\|x\|_u$，且此式对$x=0$也成立。

下面证明定理中所定义的$\|\cdot\|_{m,u}$满足矩阵范数定义中的三个条件。

（1）如果$A\neq 0$，则必有e_i使$Ae_i\neq 0$，于是由$\|Ax\|_m\leq\|A\|_{m,u}\|x\|_u$，有

$0<\|Ae_i\|_m \le \|A\|_{m,u}\|e_i\|_u$，因此当$A\neq 0$时，$\|A\|_{m,u}>0$；显然当$A=0$时，$\|A\|_{m,u}=0$。

（2）任取$k\in C$，有 $\|kA\|_{m,u}=\max\limits_{\|x\|_u=1}\|kAx\|_m=\max\limits_{\|x\|_u=1}|k|\|Ax\|_m=|k|\|A\|_{m,u}$

（3）任取$A,B\in C^{m\times n}$，设x_0满足$\|x_0\|_u=1$并且$\|(A+B)x_0\|_m=\|A+B\|_{m,u}$，则由$\|Ax\|_m\le\|A\|_{m,u}\|x\|_u$，有

$$\begin{aligned}\|A+B\|_{m,u}&=\|(A+B)x_0\|_m\le\|Ax_0\|_m+\|Bx_0\|_m\\&\le\|A\|_{m,u}\|x_0\|_u+\|B\|_{m,u}\|x_0\|_u=\|A\|_{m,u}+\|B\|_{m,u}\end{aligned}$$

由$\|Ax\|_m\le\|A\|_{m,u}\|x\|_u$可见，$\|\cdot\|_m$，$\|\cdot\|_u$和$\|\cdot\|_{m,u}$相容。得证。

定义 6.2.3 设$\|\cdot\|_m$和$\|\cdot\|_u$分别是C^m和C^n上的向量范数，由

$$\|A\|_{m,u} = \max_{\|x\|_u=1} \|Ax\|_m$$

定义的非负实值函数$\|\cdot\|_{m,u}$叫做$C^{m\times n}$上的算子范数或称为由向量范数$\|\cdot\|_m$和$\|\cdot\|_u$导出的矩阵范数。

■由$\|Ax\|_m \le \|A\|_{m,u}\|x\|_u$知：算子范数与诱导它的向量范数是相容的。

算子范数之间的相容性

定理 6.2.5 设$\|\cdot\|_m$，$\|\cdot\|_u$和$\|\cdot\|_w$分别是C^m，C^n和C^k上的向量范数，如果按照$\|A\|_{m,u}=\max\limits_{\|x\|_u=1}\|Ax\|_m$分别定义$C^{m\times n}$，$C^{n\times k}$和$C^{m\times k}$上的算子范数$\|\cdot\|_{m,u}$，$\|\cdot\|_{u,w}$和$\|\cdot\|_{m,w}$，则$\|AB\|_{m,w}\le\|A\|_{m,u}\|B\|_{u,w},\forall A\in C^{m\times n},\forall B\in C^{n\times k}$

证明：设$x_0\in C^k$并且$\|x_0\|_w=1$，$\|ABx_0\|_m=\|AB\|_{m,w}$，则由$\|Ax\|_m\le\|A\|_{m,u}\|x\|_u$得

$$\|AB\|_{m,w}=\|ABx_0\|_m=\|A(Bx_0)\|_m\le\|A\|_{m,u}\|Bx_0\|_u$$

$$\le\|A\|_{m,u}\|B\|_{u,w}\|x_0\|_w=\|A\|_{m,u}\|B\|_{u,w}$$

定理 6.2.6 设$\|\cdot\|$是C^n上的向量范数，则在$C^{n\times n}$上由向量范数$\|\cdot\|$导出的矩阵范数$\|\cdot\|$是相容矩阵范数，即

$$\|AB\| \le \|A\|\|B\|, \forall A, B \in C^{n\times n}$$

证明：令$m=n=k$，且$\|\cdot\|_{m,u}$，$\|\cdot\|_{u,w}$和$\|\cdot\|_{m,w}$为C^n上的同一个向量范数$\|\cdot\|$，由定理 6.2.5 即得。

$$\|A\|_p$$

如果把C^n**上的向量范数**$\|\cdot\|_p\ (p=1,2,\infty)$**限制到**$C^m$**上，恰好是**$C^m$**上的向量范数**$\|\cdot\|_p$**。由定理** 6.2.4**，可以得到**$C^{m\times n}$**上的算子范数**$\|\cdot\|_p$

$$\|A\|_p=\max_{\|x\|_p=1}\|Ax\|_p,A\in C^{m\times n}\ (p=1,2,\infty)$$

并且由定理 6.2.5 **知，这些算子范数都是相容的，即**

$$\|AB\|_p\le\|A\|_p\|B\|_p,A\in C^{m\times n},B\in C^{n\times k}$$

定理 6.2.7 设 $A=\left(a_{ij}\right)\in C^{m\times n}$，则有

$$\|A\|_1=\max_{1\le j\le n}\sum_{i=1}^{m}\left|a_{ij}\right|$$

$$\|A\|_2=\left(\lambda_{\max}\left(A^H A\right)\right)^{\frac{1}{2}}$$

$$\|A\|_\infty=\max_{1\le i\le m}\sum_{j=1}^{n}\left|a_{ij}\right|$$

通常将 $\|A\|_1$ 称为 A 的列和范数，$\|A\|_2$ 称为 A 的谱范数，$\|A\|_\infty$ 称为 A 的行和范数。

例 6.2.2 设

$$A=\begin{pmatrix} 2 & -1 & 0 \\ 0 & 2 & 3 \\ 1 & 2 & 0 \end{pmatrix}$$

计算$\|A\|_1$，$\|A\|_2$，$\|A\|_\infty$**和**$\|A\|_F$

解 $\|A\|_1=5$，$\|A\|_\infty=5$。

因为$A^H A=\begin{pmatrix} 5 & 0 & 0 \\ 0 & 9 & 6 \\ 0 & 6 & 9 \end{pmatrix}$，**所以**

$$\|A\|_F=\left(tr\left(A^H A\right)\right)^{\frac{1}{2}}=\sqrt{23}，\ \|A\|_2=\left(l_{\max}\left(A^H A\right)\right)^{\frac{1}{2}}=\sqrt{15}$$

Frobenius 范数 $\|A\|_F$

$$\|A\|_F \equiv \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{\frac{1}{2}} = \left(tr\left(A^H A \right) \right)^{\frac{1}{2}}$$

- **范数 $\|A\|_F$ 是相容范数，但不是算子范数**
- **$\|A\|_F$ 不但计算简单，而且具有如下性质**

定理 6.2.8 **设** $A \in C^{m\times n}$，U **和** V **分别为** m **阶和** n **阶酉矩阵，则**

$$\|UAV\|_F = \|A\|_F$$

证明：

$$\|UAV\|_F^2 = tr\left[(UAV)^H (UAV) \right] = tr\left(V^H A^H AV \right)$$

$$= tr\left(V^{-1} A^H AV \right) = tr\left(A^H A \right) = \|A\|_F^2$$

谱范数$\|A\|_2$　$\|A\|_2=\left(\lambda_{\max}\left(A^H A\right)\right)^{\frac{1}{2}}$

■不便计算，但有许多很好的性质，所以在理论研究中常使用。

定理 6.2.9 **设**$A\in C^{m\times n}$，**则**

$$\|A\|_2=\max_{\substack{\|x\|_2=1\\ \|y\|_2=1}}\left|y^H Ax\right|$$

$$\left\|A^H\right\|_2=\left\|A^T\right\|_2=\|A\|_2$$

$$\left\|A^H A\right\|_2=\|A\|_2^2$$

$$\|A\|_2^2\le\|A\|_1\|A\|_\infty$$

并且对 m **阶酉矩阵** U **和** n **阶酉矩阵** V **有**$\|UAV\|_2=\|A\|_2$

§6.3 矩阵序列与矩阵级数

本节利用前两节建立的范数理论讨论矩阵序列的极限和矩阵级数等分析概念。

■矩阵序列的极限

介绍矩阵序列的收敛性定义、矩阵序列收敛的充分必要条件、矩阵序列的极限运算性质等

■矩阵级数

矩阵级数定义、矩阵级数收敛、矩阵级数绝对收敛的定义，矩阵级数（绝对）收敛的充要条件等

矩阵序列的极限

定义 6.3.1 **设有矩阵序列**$\left\{A^{(k)}\right\}$**，其中**$A^{(k)}=\left(a_{ij}^{(k)}\right)\in C^{m\times n}$**。如果当**$k\to\infty$**时，矩阵**$A^{(k)}$**的每一个元素**$a_{ij}^{(k)}$**都有极限**$a_{ij}$**，即**

$$\lim_{k\to\infty} a_{ij}^{(k)}=a_{ij},\quad 1\le i\le m;1\le j\le n$$

则称矩阵序列$\left\{A^{(k)}\right\}$**是收敛的，并把矩阵**$A=\left(a_{ij}\right)\in C^{m\times n}$**称为**$\left\{A^{(k)}\right\}$**的极限，或称**$\left\{A^{(k)}\right\}$**收敛于**$A$**，记为**

$$\lim_{k\to\infty} A^{(k)}=A\text{或}A^{(k)}\to A\left(k\to\infty\right)$$

定理 6.3.1 设$\|\cdot\|$是$C^{m\times n}$上任一矩阵范数，$C^{m\times n}$中矩阵序列$\{A^{(k)}\}$收敛于 A 的充分必要条件是

$$\lim_{k\to\infty}\left\|A^{(k)}-A\right\|=0$$

关于矩阵序列的极限运算有如下性质：

（1）设$\left\{A^{(k)}\right\}$，$\left\{B^{(k)}\right\}$是$C^{m\times n}$中矩阵序列，且$\lim\limits_{k\to\infty} A^{(k)} = A$，$\lim\limits_{k\to\infty} B^{(k)} = B$，则

$$\lim_{k\to\infty}\left(aA^{(k)} + bB^{(k)}\right) = aA + bB$$

其中$a,b\in C$是常数。

（2）设$\left\{A^{(k)}\right\}$与$\left\{B^{(k)}\right\}$是分别是$C^{m\times n}$与$C^{n\times k}$中矩阵序列，且$\lim\limits_{k\to\infty} A^{(k)} = A$，$\lim\limits_{k\to\infty} B^{(k)} = B$，则

$$\lim_{k\to\infty} A^{(k)}B^{(k)} = AB$$

由定义 6.3.1 和定理 6.3.1 容易验证。

定理 6.3.2 设$A\in C^{n\times n}$，$\lim\limits_{k\to\infty} A^k = 0$的充分必要条件是$r(A)<1$。

证明：对$A\in C^{n\times n}$，**存在 n 阶可逆矩阵 P 使得**

$$P^{-1}AP = J = diag\left(J_1, J_2, \cdots, J_s\right)$$

其中

$$J_i = \begin{pmatrix} l_i & 1 & & & 0 \\ & l_i & 1 & & \\ & & \ddots & \ddots & \\ & 0 & & l_i & 1 \\ & & & & l_i \end{pmatrix}_{n_i\times n_i}$$

则

$$P^{-1}A^kP = J^k = diag\left(J_1^k, J_2^k, \cdots, J_s^k\right)$$

因此$\lim\limits_{k\to\infty} A^k = 0 \Leftrightarrow \lim\limits_{k\to\infty} J^k = 0 \Leftrightarrow \lim\limits_{k\to\infty} J_i^k = 0\left(i = 1, 2, \cdots, s\right)$**。而**

$$J_i^k = \begin{pmatrix} f_k\left(l_i\right) & f_k^{'}\left(l_i\right) & \frac{1}{2!} f_k^{''}\left(l_i\right) & \cdots & \frac{1}{\left(n_i - 1\right)!} f_k^{\left(n_i - 1\right)}\left(l_i\right) \\ & f_k\left(l_i\right) & f_k^{'}\left(l_i\right) & \ddots & \vdots \\ & & & & \vdots \\ & & \ddots & \ddots & \frac{1}{2!} f_k^{''}\left(l_i\right) \\ & 0 & & \ddots & f_k^{'}\left(l_i\right) \\ & & & & f_k\left(l_i\right) \end{pmatrix}$$

其中 $f_k\left(l\right) = l^k$**。因为对任一多项式**$g\left(l\right)$**，当**$k \to \infty$**时，**$g\left(k\right)\left|l_i\right|^k \to 0$

$\Leftrightarrow \left|l_i\right| < 1$**，而**$\left|l_i\right| < 1\left(i = 1, 2, \cdots, s\right) \Leftrightarrow r\left(A\right) < 1$**。**

定理 6.3.3 设$A\in C^{n\times n}$，如果存在$C^{n\times n}$上的一种相容矩阵范数$\|\cdot\|$使$\|A\|<1$，则$\lim\limits_{k\to\infty} A^k = 0$。

定理 6.3.4 设$\|\cdot\|$是$C^{n\times n}$上的相容矩阵范数，则对任意$A\in C^{n\times n}$，有

$$r(A)=\lim_{k\to\infty}\left\|A^k\right\|^{\frac{1}{k}}$$

证明：因为$\left(r(A)\right)^k=r\left(A^k\right)\le\left\|A^k\right\|$，所以对所有$k=1,2,\cdots$，有$r(A)\le\left\|A^k\right\|^{\frac{1}{k}}$。另一方面，对任意$e>0$，矩阵$\tilde{A}=\left[r(A)+e\right]^{-1}A$的谱半径严格小于1。由定理6.3.2知$\lim\limits_{k\to\infty}\tilde{A}^k=0$，于是当$k\to\infty$时$\left\|\tilde{A}^k\right\|\to 0$。因此，存在正整数 K 使得$k>K$时$\left\|\tilde{A}^k\right\|<1$，即对所有$k>K$有$\left\|A^k\right\|\le\left[r(A)+e\right]^k$或$\left\|A^k\right\|^{\frac{1}{k}}\le r(A)+e$。故$\lim\limits_{k\to\infty}\left\|A^k\right\|^{\frac{1}{k}}=r(A)$。

矩阵级数

定义 6.3.2 设$\left\{A^{(k)}\right\}$是$C^{m\times n}$的矩阵序列，其中$A^{(k)}=\left(a_{ij}^{(k)}\right)\in C^{m\times n}$，无穷和

$$A^{(1)}+A^{(2)}+\cdots+A^{(k)}+\cdots$$

称为**矩阵级数**，记为$\sum_{k=1}^{\infty}A^{(k)}$。对正整数$k\geq 1$，记$S^{(k)}=\sum_{i=1}^{k}A^{(i)}$，称$S^{(k)}$为矩阵级数$\sum_{k=1}^{\infty}A^{(k)}$的**部分和**。如果矩阵序列$\left\{S^{(k)}\right\}$收敛，且有极限 S，即$\lim_{k\to\infty}S^{(k)}=S$，则称矩阵级数$\sum_{k=1}^{\infty}A^{(k)}$**收敛**，并称 S 为矩阵级数$\sum_{k=1}^{\infty}A^{(k)}$的**和**，记为$\sum_{k=1}^{\infty}A^{(k)}=S$。不收敛的矩阵级数称为**发散**的。

■ $\sum_{k=1}^{\infty} A^{(k)}$ **收敛** $\Leftrightarrow mn$ **个数项级数** $\sum_{k=1}^{\infty} a_{ij}^{(k)}\left(i=1,2,\cdots,m,\ j=1,2,\cdots n\right)$**都收敛**

■ **若矩阵级数** $\sum_{k=1}^{\infty} A^{(k)}$ **收敛，则** $\lim_{k\to\infty} A^{(k)}=0$ **；**

■ **若矩阵级数** $\sum_{k=1}^{\infty} A^{(k)}=S_1$**，** $\sum_{k=1}^{\infty} B^{(k)}=S_2$**，** $a,b\in C$ **是常数，则**

$$\sum_{k=1}^{\infty}\left(aA^{(k)}+bB^{(k)}\right)=aS_1+bS_2$$

■ **设** $P\in C^{m\times n}$**，** $Q\in C^{n\times n}$**，若矩阵级数** $\sum_{k=1}^{\infty} A^{(k)}$ **收敛，则** $\sum_{k=1}^{\infty} PA^{(k)}Q$ **收敛，且**

$$\sum_{k=1}^{\infty} PA^{(k)}Q=P\left(\sum_{k=1}^{\infty} A^{(k)}\right)Q$$

定义 6.3.3 设 $\sum_{k=1}^{\infty} A^{(k)}$ 是矩阵级数，其中 $A^{(k)}=\left(a_{ij}^{(k)}\right)\in C^{m\times n}$，如果 mn 个数项级数 $\sum_{k=1}^{\infty} a_{ij}^{(k)}\,(i=1,2,\cdots,m,\ j=1,2,\cdots n)$ 都绝对收敛，则称矩阵级数 $\sum_{k=1}^{\infty} A^{(k)}$ 绝对收敛。

定理 6.3.5 矩阵级数 $\sum_{k=1}^{\infty} A^{(k)}$（其中 $A^{(k)}=\left(a_{ij}^{(k)}\right)\in C^{m\times n}$）绝对收敛的充分必要条件是数项级数 $\sum_{k=1}^{\infty}\left\|A^{(k)}\right\|$ 收敛，其中 $\|\cdot\|$ 是 $C^{m\times n}$ 上的任一矩阵范数。

与数项级数类似，若矩阵级数 $\sum_{k=1}^{\infty} A^{(k)}$ 绝对收敛，则它一定收敛，并且任意交换各项的求和次序所得的新级数仍收敛，和也不改变。

对矩阵级数也有幂级数的概念，矩阵幂级数是研究矩阵函数的重要工具。

定义 6.3.4 设 $A \in C^{n\times n}$，形如

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

的矩阵级数~~成~~为矩阵幂级数。

称为

由定理 6.3.5 可得如下定理：

定理 6.3.6 设 $A \in C^{n\times n}$，如果数项级数 $\sum_{k=0}^{\infty} |c_k| \|A\|^k$ 收敛，则矩阵幂级数 $\sum_{k=0}^{\infty} c_k A^k$ 绝对收敛，其中 $\|\cdot\|$ 是 $C^{n\times n}$ 上的某种相容矩阵范数。

推论 6.3.1 设 $A\in C^{n\times n}$，如果 $C^{n\times n}$ 上的某种相容矩阵范数 $\|\cdot\|$ 使得 $\|A\|$ 在幂级数

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

的收敛圆内，则矩阵幂级数 $\sum_{k=0}^{\infty} c_k A^k$ 绝对收敛。

定理 6.3.7 设 $A\in C^{n\times n}$，并且幂级数 $\sum_{k=0}^{\infty} c_k z^k$ 的收敛半径为 R。如果 $r(A)<R$，则矩阵幂级数 $\sum_{k=0}^{\infty} c_k A^k$ 绝对收敛；如果 $r(A)>R$，则矩阵幂级数 $\sum_{k=0}^{\infty} c_k A^k$ 发散。

推论 6.3.2 如果幂级数$\sum_{k=0}^{\infty} c_k z^k$在整个平面上都收敛，则对任意$A \in C^{n \times n}$，矩阵幂级数$\sum_{k=0}^{\infty} c_k A^k$收敛。

定理 6.3.8 设$A \in C^{n \times n}$，如果$\|A\| < 1$，则矩阵$I - A$非奇异，并且

$$\left\|(I - A)^{-1}\right\| \leq \frac{1}{1 - \|A\|}$$

其中$\|\cdot\|$是$C^{n \times n}$上的相容矩阵范数且满足$\|I\| = 1$

Neumann级数

定理 6.3.9 设 $A\in C^{n\times n}$，则矩阵幂级数 $\sum_{k=0}^{\infty} c_k A^k$ 收敛的充分必要条件是 $r(A)<1$，并且其和为 $(I-A)^{-1}$，即

$$\sum_{k=0}^{\infty} A^k=(I-A)^{-1}$$

此外，如果 $\|A\|<1$，则有

$$\left\|(I-A)^{-1}-\sum_{k=0}^{m} A^k\right\|\le\frac{\|A\|^{m+1}}{1-\|A\|}$$

其中，$\|\cdot\|$ 是 $C^{n\times n}$ 上的相容矩阵范数且满足 $\|I\|=1$

§6.4 矩阵扰动分析

- 为了解决科学和工程技术中的实际问题，人们依据物理、力学等规律建立问题的数学模型，并根据数学模型提出求解数学问题的数值计算方法。
- 在数学问题的求解过程中，数值计算方法引起的截断误差和计算环境引起的舍入误差会影响计算结果的精度。
- 为了分析这些误差对数学问题解的影响，人们将其归结为原始数据的扰动（或摄动）对解的影响。
- 我们需要研究该扰动引起了问题解的多大变化，即问题解的稳定性。

考虑一个 2 阶线性方程组

$$\begin{pmatrix} 1 & 0.99 \\ 0.99 & 0.98 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

可以验证，该方程组的精确解为 $x_1=100$**，** $x_2=-100$

如果系数矩阵有一扰动 $\begin{pmatrix} 0 & 0 \\ 0 & 0.01 \end{pmatrix}$**，并且右端项也有一扰动** $\begin{pmatrix} 0 \\ 0.001 \end{pmatrix}$**，则扰动后的线性方程组为**

$$\begin{pmatrix} 1 & 0.99 \\ 0.99 & 0.99 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 1 \\ 1.001 \end{pmatrix}$$

可以验证，这个方程组的精确解为 $x_1+dx_1=-0.1$**，** $x_2+dx_2=10/9$

可见，系数矩阵和右端项的微小扰动引起了解的强烈变化。

- 原始数据的扰动对问题解的影响程度取决于问题本身的固有性质。
  - 如果原始数据的小扰动引起问题解的很大变化，则称该问题是病态的（敏感的）或不稳定的；否则，就称该问题时良态的（不敏感的）或稳定的
- 矩阵扰动分析就是研究矩阵元素的变化对矩阵问题解的影响，它对矩阵论和矩阵计算都由重要意义。
- 本节内容：
  - 矩阵 A 的逆矩阵的扰动分析
  - 以 A 为系数矩阵的线性方程组解的扰动分析
  - 矩阵特征值的扰动分析*
- 假定$\|\cdot\|$是$C^{n\times n}$上满足$\|I\|=1$的相容范数矩阵。

矩阵逆的扰动分析

设矩阵$A\in C^{n\times n}$并且A非奇异，经扰动变为$A+E$，其中$E\in C^{n\times n}$称为扰动矩阵。我们需要解决：

- 什么条件下，$A+E$非奇异
- 当$A+E$非奇异时，$\left(A+E\right)^{-1}$与A^{-1}的近似程度

定理 6.4.1 设$A, E \in C^{n\times n}$， $B = A + E$，如果 A 与 B 均非奇异，则

$$\frac{\|B^{-1}-A^{-1}\|}{\|A^{-1}\|} \le \|A\|\|B^{-1}\|\frac{\|E\|}{\|A\|}$$

证明：由$B^{-1}-A^{-1}=A^{-1}(A-B)B^{-1}=-A^{-1}EB^{-1}$，可得

$$\|B^{-1}-A^{-1}\| \le \|A^{-1}\|\|B^{-1}\|\|E\|$$

于是

$$\frac{\|B^{-1}-A^{-1}\|}{\|A^{-1}\|} \le \|B^{-1}\|\|E\| \le \|A\|\|B^{-1}\|\frac{\|E\|}{\|A\|}$$

定理 6.4.2 设 $A\in C^{n\times n}$ 是非奇异矩阵，$E\in C^{n\times n}$ 满足条件

$$\left\|A^{-1}E\right\|<1$$

则 $A+E$ 非奇异，并且有

$$\left\|(A+E)^{-1}\right\|\le\frac{\left\|A^{-1}\right\|}{1-\left\|A^{-1}E\right\|}$$

$$\frac{\left\|(A+E)^{-1}-A^{-1}\right\|}{\left\|A^{-1}\right\|}\le\frac{\left\|A^{-1}E\right\|}{1-\left\|A^{-1}E\right\|}$$

证明：灵活运用定理 6.3.8 和定理 6.3.9。略。

因为$\|A^{-1}E\| \le \|A^{-1}\|\|E\|$，所以由定理 6.4.2 可得如下推论：

推论 6.4.1 设$A \in C^{n\times n}$是非奇异矩阵，$E \in C^{n\times n}$满足条件$\|A^{-1}\|\|E\|<1$，则$A+E$非奇异，并且有

$$\frac{\left\|(A+E)^{-1}-A^{-1}\right\|}{\|A^{-1}\|} \le \frac{k(A)\dfrac{\|E\|}{\|A\|}}{1-k(A)\dfrac{\|E\|}{\|A\|}}$$

其中$k(A)=\|A\|\|A^{-1}\|$，称为 A 关于求逆的条件数。

■ $k(A)$反映了A^{-1}对于 A 的扰动的敏感性，$k(A)$愈大，$(A+E)^{-1}$与A^{-1}的相对误差就愈大。如果$k(A)$很大，则矩阵 A 关于求逆是病态的。

线性方程组解的扰动分析

对线性方程组

$$Ax = b$$

有

如果系数矩阵 A **和右端项** b **分别由扰动** E **和**db**，则扰动后方程组为**

$$(A+E)(x+dx) = b+db$$

问题：

- **什么条件下扰动后的方程组有唯一解？**
- **估计解的相对误差。**

定理 6.4.3 设$A\in C^{n\times n}$是非奇异矩阵，b，$db\in C^{n}$，x 是方程组$Ax=b$的解，如果$E\in C^{n\times n}$满足条件$\left\|A^{-1}\right\|\left\|E\right\|<1$，则方程组$(A+E)(x+dx)=b+db$有唯一解$x+dx$，并且有

$$\frac{\left\|dx\right\|}{\left\|x\right\|}\leq\frac{k(A)}{1-k(A)\frac{\left\|E\right\|}{\left\|A\right\|}}\left(\frac{\left\|E\right\|}{\left\|A\right\|}+\frac{\left\|db\right\|}{\left\|b\right\|}\right)$$

其中$k(A)=\left\|A\right\|\left\|A^{-1}\right\|$，向量范数$\left\|\cdot\right\|$与矩阵范数$\left\|\cdot\right\|$相容。

证明：因为$\|A^{-1}\|\|E\|<1$，由推论 6.4.1 知$A+E$非奇异，因此方程组

$$(A+E)(x+dx)=b+db$$

有唯一解$x+dx$。由$Ax=b$得$x=A^{-1}b$，而从$(A+E)(x+dx)=b+db$有

$$dx=(A+E)^{-1}(b+db)-x=(A+E)^{-1}b+(A+E)^{-1}db-x$$

$$=(A+E)^{-1}db-x+(I+A^{-1}E)^{-1}A^{-1}b$$

$$=(A+E)^{-1}db-\left[I-(I+A^{-1}E)^{-1}\right]x=(A+E)^{-1}db-(I+A^{-1}E)^{-1}A^{-1}Ex$$

由$\|A^{-1}E\|\le\|A^{-1}\|\|E\|$，定理 6.3.8 和定理 6.4.2 可得

$$\|dx\|\le\frac{\|A^{-1}\|\|db\|}{1-\|A^{-1}E\|}+\frac{\|A^{-1}E\|\|x\|}{1-\|A^{-1}E\|}\le\frac{\|A^{-1}\|\|x\|}{1-\|A^{-1}\|\|E\|}\left(\frac{\|db\|}{\|x\|}+\|E\|\right)$$

因为$\|b\| \le \|A\|\|x\|$，则由上式即得

$$\frac{\|dx\|}{\|x\|} \le \frac{\|A^{-1}\|}{1-\|A^{-1}\|\|E\|}\left(\frac{\|db\|}{\|x\|}+\|E\|\right) \le \frac{k(A)}{1-k(A)\frac{\|E\|}{\|A\|}}\left(\frac{\|db\|}{\|b\|}+\frac{\|E\|}{\|A\|}\right)$$

■ $k(A)$反映了线性方程组$Ax=b$的解x的相对误差对于A和b的相对误差的依赖程度，因此也被称为求解线性方程组$Ax=b$的条件数。如果$k(A)$很大，则线性方程组$Ax=b$是病态的。

在求矩阵的逆或求解线性方程组时，可以通过变换降低矩阵的条件数，即所谓预处理或预条件。例如，对线性方程组

$\begin{pmatrix} 1 & 10^5 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 10^5 \\ 2 \end{pmatrix}$，因为 $A = \begin{pmatrix} 1 & 10^5 \\ 1 & 1 \end{pmatrix}$，所以 $A^{-1} = \frac{1}{10^5 - 1}\begin{pmatrix} -1 & 10^5 \\ 1 & -1 \end{pmatrix}$。

取矩阵范数 $\|\cdot\|$ 为 $\|\cdot\|_\infty$，则 $k(A) = \frac{\left(1 + 10^5\right)^2}{10^5 - 1} \approx 10^5$，因此矩阵 A 求逆和求解方程组 $Ax = b$ 都是病态的。如果用 10^5 除方程组的第一个方程，得 $Bx=c$，即 $\begin{pmatrix} 10^{-5} & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$，其中 $B = \begin{pmatrix} 10^{-5} & 1 \\ 1 & 1 \end{pmatrix}$，则 $k(B) \approx 4$，于是方程组 $Bx=c$ 是良态的。