第2章　范数理论及其应用

2.1 向量范数

2.2 矩阵范数

2.3 范数的一些应用

2.1 向量范数

1. 向量范数的概念及基本性质

2. 向量范数具体实例

3. 向量范数的等价性

4. 基于向量范数的收敛性

5*. 机器学习中的距离度量与相似性度量

1. 向量范数的概念及基本性质

假设V是数域K上的线性空间，对V的任一向量x，定义一个实值函数$\|x\|$，如果这个函数满足：

(1) 非负性：$\|x\|\geq 0$，且$\|x\|=0$当且仅当$x=0$；

(2) 齐次性：$\|kx\|=|k|\,\|x\|$，$k\in K$，$x\in V$；

(3) 三角不等式：$\|x+y\|\leq\|x\|+\|y\|$，$x,y\in V$，

则称$\|x\|$为V上的向量范数。

注：(2)中$|k|$表示绝对值($K=R$)或模($K=C$)。

向量的范数具有下列简单性质：

(1) 当 $\|x\| \neq 0$ 时，$\left\|\frac{1}{\|x\|}x\right\| = 1$ $\quad \because \left\|\frac{1}{\|x\|}x\right\| = \frac{1}{\|x\|}\|x\| = 1$

(2) $\forall x \in V$，$\|-x\| = \|x\|$ $\quad \because \|-x\| = |-1|\|x\| = \|x\|$

(3) $\forall x, y \in V$，$\|x\| - \|y\| \leq \|x - y\|$

$$\because \|x\| = \|(x-y)+y\| \leq \|x-y\| + \|y\| \Rightarrow \|x\| - \|y\| \leq \|x-y\|$$

(4) $\forall x, y \in V$，$\big|\|x\| - \|y\|\big| \leq \|x - y\|$

$$\|\boldsymbol{x}\| = \|(\boldsymbol{x}-\boldsymbol{y})+\boldsymbol{y}\| \leq \|\boldsymbol{x}-\boldsymbol{y}\| + \|\boldsymbol{y}\| \quad \Rightarrow \|\boldsymbol{x}\| - \|\boldsymbol{y}\| \leq \|\boldsymbol{x}-\boldsymbol{y}\|$$

同样 $\quad \|\boldsymbol{y}\| - \|\boldsymbol{x}\| \leq \|\boldsymbol{y}-\boldsymbol{x}\| = \|\boldsymbol{x}-\boldsymbol{y}\|$

虽然向量范数是定义在一般的线性空间上的，但是由于前面的讨论，我们知道任何*n*维线性空间在一个基下都代数同构于常用的*n*维复(或实)列向量空间，因此下面我们仅仅讨论n维复(或实) 列向量空间就足够了

下面讨论如下：

1.设$\|\cdot\|$为线性空间 V_n 的范数，任取它的一个基 $\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_n$,则对于任意向量 $\mathbf{x}$,它可以表示为

$$\mathbf{x}=\xi_1\mathbf{x}_1+\xi_2\mathbf{x}_2+\ldots+\xi_n\mathbf{x}_n$$

其中，$(\xi_1,\xi_2,\ldots,\xi_n)^T$ 为 x 的坐标。

由此定义 C^n(或 R^n)中的范数如下：

$$\|\xi\|_C=\varphi(\xi)=\|\xi_1\mathbf{x}_1+\xi_2\mathbf{x}_2+\ldots+\xi_n\mathbf{x}_n\|$$

则容易验证$\|\xi\|_C$确实为 C^n 中的范数.

2.反之，若$\|\xi\|_C$为 C^n 中的范数，定义 V_n 的范数如下：

$$\|\mathbf{x}\|=\phi(\mathbf{x})=\|\xi\|_C$$

其中 $\mathbf{x}=\xi_1\mathbf{x}_1+\xi_2\mathbf{x}_2+\ldots+\xi_n\mathbf{x}_n$。

则容易验证 $\phi(\mathbf{x})$确实为 V_n 的范数。

这个例子充分说明了一般线性空间的范数和 n 维复(或实)列向量空间的范数之间的关系。这也是为我们只讨论 n 维复(或实)列向量空间的范数的理由.

范数与函数

范数首先是一个函数，它将线性空间的任意向量映射为非负实数。

性质1. 范数是凸函数，

即 $\|(1-\lambda)x+\lambda y\|\leq(1-\lambda)\|x\|+\lambda\|y\|$

其中 $0\leq \lambda \leq 1$。

向量的范数类似于向量长度。

性质2. (范数的乘法)

若$\|\cdot\|$为线性空间V上的向量范数，

则$k\|\cdot\|$ 仍然为向量范数, 其中$k > 0$.

定义：设 $x = (\xi_1,\cdots,\xi_n)^T \in \mathbf{C}^n$，记 $|x| = (|\xi_1|,\cdots,|\xi_n|)^T \in R^n$，

如果 $\mathbf{C}^n$ 上的范数 v 满足条件：

$$v(x) = v(|x|),\quad \forall x \in \mathbf{C}^n。$$

则称 v 为绝对范数。

引理 范数 v 为绝对范数的充要条件是：

当 $|x| \leqslant |y|$ 时，则有 $v(x) \leqslant v(y)$。 (♣)

满足条件(♣)的范数又称为单调范数。

可见，绝对范数就是单调范数

以上引理结论的可查Ronger著《矩阵分析》参考书可得。在此不赘述。

性质 3. 设$\|\cdot\|_{comp}$为 R^m 上的范数，且对 $\mathbf{x}\in(R^+)^m$ 为单调增加的(即,若 $\mathbf{x},\mathbf{y}\in(R^+)^m$，且 $x_i\le y_i$,那么$\|\mathbf{x}\|_{comp}\le\|\mathbf{y}\|_{comp}$成立.),那么，对于给定的 m 个 n 维线性空间 V 上的范数$\|\cdot\|_i$,i=1,2,…,m,我们可以定义一个复合范数为

$$\|\mathbf{x}\|=\|U(\mathbf{x})\|_{comp},$$

其中，$U(\mathbf{x})=(\|\mathbf{x}\|_1,\|\mathbf{x}\|_2,\ldots,\|\mathbf{x}\|_m)^T$.

证明：非负性和齐次性是显然的，仅需证明三角不等式。

$$\|\mathbf{x}+\mathbf{y}\|=\|U(\mathbf{x}+\mathbf{y})\|_{comp}$$

$$\le\|U(\mathbf{x})+U(\mathbf{y})\|_{comp}(\text{因 } U(\mathbf{x}+\mathbf{y}) \le U(\mathbf{x})+U(\mathbf{y}))$$

$$\le\|U(\mathbf{x})\|_{comp}+\|U(\mathbf{y})\|_{comp}$$

$$=\|\mathbf{x}\|+\|\mathbf{y}\|$$

例如. 若$\|\cdot\|_f$和$\|\cdot\|_g$为线性空间 V 上的两个向量范数，则

(1). $\|\cdot\|_f + \|\cdot\|_g$ 为 V 上向量范数。

(2). $\max\{\|\cdot\|_f, \|\cdot\|_g\}$ 为 V 上向量范数。

(3) $[(\|\cdot\|_f)^2+(\|\cdot\|_g)^2]^{1/2}$ 为 V 上向量范数。

性质 4. (范数的合成)

设 n 维线性空间 $V= V_1\oplus V_2\oplus\ldots\oplus V_m$,

且 $\|\cdot\|_i, i=1,2,\ldots,m$, 为线性子空间 V_i 上

的范数， 而 $\|\cdot\|_{comp}$ 为 R^m 上的范数， 且对

$\mathbf{x}\in (R^+)^m$ 为单调增加的(即,若 $\mathbf{x},\mathbf{y}\in(R^+)^m$,

且 $x_i\leq y_i$,那么 $\|\mathbf{x}\|_{comp}\leq\|\mathbf{y}\|_{comp}$ 成立.),

则对任意 $x\in V$,存在唯一的分解

$$\mathbf{x}=\mathbf{x}_1+\mathbf{x}_2+\ldots+\mathbf{x}_m$$

其中 $\mathbf{x}_i\in V_i$,

这时定义 $\mathbf{x}$ 的范数为 $\|\mathbf{x}\|=\|U(\mathbf{x})\|_{comp}$,

其中， $U(\mathbf{x})=(\|\mathbf{x}_1\|_1,\|\mathbf{x}_2\|_2, \ldots,\|\mathbf{x}_m\|_m)^T$.

证明类似于性质 3.(略)

定义：线性空间 V 的闭凸集Ω若满足,$x\in\Omega$,则$\lambda\cdot x\in\Omega$,其中$|\lambda|\leq1$，那么Ω为均衡闭凸集。

性质 5. (范数与凸集,又称为范数的几何性质)

若$\|\cdot\|$为线性空间 V 上的

向量范数，集合$\Omega=\{x:\|x\|\leq 1\}$为 V 上均衡闭凸集。

反之，若Ω为 V 上的均衡闭凸集，且Ω含有内点，

即包含一个小的单位球。则可以定义函数 $P(x)$

如下：当 $x\neq0$ 时，$P(x)=\min\{\lambda>0:x/\lambda\in\Omega\}$

当 $x=0$ 时，$P(x)=0$.则 $P(x)$为 V 上的范数。

* 证明：1). 显然 $P(x)\geq 0$, 且 $P(0)=0$.

下面我们证明若 $P(x)=0$, 则 $x=0$;

用反证法， 设 $x\neq 0$， 则由 $P(x)$的定义，任给 $\lambda>P(x)=0$, 则有 $x/\lambda\in\Omega$。 因为Ω为有界集。即存在常数 $M>0$ 使得

对任意 $y\in\Omega$, $\|y\|\leq M$. 其中 $\|\cdot\|$为某一给定的范数。

令 $y=x/\lambda$,则得到$\|x/\lambda\|\leq M$,即$\|x\|\leq\lambda\cdot M$,由于λ为任意大于 0 的数，若令$\lambda\to 0$ 则有$\|x\|=0$。因$\|\cdot\|$为范数，

从而 $x=0$. 这样，我们就证明了 1).

2). 若 $x=0$，则 $P(kx)=|k|P(x)$ 显然成立。

假设 $x\neq 0$，由于 $x/P(x)\in\Omega$, 且

任何 $\lambda\geq P(x)$, $x/\lambda\in\Omega$;

而任何 $\lambda< P(x)$, $x/\lambda\notin\Omega$.

显然 $kx/P(kx)\in\Omega$，则 $[(k/|k|)\cdot\{x/[P(kx)/|k|]\}\in\Omega$

注意 $k/|k|$ 的幅度为 1, 从而由 Ω 的均衡性，

我们有 $x/[P(kx)/|k|]\in\Omega$，这样由定义有

$P(x)\leq P(kx)/|k|$，即

$$|k|\cdot P(x)\leq P(kx). \qquad (\heartsuit)$$

同样由于 $x/P(x)\in\Omega$, 注意到 $k/|k|$ 的幅度为 1，

从而 $(kx)/(|k|\,P(x))\in\Omega$, 由定义有

$$P(kx)\leq |k|\cdot P(x) \qquad (\spadesuit)$$

联合($\heartsuit$)和($\spadesuit$)，我们有 $P(kx)=|k|\,P(x)$.

(3). 设 $x\neq 0, y\neq 0$, 则 $x/P(x)\in\Omega$, $y/P(y)\in\Omega$,

令 $\lambda=P(y)/(P(x)+P(y))$, 由于Ω为凸集,

从而

$(x+y)/(P(x)+P(y))=(1-\lambda)\cdot x/P(x)+\lambda\cdot y/(P(y)\in\Omega$,

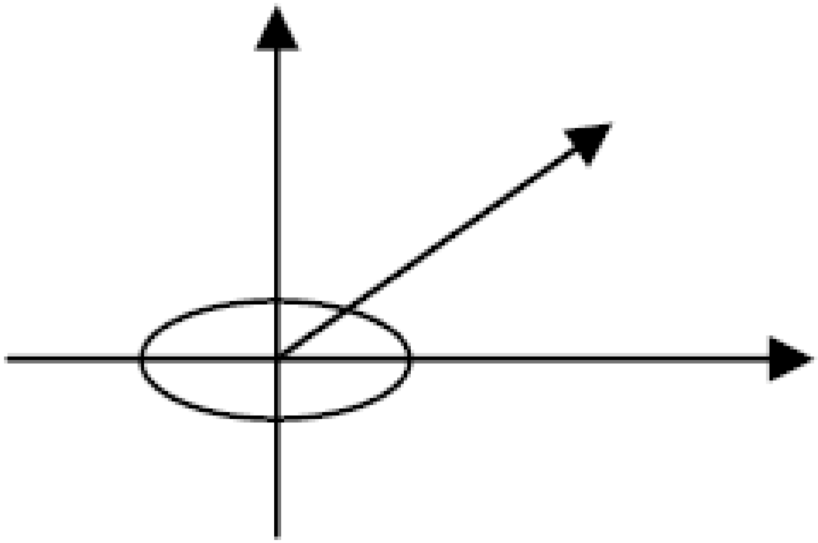
这样有 $P(x+y)$的定义，我们有

$P(x+y)\leq P(x)+P(y)$.

当 x 和 y 有一个或全部为 0 时，显然三角不等式仍然成立。

联合 1), 2)和 3), 从而 $P(x)$为范数。

这个性质说明了范数和均衡凸集之间的一一对应关系。



均衡凸集与范数

2. 向量范数具体实例

例1 在C^n中，向量$x=(x_1,x_2,\ldots,x_n)$的长度

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

是一种范数。一般称为2-范数，记为$\|x\|_2$。

例2 在C^n中，向量$x=(x_1,x_2,\ldots,x_n)$的长度

$$\|x\|=|x_1|+|x_2|+\cdots+|x_n|$$

是一种范数。一般称为1-范数，记为$\|x\|_1$。

例3 在C^n中，向量$x=(x_1,x_2,\ldots,x_n)$的长度

$$\|x\|=\max_i|x_i|$$

是一种范数。一般称为∞-范数，记为$\|x\|_\infty$。

P-范数. **在C^n中，向量$x=(x_1,x_2,\ldots,x_n)$的长度**

$$\|x\|_p=\left(|x_1|^p+|x_2|^p+\cdots+|x_n|^p\right)^{1/p}$$

当p≥1是一种范数。一般称为p-范数，记为$\|x\|_p$。

当p等于2、等于1、趋于无穷时，分别对应2-范数、1-范数、∞-范数。
当0≤p<1时不是严格意义上的向量范数(不满足三角不等式)，但在实际应用中仍然许多重要应用，如p=0时提取的是x的非零分量的个数，在稀疏信号的表示中有很多应用。

引理 2.1 对任意实数$\alpha \geq 0$和$\beta \geq 0$，都有$\alpha\beta \leq \frac{\alpha^p}{p} + \frac{\beta^q}{q}$，其中

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

定理 2.2 对任意$\xi_k, \eta_k \in C \quad (k=1,2,\cdots,n)$，有

$$\sum_{k=1}^{n}|\xi_k||\eta_k| \leq \left(\sum_{k=1}^{n}|\xi_k|^p\right)^{\frac{1}{p}}\left(\sum_{k=1}^{n}|\eta_k|^q\right)^{\frac{1}{q}}$$

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

称式（2.1）为 Holder 不等式。当 p=q=2 时，即得

Cauchy-Schwarz 不等式$(\sum_{k=1}^{n}|\xi_k||\eta_k|)^2 \leq (\sum_{k=1}^{n}|\xi_k|^2)(\sum_{k=1}^{n}|\eta_k|^2)$ （1.5）

（下面证明p 范数的三角不等式）

证 易知非负性和齐次性成立，当p=1时，例2.2中已证明三角不等式成立。下设p>1，则对

$$y=(\eta_1,\eta_2,\cdots\eta_n)^T\in C^n.$$

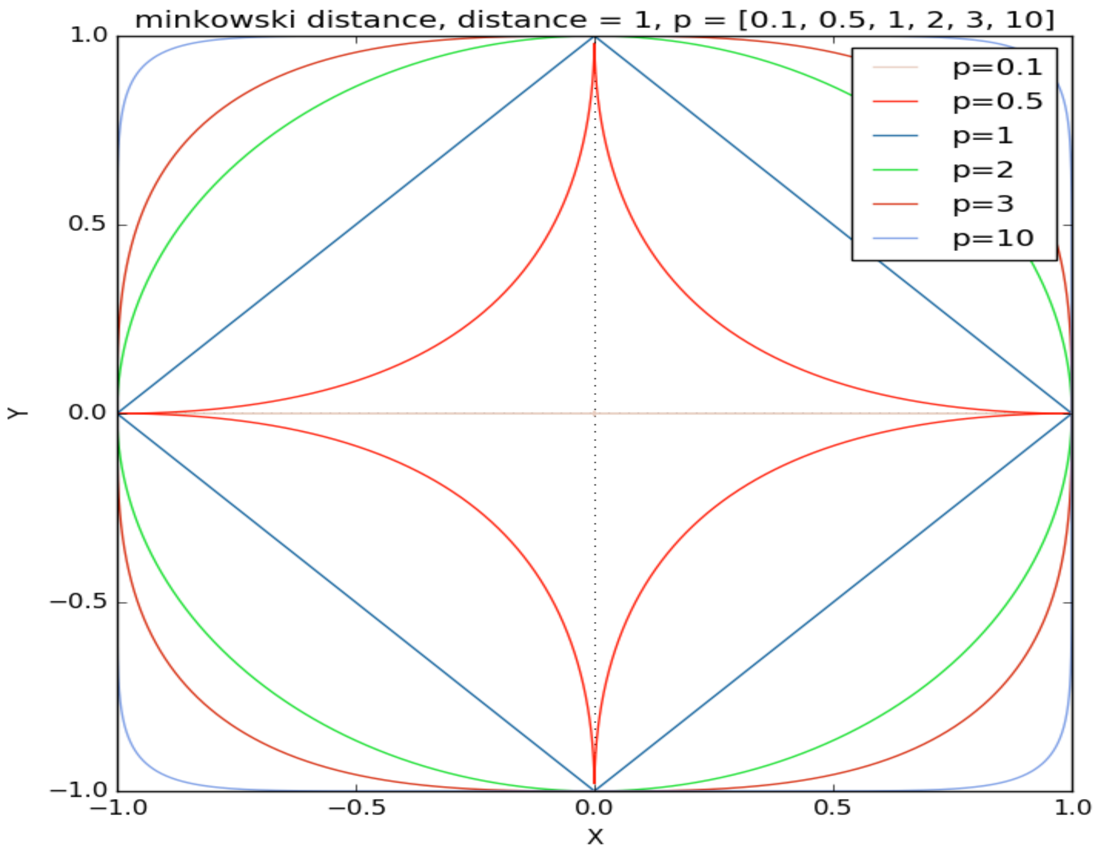
，利用定理2.2得

$$\|x+y\|_p^p=\sum_{k=1}^n|\xi_k+\eta_k|^p\le\sum_{k=1}^n|\xi_k||\xi_k+\eta_k|^{p-1}+\sum_{k=1}^n|\eta_k||\xi_k+\eta_k|^{p-1}\le$$

$$(\sum_{k=1}^n|\xi_k|^p)^{\frac{1}{p}}(\sum_{k=1}^n|\xi_k+\eta_k|^{q(p-1)})^{\frac{1}{q}}+(\sum_{k=1}^n|\eta_k|^p)^{\frac{1}{p}}(\sum_{k=1}^n|\xi_k+\eta_k|^{q(p-1)})^{\frac{1}{q}}$$

$$=(\|x\|_p+\|y\|_p)\|x+y\|_p^{\frac{p}{q}}$$

故$\|x+y\|_p=\|x+y\|_p^{p-\frac{p}{q}}\le\|x\|_p+\|y\|_p$



不同p-范数时当 $\|x\|_p=1$ 时在二维平面构成的图形。注意当p<1时不是凸的范数，在目前的压缩感知的稀疏信号重构中有着重要作用。

例4：设A为n阶实对称正定矩阵，对$x \in Rn$，

定义$\|x\|_A = \left(x^T A x\right)^{1/2}$称为加权范数或椭圆范数

由正定矩阵定义可知$\|x\|_A = 0 \Leftrightarrow x = 0$; $\|x\|_A \neq 0 \Leftrightarrow x \neq 0$

对任意数 $\alpha \in R$，有

$$\|\alpha x\|_A = \sqrt{(\alpha x)^T A \alpha x} = \sqrt{\alpha^2 x^T A x} = |\alpha| \sqrt{x^T A x} = |\alpha| \|x\|_A$$

由 A正定且实对称 $\Rightarrow \exists$ 正交矩阵Q，使得

$$Q^T A Q = diag(\lambda_1, \cdots, \lambda_n), \quad \lambda_i > 0, i = 1, \cdots, n$$

定义 $B = diag(\sqrt{\lambda_1}, \cdots, \sqrt{\lambda_n}) Q^T$ 可得 $A = B^T B$

$$\because \|x\|_A = \left(x^T B^T B x\right)^{1/2} = \left[(Bx)^T (Bx)\right]^{1/2} = \|Bx\|_2$$

$$\therefore \|x+y\|_A = \|B(x+y)\|_2 \le \|Bx\|_2 + \|By\|_2 = \|x\|_A + \|y\|_A$$

例5：设$\|y\|_\alpha$ 是 C^m中的一个向量范数，给定矩阵 $A\in C^{m\times n}$ ，它的n个列向量线性无关。对于C^m 中的一个向量$x=(x_1,x_2,\cdots,x_n)^T$，规定$\|x\|_\beta=\|Ax\|_\alpha$ 则 $\|x\|_\beta$ 也是 C^m 中的一个向量范数。

证：1）设$A=(a_1,a_2,\cdots,a_n)$，由假设知$a_1,a_2,\cdots,a_n$ 线性无关。

当 $x\neq 0$ $\quad Ax=(a_1,\cdots,a_n)\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=a_1x_1+\cdots+a_nx_n\neq 0$

又因为$\|y\|_\alpha$是C^m 中的一个向量范数，有 $\|Ax\|_\alpha>0$

即 $\|x\|_\beta>0$

当 $x=0$ 时，$Ax=0$，所以 $\|x\|_\beta=\|Ax\|_\alpha=0$

2）$\forall k\in C$，$\|kx\|_\beta=\|A(kx)\|_\alpha=\|kAx\|_\alpha=|k|\|Ax\|_\alpha=|k|\|x\|_\beta$

3）$\forall x=\left(x_1,x_2,\cdots,x_n\right)^T, y=\left(y_1,y_2,\cdots,y_n\right)^T\in C^n$ 有

$$\|x+y\|_\beta=\|A(x+y)\|_\alpha=\|Ax+Ay\|_\alpha\le\|Ax\|_\alpha+\|Ay\|_\alpha=\|x\|_\beta+\|y\|_\beta$$

所以，$\|x\|_\beta$ 是 C^n 中的一个向量范数。

由此可知，当给定 $A\in C^{m\times n}$ 时，可以由 C^m 中的一个向量范数确定 C^n 中的一个向量范数。

例6 在$C[a,b]$中，对于函数$f(t)$，

$$\|f(t)\|_p=\left(\int_a^b|f(t)|^p\,dt\right)^{1/p}$$

是一种范数。

当p等于2、等于1、趋于无穷时，分别称为L_2范数、L_1范数、L_∞范数。

例7 给定n维线性空间V的一组基$x_1,x_2,\ldots,x_n$。设x在这组基下的坐标为$y=(y_1,y_2,\ldots,y_n)^T$，则

$$\|x\|_p=\|y\|_p$$

也是一种范数。

3. 范数等价

定义：有限维线性空间 V^n 中任意两个向量范数 $\|x\|_\alpha$ 和 $\|x\|_\beta$，如果存在着正常数 c_1 和 c_2，使得 $c_1\|x\|_\beta \le \|x\|_\alpha \le c_2\|x\|_\beta \quad (\forall x \in V^n)$

则称范数 $\|x\|_\alpha$ 与 $\|x\|_\beta$ 等价

1）自反性：$1\cdot\|x\|_\alpha \le \|x\|_\alpha \le 1\cdot\|x\|_\alpha, \forall x \in V^n$

2）对称性：$\dfrac{1}{c_2}\|x\|_\alpha \le \|x\|_\beta \le \dfrac{1}{c_1}\|x\|_\alpha, \forall x \in V^n$

3）传递性：

$$\left.\begin{array}{l} c_1\|x\|_\beta \le \|x\|_\alpha \le c_2\|x\|_\beta \\ c_3\|x\|_\gamma \le \|x\|_\beta \le c_4\|x\|_\gamma \end{array}\right\} \quad \forall x \in V^n$$

$$\Rightarrow c_5\|x\|_\gamma \le \|x\|_\alpha \le c_6\|x\|_\gamma$$

例　向量空间 V^n 中，对 $\forall x=(\xi_1,\xi_2,\cdots,\xi_n)^T$ ，有

（1）$\|x\|_1=\sum|\xi_i|\le n\bullet\max_i|\xi_i|=n\|x\|_\infty$　$\|x\|_1\ge\sum|\xi_i|=1\bullet\|x\|_\infty$

$\therefore 1\bullet\|x\|_\infty\le\|x\|_1\le n\bullet\|x\|_\infty$

（2）$\|x\|_2=\left(\sum|\xi_i|^2\right)^{\frac{1}{2}}\le\left(n\bullet\max_i|\xi_i|^2\right)^{\frac{1}{2}}=\sqrt{n}\|x\|_\infty$

$\|x\|_2\ge\left(\max_i|\xi_i|^2\right)^{\frac{1}{2}}=1\bullet\|x\|_\infty$　$\therefore 1\bullet\|x\|_\infty\le\|x\|_2\le n\bullet\|x\|_\infty$

（3）$\frac{1}{\sqrt{n}}\bullet\|x\|_2\le\|x\|_1\le n\bullet\|x\|_2$

上面的例子充分说明了不同向量p范数之间是存在等价关系的。我们可以证明任意的两个向量范数等价.

定理：有限维线性空间中任意两个向量范数都等价。

证明思路 1)范数等价为等价关系,满足传递性;

2)任意范数为坐标函数的连续函数;

3)在单位超球面上有大于零的极大极小值,

与2-范数等价。

4. 基于向量范数的收敛性

定义：若$\left\{x^{(k)}\right\}(k=1,2,\cdots)$ 是线性空间V^n 中的向量序列，如果存在 $x\in V^n$ ，使得 $\lim\limits_{k\to+\infty}\left\|x^{(k)}-x\right\|_\alpha=0$

则称序列 $\left\{x^{(k)}\right\}$ 按 $\alpha-$ 范数收敛于 x

定理：向量空间C^n 中，

$$\lim_{k\to+\infty}x^{(k)}=x\Leftrightarrow\forall\|x\|,\lim_{k\to+\infty}\left\|x^{(k)}-x\right\|=0$$

(在证明之前我们需要说明的是：向量序列收敛性一个直观定义就是每个分量构成的数列收敛，这个定理就是证明二者等价）

证明：只需对 $\|x\| = \|x\|_1$ 证明即可。

$$x^k \to x \Leftrightarrow \xi_i^{(k)} \to \xi_i (i = 1, 2, \cdots, n)$$

$$\Leftrightarrow \left|\xi_i^{(k)} - \xi_i\right| \to 0 (i = 1, 2, \cdots, n)$$

$$\Leftrightarrow \sum_{i=1}^{n} \left|\xi_i^{(k)} - \xi_i\right| \to 0$$

$$\Leftrightarrow \left\|x^{(k)} - x\right\|_1 \to 0$$

*5. 机器学习中的距离度量与相似性度量

在做分类时常常需要估算不同样本之间的相似性度量(Similarity Measurement)，这时通常采用的方法就是计算样本间的“距离”(Distance)。采用什么样的方法计算距离是很讲究，甚至关系到分类的正确与否。

欧几里得距离(Euclidean Distance)

欧氏距离是最常见的距离度量，衡量的是多维空间中各个点之间的绝对距离。公式如下：

$$dist(X,Y)=\sqrt{\sum_{i=1}^{n}(x_i-y_i)^2}$$

明可夫斯基距离(Minkowski Distance)

明氏距离是欧氏距离的推广，是对多个距离度量公式的概括性的表述。公式如下：

$$dist(X,Y)=\left(\sum_{i=1}^{n}|x_i-y_i|^p\right)^{1/p}$$

这里的 p 值是一个变量，当 p=2 的时候就得到了上面的欧氏距离。

曼哈顿距离(Manhattan Distance)

曼哈顿距离来源于城市区块距离，是将多个维度上的距离进行求和后的结果，即当上面的明氏距离中 p=1 时得到的距离度量公式，如下：

$$dist(X,Y)=\sum_{i=1}^{n}|x_i-y_i|$$

切比雪夫距离(Chebyshev Distance)

$$dist(X,Y) = \lim_{p \to \infty} \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{1/p} = max|x_i - y_i|$$

马哈拉诺比斯距离(Mahalanobis Distance)

有 M 个样本向量 $X_1 \sim X_m$，协方差矩阵记为 S，均值记为向量 μ，则其中样本向量 X 到 u 的马氏距离表示为

$$D(X) = \sqrt{(X-\mu)^T S^{-1} (X-\mu)}$$

而其中向量 X_i 与 X_j 之间的马氏距离定义为

$$D(X_i - X_j) = \sqrt{(X_i - X_j)^T S^{-1} (X_i - X_j)}$$

若协方差矩阵是单位矩阵（各个样本向量之间独立同分布）,则公式就成了

$$D(X_i - X_j) = \sqrt{(X_i - X_j)^T (X_i - X_j)}$$

也就是欧氏距离了。

若协方差矩阵是对角矩阵，公式变成了标准化欧氏距离。

汉明距离(Hamming distance)

定义：两个等长字符串s_1与s_2之间的汉明距离定义为将其中一个变为另外一个所需要做的最小替换次数。

例如字符串“1111”与“1001”之间的汉明距离为2。

其它相似度度量(不是距离度量)

相似度度量（Similarity），即计算个体间的相似程度，与距离度量相反，相似度度量的值越小，说明个体间相似度越小，差异越大。

向量空间余弦相似度(Cosine Similarity)

余弦相似度用向量空间中两个向量夹角的余弦值作为衡量两个个体间差异的大小。相比距离度量，余弦相似度更加注重两个向量在方向上的差异，而非距离或长度上。公式如下：

$$sim(X,Y) = cos\theta = \frac{\vec{x} \cdot \vec{y}}{||x|| \cdot ||y||}$$

皮尔森相关系数(Pearson Correlation Coefficient)

即相关分析中的相关系数 r，分别对 X 和 Y 基于自身总体标准化后计算空间向量的余弦夹角。公式如下：

$$r(X,Y)=\frac{n\sum xy-\sum x\sum y}{\sqrt{n\sum x^2-(\sum x)^2}\cdot\sqrt{n\sum y^2-(\sum y)^2}}$$

Jaccard 相似系数(Jaccard Coefficient)

Jaccard 系数主要用于计算符号度量或布尔值度量的个体间的相似度，因为个体的特征属性都是由符号度量或者布尔值标识，因此无法衡量差异具体值的大小，只能获得“是否相同”这个结果，所以 Jaccard 系数只关心个体间共同具有的特征是否一致这个问题。如果比较 X 与 Y 的 Jaccard 相似系数，只比较 x_n 和 y_n 中相同的个数，公式如下：

$$Jaccard(X,Y)=\frac{X\cap Y}{X\cup Y}$$

2.2 矩阵范数

1. 矩阵范数的概念
2. 相容范数
3. 从属范数

定义：矩阵空间 $C^{m\times n}$ 中， $\forall A\in C^{m\times n}$ ，

定义实数值$\|A\|$，且满足以下条件

1)正定条件： $\|A\|\geq 0$，且 $\|A\|=0\Leftrightarrow A=0_{m\times n}$

2)齐次条件： $\|kA\|=|k|\cdot\|A\|$， $\forall k\in K$

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

则称$\|A\|$为A的广义范数。

若对于 $C^{m\times n},C^{n\times l}$ 及 $C^{m\times n}$ 上的同类广义矩阵范数有

4)相容条件： $\|AB\|\leq\|A\|\|B\|$， $B\in C^{n\times l}$

则称$\|A\|$为A的**矩阵范数**

例　　已知 $A=(a_{ij})_{n\times n}\in C^{n\times n}$,试证明下面两个函数

$$\|A\|_{m_1}=\sum_{i,j=1}^{n}|a_{ij}|,\quad \|A\|_{m_\infty}=n\cdot\max_{i,j}|a_{ij}|$$

都是 $C^{n\times n}$ 上的矩阵范数.

2. 相容范数

对于$C^{m\times n}$上的矩阵范数$\|A\|_M$和C^m与C^n上的同类向量范数$\|x\|_v$，如果满足

$$\|Ax\|_v \le \|A\|_M \|x\|_v, \quad \forall A \in C^{m\times n}, \forall x \in C^n,$$

则称矩阵范数$\|A\|_M$和向量范数$\|x\|_v$是相容的。

例1、设 $A=\left(a_{ij}\right)_{m\times n}\in C^{m\times n}, x=\left(\xi_1,\cdots,\xi_n\right)^T$，证明

$\|A\|_{m1}=\sum\limits_{i,j=1}^{n}\left|a_{ij}\right|$ 是矩阵函数，且与 $\|x\|_1$ 相容

证明：（1）～（3）显然成立，

$$\|Ax\|_1=\sum_{i=1}^{m}\left|a_{i1}\xi_1+\cdots+a_{in}\xi_n\right|\le\sum_{i=1}^{m}\left(\left|a_{i1}\right|\left|\xi_1\right|+\cdots+\left|a_{in}\right|\left|\xi_n\right|\right)$$

$$\le\sum_{i=1}^{m}\left[\left(\left|a_{i1}\right|+\cdots+\left|a_{in}\right|\right)\left(\left|\xi_1\right|+\cdots+\left|\xi_n\right|\right)\right]$$

$$=\left[\sum_{i=1}^{m}\left(\left|a_{i1}\right|+\cdots+\left|a_{in}\right|\right)\right]\left(\left|\xi_1\right|+\cdots+\left|\xi_n\right|\right)=\|A\|_{m1}\|x\|_1$$

因此，$\|A\|_{m1}$ 与 $\|x\|_1$ 相容。

（4）证明相容性

划分 $B_{n\times l}=(b_1,\cdots b_l)$，则 $AB=(Ab_1,\cdots,Ab_l)$，且有

$$
\begin{aligned}
\|AB\|_{m1} &= \|Ab_1\|_1+\cdots+\|Ab_l\|_1 \\
&\le \|A\|_{m1}\|b_1\|_1+\cdots+\|A\|_{m1}\|b_l\|_1 \\
&= \|A\|_{m1}\left(\|b_1\|_1+\cdots+\|b_l\|_1\right) \\
&= \|A\|_{m1}\|B\|_{m1}
\end{aligned}
$$

因此，$\|A\|_{m1}=\sum_{i,j=1}^{n}|a_{ij}|$ 是矩阵函数，且与 $\|x\|_1$ 相容

例2、设 $A=\left(a_{ij}\right)_{m\times n}\in C^{m\times n}, x=\left(\xi_1,\cdots,\xi_n\right)^T$，证明

$$\|A\|_{m\infty}=n\bullet\max_{i,j}\left|a_{ij}\right|$$ 是矩阵函数，且与 $\|x\|_\infty$相容

证明：（1）～（3）成立，设 $B=\left(b_{ij}\right)_{n\times l}$，则

$$\|AB\|_{m\infty}=l\bullet\max_{i,j}\left|\sum_{k=1}^{n}a_{ik}b_{kj}\right| \le l\bullet\max_{i,j}\left(\sum_{k=1}^{n}\left|a_{ik}\right|\left|b_{kj}\right|\right)$$

$$\le l\bullet n\bullet\max_{i,j}\left(\left|a_{ik}\right|\left|b_{kj}\right|\right) \le \left(n\bullet\max_{i,j}\left|a_{ij}\right|\right)\bullet\left(l\bullet\max_{i,j}\left|b_{ij}\right|\right)=\|A\|_{m\infty}\|B$$

$$\|Ax\|_\infty=\max_i\left|\sum_{k=1}^{n}a_{ik}\xi_k\right| \le \max_i\sum_{k=1}^{n}\left|a_{ik}\right|\left|\xi_k\right|$$

$$\le \max_i\sum_{k=1}^{n}\left|a_{ik}\right|\bullet\max_i\left|\xi_i\right| \le \left(n\bullet\max_i\left|a_{ik}\right|\right)\bullet\max_i\left|\xi_i\right|=\|A\|_{m\infty}\|x$$

例3、设 $A=\left(a_{ij}\right)_{m\times n}\in C^{m\times n}, x=\left(\xi_1,\cdots,\xi_n\right)^T$，证明

$$\|A\|_{m2}=\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\left|a_{ij}\right|^2\right)^{\frac{1}{2}}$$ 是矩阵函数，且与 $\|x\|_2$ 相容

证明：（1）～（2）成立，$B=\left(b_{ij}\right)_{n\times l}$

设 $B_{m\times n}$，划分 $A=\left(a_1,\cdots,a_n\right), B=\left(b_1,\cdots,b_n\right)$，则有

$$\|A+B\|_{m2}^2=\|a_1+b_1\|_2^2+\cdots+\|a_n+b_n\|_2^2$$

$$\le\left(\|a_1\|_2+\|b_1\|_2\right)^2+\cdots+\left(\|a_n\|_2+\|b_n\|_2\right)^2$$

$$\le\|A\|_{m2}^2+2\left(\|a_1\|_2\|b_1\|_2+\cdots+\|a_n\|_2\|b_n\|_2\right)+\|B\|_{m2}^2$$

$$\le\|A\|_{m2}^2+2\left(\sum\|a_i\|_2^2\right)^{\frac{1}{2}}\left(\sum\|b_i\|_2^2\right)^{\frac{1}{2}}+\|B\|_{m2}^2=\left(\|A\|_{m2}+\|B\|_{m2}\right)$$

设 $B_{n\times l}$， $AB=\left(\sum_{k=1}^{n}a_{ik}b_{kj}\right)_{m\times l}$ ，则有

$$\|AB\|_{m2}=\sum_{i,j=1}^{n}\left|\sum_{k=1}^{n}a_{ik}b_{kj}\right|^2 \le \sum_{i,j=1}^{n}\left(\sum_{k=1}^{n}|a_{ik}||b_{kj}|\right)^2$$

$$\le \sum_{i,j=1}^{n}\left[\left(\sum_{k=1}^{n}|a_{ik}|^2\right)\bullet\left(\sum_{k=1}^{n}|b_{kj}|^2\right)\right] \le \sum_{i=1}^{n}\left[\left(\sum_{k=1}^{n}|a_{ik}|^2\right)\bullet\sum_{j=1}^{n}\left(\sum_{k=1}^{n}|b_{kj}|^2\right)\right]$$

$$\le \left(\sum_{i,k=1}^{n}|a_{ik}|^2\right)\bullet\left(\sum_{k,j=1}^{n}|b_{kj}|^2\right) = \|A\|_{m2}\bullet\|B\|_{m2}$$

特别的，取 $B=x\in C^{n\times 1}$ ，则有

$$\|Ax\|_2=\|AB\|_{m2} \le \|A\|_{m2}\bullet\|B\|_{m2} = \|A\|_{m2}\bullet\|x\|_2$$

注： 设$A=(a_{ij})\in C^{m\times n}$，则

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

是矩阵范数，且与向量的2-范数相容。

该范数称为**Frobenius范数**，或简称为***F*-范数**。

定理3：设 $A_{m\times n}$ 及酉矩阵 $P_{m\times m}$ 和 $Q_{n\times n}$，有

$$\|PA\|_F=\|A\|_F=\|AQ\|_F$$

证明：

$$\|PA\|_F^2=\mathrm{tr}\left((PA)^H(PA)\right)=\mathrm{tr}\left(A^HP^HPA\right)$$

$$=\mathrm{tr}\left(A^HA\right)$$

$$\|AQ\|_F^2=\mathrm{tr}\left((AQ)^H(AQ)\right)=\mathrm{tr}\left(Q^HA^HAQ\right)$$

$$=\mathrm{tr}\left(AQQ^HA^H\right)=\mathrm{tr}\left(A^HA\right)$$

推论：酉（正交）相似的矩阵的F-范数相等

引理：对 $C^{m\times n}$ 的矩阵范数 $\|A\|$，存在向量范数 $\|x\|_V$

使得 $\|Ax\|_V \le \|A\|\bullet\|x\|_V$

例7：对$C^{m\times n}$的矩阵范数$\|A\|$,任取非零列向量 $y\in C^n$

（1） $\|x\|_V = \|xy^{\mathrm{H}}\|$ 是 C^n 的向量范数；

（2） $\|A\|$ 与 $\|x\|_V$ 相容。

证明：（1）略；（2）

$$\|Ax\|_V = \|(Ax)y^H\| = \|A(xy^H)\| \le \|A\|\bullet\|xy^H\| = \|A\|\bullet\|x\|_V$$

3. 从属范数

定理1：已知C^m与C^n上的同类向量范数$\|x\|_v$，设$A\in C^{m\times n}$，则按如下方式定义的的函数

$$\|A\|_M=\max_{\|x\|_v=1}\|Ax\|_v=\max_{x\neq 0}\frac{\|Ax\|_v}{\|x\|_v}$$

是矩阵范数，并且与已知的向量范数$\|x\|_v$相容。

定理1中给出的矩阵范数称为由向量范数诱导出的矩阵范数，简称为从属范数。

对于任意从属范数，$\|I\|=1$，

但对于一般的矩阵范数，只有$\|I\|\geq 1$。

三、从属范数

定理：对 C^m 与 C^n 上的同类向量范数 $\|x\|_V$，定义

$$\|A\| = \max_{\|x\|_V=1} \|Ax\|_V \qquad \left(\forall A_{m\times n}, x \in C^n\right)$$

则 $\|A\|$ 是 $C^{m\times n}$ 中矩阵A 的范数，且 $\|A\|$ 与 $\|x\|_V$ 相容

$\|A\|$称为由 $\|x\|_V$导出的矩阵范数（或称为从属范数）

等价定义：$\displaystyle \max_{\|x\|_V=1} \|Ax\|_V = \max_{x\neq\theta} \frac{\|Ax\|_V}{\|x\|_V}$

证明:(1) $A\neq 0:\exists x_0$ 满足 $\|x_0\|_V=1,\text{st } Ax_0\neq\theta$ 从而

$$\|A\|\geq\|Ax_0\|_V>0$$

$$A=0:\|A\|=\max_{\|x\|_V=1}\|Ax\|_V=\max_{\|x\|_V=1}\|0\|_V=0$$

(2) 略

(3) 对 $A+B:\exists x_1,\|x_1\|_V=1,\text{st}\max_{\|x\|_V=1}\|(A+B)x_1\|_V=\|(A+B)x_1\|_V$

$$\|A+B\|=\|Ax_1+Bx_1\|_V\leq\|Ax_1\|_V+\|Bx_1\|_V\leq\|A\|+\|B\|$$

(4) 先证 $\|Ay\|_V\leq\|A\|\|y\|_V\quad\left(y\in C^n\right)$

$y=\theta$ ： 显然成立。

$y \neq \theta$：定义 $y_0 = \dfrac{y}{\|y\|_V}$ 满足

$$\|y_0\|_V = 1 \Rightarrow \|Ay_0\|_V \leq \max_{\|y\|_V=1} \|Ay\|_V = \|A\|$$

故 $\|Ay\|_V = \left\|A\left(\|y\|_V\, y_0\right)\right\|_V = \|Ay_0\|_V\, \|y\|_V \leq \|A\| \bullet \|y\|_V$

对 $AB: \exists x_2, \|x_2\|_V = 1, st \max\limits_{\|x\|_V=1} \|(A+B)x\|_V = \|(A+B)x_2\|_V$

$$\|AB\| = \|(AB)x_2\|_V = \|A(Bx_2)\|_V \leq \|A\| \bullet \|Bx_2\|_V \leq \|A\| \bullet \|B\|$$

故定理成立。

注：

（1）一般的矩阵范数：$\because I = I \bullet I$

$$\|I\| \le \|I\| \bullet \|I\| \qquad \therefore \|I\| \ge 1$$

例如：$\|I\|_{m1} = n, \quad \|I\|_F = \sqrt{n}$

（2）矩阵的从属范数：$\|I\| = \max\limits_{\|x\|_V = 1} \|Ix\|_V = 1$

（3）常用的从属范数：

$$\|x\|_V \qquad \|x\|_1 \qquad \|x\|_2 \qquad \|x\|_\infty$$

$$\|A\|_M \qquad \|A\|_1 \qquad \|A\|_2 \qquad \|A\|_\infty$$

定理2：由向量的1-范数、2-范数和∞-范数分别诱导出的矩阵范数分别是

$$(1)\ \|A\|_1=\max_j\sum_{i=1}^{m}|a_{ij}|;$$

$$(2)\ \|A\|_2=\max_j\sqrt{\lambda_j\left(A^*A\right)};$$

$$(3)\ \|A\|_\infty=\max_i\sum_{j=1}^{n}|a_{ij}|.$$

通常依次称为**列和范数**、**谱范数**和**行和范数**。

先证列和范数： $\|A\|_1 = \max_j \left\{ \sum_{i=1}^{m} |a_{ij}| \right\}$

证明： （1）记 $t = \max_j \left\{ \sum_{i=1}^{m} |a_{ij}| \right\}$，

如果 $x \in C^n$ 满足 $\|x\|_1 = 1$ ，则

$$\|Ax\|_1 = \sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij}\xi_j \right| \le \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}| \cdot |\xi_j| = \sum_{j=1}^{n} \left[|\xi_j| \cdot \left(\sum_{i=1}^{m} |a_{ij}| \right) \right]$$

$$\le \sum_{j=1}^{n} \left\lfloor |\xi_j| t \right\rfloor \le t \cdot \|x\|_1 = t \qquad \therefore \|A\|_1 \le t$$

选k使得 $t=\max\limits_j\left\{\sum\limits_{i=1}^m |a_{ij}|\right\}=\sum\limits_{i=1}^m |a_{ik}|$ ，令

$e_k=(0,\cdots,0,1,0,\cdots,0)^T$ 则 $\|e_k\|_1=1$ ，而且

$$\|A\|_1\ge\|Ae_k\|_1=\sum_{i=1}^m |a_{ik}| =\max_j\left\{\sum_{i=1}^m |a_{ij}|\right\}=t$$

所以 $\|A\|_1=\max\limits_j\left\{\sum\limits_{i=1}^m |a_{ij}|\right\}$

其余谱范数和行和范数类似证明。

定理3：谱范数和F-范数都是酉不变范数，
即对于任意酉矩阵P和Q，有$\|PAQ\|=\|A\|$。

证明略.

2.3 范数的一些应用

1. 逆矩阵的扰动分析
2. 谱半径及其性质

定理6：设 $A\in C^{n\times n}$,且对 $C^{n\times n}$ 的矩阵范数 $\|\bullet\|$ ，满足 $\|A\|<1$,则矩阵 $I-A$ 非奇异，且有

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

证明：选取向量范数 $\|x\|_v$ ，使得 $\|A\|$ 与 $\|x\|_v$ 相容。

如果 $\det(I-A)=0$,则 $(I-A)x=0$ 有非零解 x_0

$$x_0=Ax_0\Rightarrow\|x_0\|_V=\|Ax_0\|_V\le\|A\|\bullet\|x_0\|_V<\|x_0\|_V$$

产生了矛盾，故 $\det(I-A)\neq 0$ ， $I-A$ 可逆。

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

$$\Rightarrow (I-A)^{-1}=I+(I-A)^{-1}A$$

$$\Rightarrow \left\|(I-A)^{-1}\right\| \le \|I\|+\left\|(I-A)^{-1}A\right\|$$

$$\Rightarrow \left\|(I-A)^{-1}\right\| \le \|I\|+\left\|(I-A)^{-1}\right\|\|A\|$$

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

定理7：设 $A\in C^{n\times n}$,且对$C^{n\times n}$的矩阵范数 $\|\bullet\|$ ，满足 $\|A\|<1$,则矩阵$I-A$非奇异，且有

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

证明：已证 $\|A\|<1$，则 $I-A$可逆。

恒等式： $$(I-A)-I=-A$$

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

左乘A： $A-A(I-A)^{-1}=-A^2(I-A)^{-1}$

左乘 A： $A-A(I-A)^{-1}=-A^2(I-A)^{-1}$

$$\Rightarrow \left\|A(I-A)^{-1}\right\| \le \|A\|+\|A\|\bullet\left\|A(I-A)^{-1}\right\|$$

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

$$\Rightarrow \left\|I-(I-A)^{-1}\right\| = \left\|-A(I-A)^{-1}\right\| \le \frac{\|A\|}{1-\|A\|}$$

定理8：设$A\in C^{n\times n}$可逆，$B\in C^{n\times n}$，且对某种矩阵范数有$\|A^{-1}B\|<1$，则有

(1) $A+B$非奇异；

(2) 设$F=I-(I+A^{-1}B)^{-1}$，则 $\|F\|\le \dfrac{\|A^{-1}B\|}{1-\|A^{-1}B\|}$；

(3) $$\frac{\|A^{-1}-(A+B)^{-1}\|}{\|A^{-1}\|}\le \frac{\|A^{-1}B\|}{1-\|A^{-1}B\|}.$$

称cond(A)=$\|A\|\ \|A^{-1}\|$为矩阵A的**条件数(condition number)**，它反映了A^{-1}对扰动的敏感程度。

证明：利用定理6和定理7可证明

矩阵范数在线性方程组求解中的误差分析

设线性方程组 Ax=b 的扰动解x+δx满足方程为

$$(A+\delta A)(x+\delta x)=(b+\delta b)$$

那么成立

$$\frac{\|\delta x\|}{\|x\|}\le cond(A)\left(\frac{\|\delta A\|}{\|A\|}+\frac{\|\delta b\|}{\|b\|}\right)$$

证明：由题设可得

$$Ax + A\cdot\delta x + \delta A\cdot x + \delta A\cdot\delta x = b + \delta b$$

即

$$A\cdot\delta x + \delta A\cdot x \approx \delta b$$

从而

$$\delta x \approx A^{-1}\delta b - A^{-1}\delta A\cdot x$$

两边取范数并放大后可得

$$\|\delta x\| \le \|A^{-1}\|\cdot\|\delta b\| + \|A^{-1}\|\cdot\|\delta A\|\cdot\|x\|$$

再根据$\|A\|\cdot\|x\|\ge\|b\|$可得

$$\frac{\|\delta x\|}{\|x\|} \le \frac{\|A^{-1}\|\cdot\|\delta b\|}{\|x\|} + \|A^{-1}\|\cdot\|\delta A\|$$

$$\le (\|A^{-1}\|\cdot\|A\|)\frac{\|\delta b\|}{\|b\|} + (\|A^{-1}\|\cdot\|A\|)\frac{\|\delta A\|}{\|A\|}$$

2. 谱半径及其性质

设$A\in C^{n\times n}$的n个特征值为$\lambda_1,\lambda_2,\ldots,\lambda_n$，称

$$\rho(A)=\max_i|\lambda_i|$$

为矩阵A的**谱半径(spectral radius)**。

例1 $\rho(A^k)=[\rho(A)]^k$。

例2 矩阵A的2-范数为

$$\|A\|_2=\left(\rho(A^*A)\right)^{\frac{1}{2}}=\left(\rho(AA^*)\right)^{\frac{1}{2}},$$

若A是Hermite矩阵，则

$$\|A\|_2=\rho(A).$$

定理9：设$A \in C^{n \times n}$，则对任意一种矩阵范数都有

$$\rho(A) \le \|A\|.$$

证明：对矩阵范数构造相应的相容向量范数

$\|\cdot\|_{V_n}$，从而有

设λ为 A 的任意特征值，x 为相应特征向量，

则有 $|\lambda| \cdot \|x\|_{V_n} = \|\lambda x\|_{V_n} = \|Ax\|_{V_n}$

$$\le \|A\| \cdot \|x\|_{V_n}.$$

从而有 $|\lambda| \le \|A\|$，由题设

$\rho(A) \le \|A\|$.

定理10：设$A\in C^{n\times n}$，则对任意正数ε，必存在某种矩阵范数，使得

$$\|A\|\le\rho(A)+\varepsilon.$$

上面两个定理揭示了**谱半径与范数之间的关系**。

证明：根据定理 1.29，存在可逆矩阵

$P\in C^{n\times n}$ 使得

$P^{-1}AP=\Lambda+\tilde{I}$

$$\tilde{I}=\begin{pmatrix} 0 & \delta_1 & & & \\ & 0 & \delta_2 & & \\ & & \ddots & \ddots & \\ & & & 0 & \delta_{n-1} \\ & & & & 0 \end{pmatrix}$$

其中δ_i等于 0 或 1。 记对角元矩阵为Λ，即

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

令 $D=\mathrm{diag}(1,\varepsilon,\ldots,\varepsilon^{n-1})$,$S=PD$,那么

$$S^{-1}AS=\Lambda+\varepsilon\tilde{I}$$

显然有$\|S^{-1}AS\|_1=\|\Lambda+\varepsilon\tilde{I}\|_1\le\rho(A)+\varepsilon$

从而定义矩阵 $X\in C^{n\times n}$ 的范数为

$$\|X\|_M=\|S^{-1}XS\|_1$$

从而有

$$\|A\|_M=\|S^{-1}AS\|_1\le\rho(A)+\varepsilon$$

证毕.

注意，我们给出的范数定义和矩阵 A 有关。实际上不存在这样的矩阵范数对任意矩阵都成立。

矩阵 **A** 的矩阵范数与谱: