向量范数和矩阵范数

—— *Roger A. Horn, Charles R. Johnson 《Matrix Analysis》*

谭志豪

2018.11.7

向量范数

Definition 5.1.1. *Let V be a vector space over the field* $\mathbf{F}$ ($\mathbf{F} = \mathbf{R}$ *or* $\mathbf{C}$). *A function* $\|\cdot\| : V \to \mathbf{R}$ *is a* norm *(sometimes one says* vector norm*) if, for all $x, y \in V$ and all $c \in \mathbf{F}$,*

(1)	$\|x\| \geq 0$	*Nonnegativity*
(1a)	$\|x\| = 0$ *if and only if* $x = 0$	*Positivity*
(2)	$\|cx\| = \|c\|\|x\|$	*Homogeneity*
(3)	$\|x + y\| \leq \|x\| + \|y\|$	*Triangle Inequality*

A function $\|\cdot\| : V \to \mathbf{R}$ that satisfies axioms (1), (2), and (3) of (5.1.1) is called a *seminorm*. The seminorm of a nonzero vector can be zero.

由内积诱导的范数

Definition 5.1.3. *Let V be a vector space over the field* $\mathbf{F}$ ($\mathbf{F} = \mathbf{R}$ *or* $\mathbf{C}$). *A function* $\langle\cdot,\cdot\rangle : V \times V \to \mathbf{F}$ *is an inner product if for all* $x, y, z \in V$ *and all* $c \in \mathbf{F}$,

(1)	$\langle x, x\rangle \geq 0$	*Nonnegativity*
(1a)	$\langle x, x\rangle = 0$ *if and only if* $x = 0$	*Positivity*
(2)	$\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$	*Additivity*
(3)	$\langle cx, y\rangle = c\langle x, y\rangle$	*Homogeneity*
(4)	$\langle x, y\rangle = \overline{\langle y, x\rangle}$	*Hermitian Property*

Learning And Mining from DatA

由内积诱导的范数

Theorem 5.1.4 (Cauchy–Schwarz inequality). *Let* $\langle \cdot, \cdot \rangle$ *be an inner product on a vector space* V *over the field* $\mathbf{F}$ ($\mathbf{F} = \mathbf{R}$ *or* $\mathbf{C}$). *Then*

$$|\langle x, y\rangle|^2 \le \langle x, x\rangle \langle y, y\rangle \quad \textit{for all} \quad x, y \in V \tag{5.1.5}$$

with equality if and only if x *and* y *are linearly dependent, that is, if and only if* $x = \alpha y$ *or* $y = \alpha x$ *for some* $\alpha \in \mathbf{F}$.

$$v = \langle y, y\rangle x - \langle x, y\rangle y$$

$$\begin{aligned} 0 \le \langle v, v\rangle &= \langle \langle y, y\rangle x - \langle x, y\rangle y, \langle y, y\rangle x - \langle x, y\rangle y\rangle \\ &= \langle y, y\rangle^2 \langle x, x\rangle - \langle y, y\rangle \overline{\langle x, y\rangle} \langle x, y\rangle - \langle x, y\rangle \langle y, x\rangle \langle y, y\rangle + \langle y, y\rangle \overline{\langle x, y\rangle} \langle x, y\rangle \\ &= \langle y, y\rangle^2 \langle x, x\rangle - \langle y, y\rangle |\langle x, y\rangle|^2 \\ &= \langle y, y\rangle \left(\langle x, x\rangle \langle y, y\rangle - |\langle x, y\rangle|^2 \right) \end{aligned} \tag{5.1.6}$$

由内积诱导的范数

内积诱导的范数

Corollary 5.1.7. *If $\langle \cdot, \cdot \rangle$ is an inner product on a real or complex vector space V, then the function $\|\cdot\| : V \to [0, \infty)$ defined by $\|x\| = \langle x, x \rangle^{1/2}$ is a norm on V.*

内积诱导的范数一定满足平行四边形恒等式

$$\frac{1}{2}(\|x + y\|^2 + \|x - y\|^2) = \|x\|^2 + \|y\|^2$$

满足平行四边形恒等式的范数一定可由内积诱导

$$\langle x, y \rangle = \frac{1}{2}(\|x + y\|^2 - \|x\|^2 - \|y\|^2)$$

LAMDA Learning And Mining from DatA

常见的有限维向量范数

- l_2 范数 $\|x\|_2 = (|x_1|^2 + \cdots + |x_n|^2)^{1/2}$
  - Euclid内积诱导
  - 酉不变 $\|Ux\|_2 = \|x\|_2$
- l_1 范数 $\|x\|_1 = |x_1| + \cdots + |x_n|$
- l_∞ 范数 $\|x\|_\infty = \max\{|x_1|, \ldots, |x_n|\}$
- l_p 范数 $\|x\|_p = (|x_1|^p + \cdots + |x_n|^p)^{1/p}, \quad p \geq 1$
  - Minkowski不等式

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p$$

$$\begin{aligned}
\|f+g\|_p^p &= \int |f+g|^p \, \mathrm{d}\mu \\
&= \int |f+g| \cdot |f+g|^{p-1} \, \mathrm{d}\mu \\
&\leq \int (|f|+|g|)|f+g|^{p-1} \, \mathrm{d}\mu \\
&= \int |f||f+g|^{p-1} \, \mathrm{d}\mu + \int |g||f+g|^{p-1} \, \mathrm{d}\mu \\
&\leq \left(\left(\int |f|^p \, \mathrm{d}\mu \right)^{\frac{1}{p}} + \left(\int |g|^p \, \mathrm{d}\mu \right)^{\frac{1}{p}} \right) \left(\int |f+g|^{(p-1)\left(\frac{p}{p-1}\right)} \, \mathrm{d}\mu \right)^{1-\frac{1}{p}} \\
&= (\|f\|_p + \|g\|_p) \frac{\|f+g\|_p^p}{\|f+g\|_p}
\end{aligned}$$

LAMDA Learning And Mining from DatA

常见的无限维向量范数

- 无限维向量空间
  - 区间 $[a,b]$ 上所有连续的实值 (复值)函数

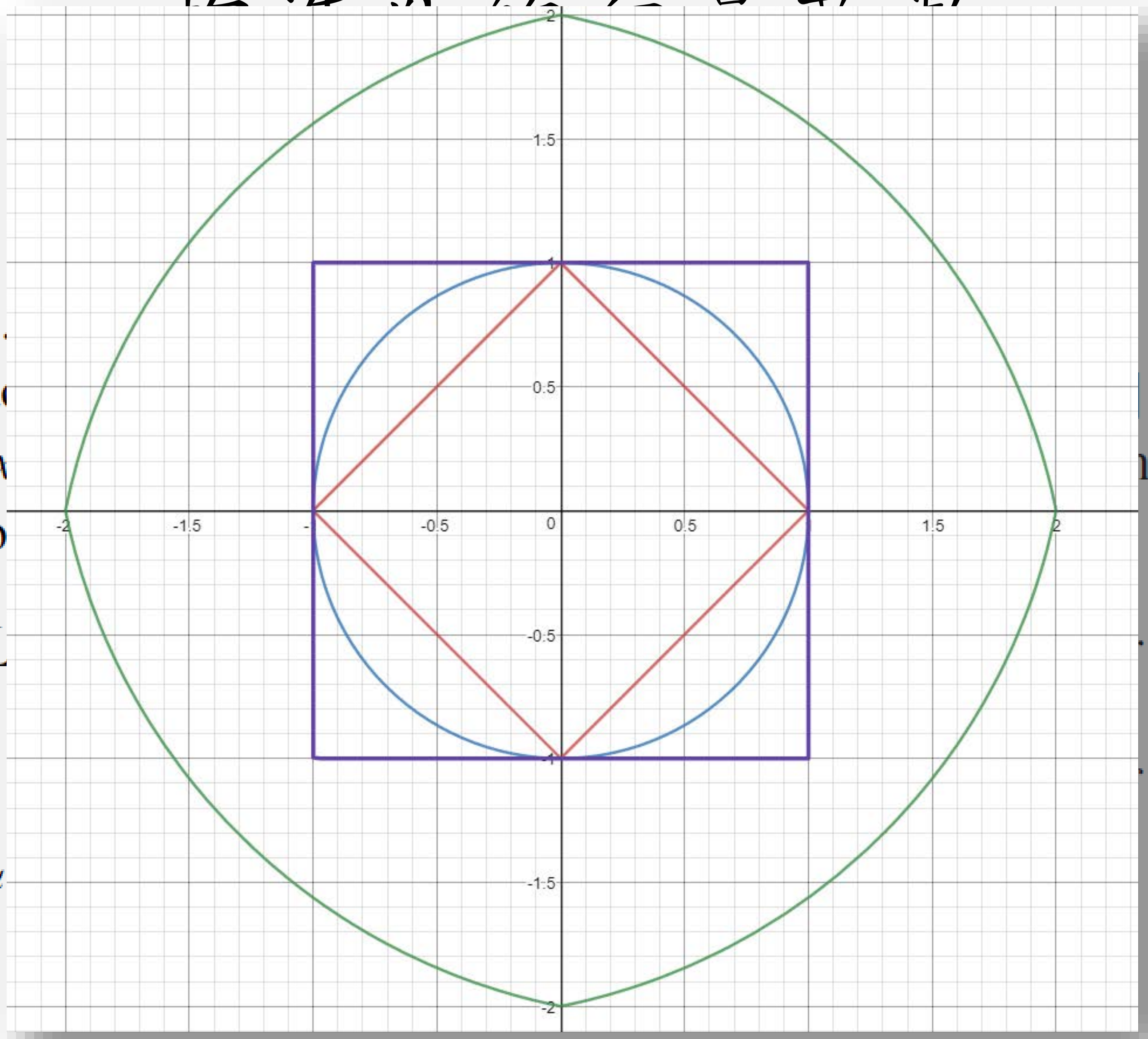
$$\|f\|_2 = \left[\int_a^b |f(t)|^2 dt\right]^{1/2}$$

$$\|f\|_1 = \int_a^b |f(t)| dt$$

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

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

$\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$



5.3.P2 Let $m = 2$ ... $\mathbf{R}^2$ that does not satisfy the monot ... $[\|x\|_\infty, \|x\|_1]^T\| =$ $\min\{|x_1|, |x_2|\} + |x$ ... homogeneity axiom for a norm on $\mathbf{R}^2$ b

Theorem 5.3.1. *L ... space V over the field* $\mathbf{F}$ *(*$\mathbf{F} = \mathbf{R}$ *or ...* $\leq \|y + z\|$ *for all vectors* $y, z \in \mathbf{R}^m$ *...* $: V \to \mathbf{R}$ *defined by* $f(x) = \|[\|x\|_\alpha$

Learning And Mining from DatA

构造新的向量范数

Let $S \in M_{m,n}$ have full column rank, so $m \geq n$. Let $\|\cdot\|$ be a given norm on $\mathbf{C}^m$ and define

$$\|x\|_S = \|Sx\| \tag{5.2.6}$$

for $x \in \mathbf{C}^n$. Then $\|\cdot\|_S$ is a norm on $\mathbf{C}^n$.

向量范数的种类

由内积诱导的范数

- Cauchy-Swartz不等式
- 平行四边形恒等式
- l_2范数酉不变、由内积诱导

常见的向量范数

- Minkowski不等式
  - Hölder不等式
- 无限维的范数

构造新范数

- 范数的单调性

Learning And Mining from DatA

向量范数的分析性质

- 向量范数可以用来度量向量序列的收敛性

Definition 5.4.1. *Let V be a real or complex vector space with a given norm $\|\cdot\|$. We say that a sequence $\{x^{(k)}\}$ of vectors in V* converges *to a vector $x \in V$ with respect to $\|\cdot\|$ if and only if $\lim_{k\to\infty}\|x^{(k)} - x\| = 0$. If $\{x^{(k)}\}$ converges to x with respect to $\|\cdot\|$, we write $\lim_{k\to\infty}x^{(k)} = x$ with respect to $\|\cdot\|$.*

- 是否存在可能对一种范数收敛而对另一种范数不收敛?

无限维情况

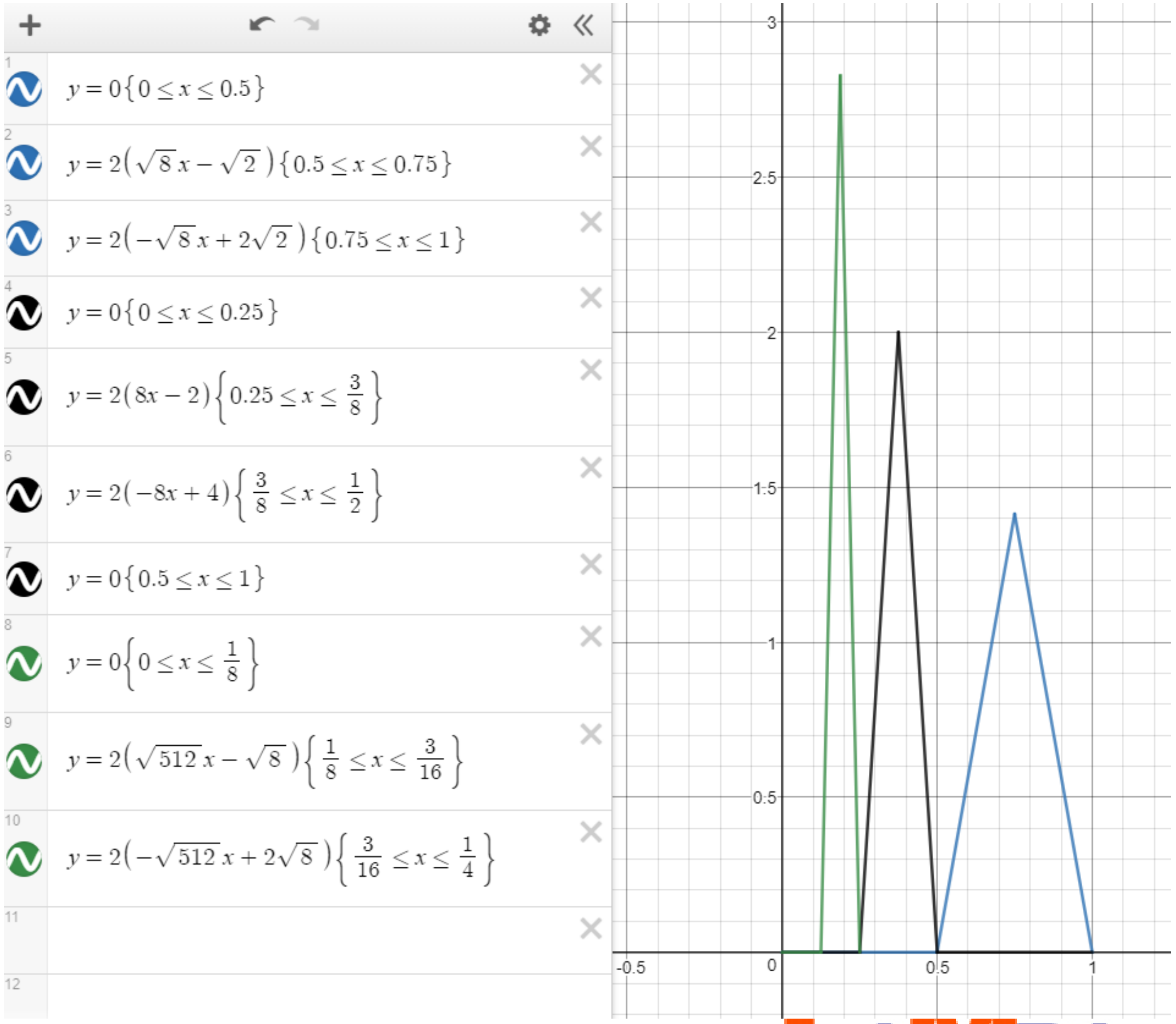
the sequence $\{f_k\}$ of functions in $C[0, 1]$

$$
\begin{aligned}
f_k(x) &= 0, & 0 \le x \le \tfrac{1}{k} \\
f_k(x) &= 2(k^{3/2}x - k^{1/2}), & \tfrac{1}{k} \le x \le \tfrac{3}{2k} \\
f_k(x) &= 2(-k^{3/2}x + 2k^{1/2}), & \tfrac{3}{2k} \le x \le \tfrac{2}{k} \\
f_k(x) &= 0, & \tfrac{2}{k} \le x \le 1
\end{aligned}
$$

$$\|f_k\|_1 = \frac{1}{2}k^{-1/2} \to 0 \text{ as } k \to \infty$$

$$\|f_k\|_2 = \frac{1}{\sqrt{3}} \text{ for all } k = 1, 2, \ldots$$

$$\|f_k\|_\infty = k^{1/2} \to \infty \text{ as } k \to \infty$$



Learning And Mining from DatA

有限维情况

- 范数都是一致连续的

Lemma 5.4.3. *Let* $\|\cdot\|$ *be a norm on a vector space* V *over the field* $\mathbf{F}$ ($\mathbf{F}=\mathbf{R}$ *or* $\mathbf{C}$), *let* $m \geq 1$ *be a given positive integer, let* $x^{(1)}, x^{(2)}, \ldots, x^{(m)} \in V$ *be given vectors, and define* $x(z) = z_1 x^{(1)} + z_2 x^{(2)} + \cdots + z_m x^{(m)}$ *for any* $z = [z_1 \ \ldots \ z_m]^T \in \mathbf{F}^m$. *The function* $g: \mathbf{F}^m \to \mathbf{R}$ *defined by*

$$g(z) = \|x(z)\| = \|z_1 x^{(1)} + z_2 x^{(2)} + \cdots + z_m x^{(m)}\|$$

is a uniformly continuous function on $\mathbf{F}^m$ *with respect to the Euclidean norm.*

$$\begin{aligned} |g(x(u)) - g(x(v))| &= |\|x(u)\| - \|x(v)\|| \leq \|x(u) - x(v)\| \\ &= \left\| \sum_{i=1}^{m} (u_i - v_i) x^{(i)} \right\| \leq \sum_{i=1}^{m} |u_i - v_i| \|x^{(i)}\| \\ &\leq \left(\sum_{i=1}^{m} |u_i - v_i|^2 \right)^{1/2} \left(\sum_{i=1}^{m} \|x^{(i)}\|^2 \right)^{1/2} = C \|u - v\|_2 \end{aligned}$$

有限维情况

Theorem 5.4.4. *Let f_1 and f_2 be real-valued functions on a finite-dimensional vector space V over the field* $\mathbf{F}$ ($\mathbf{F} = \mathbf{R}$ *or* $\mathbf{C}$), *let $\mathcal{B} = \{x^{(1)}, \ldots, x^{(n)}\}$ be a basis for V, and let $x(z) = z_1 x^{(1)} + \cdots + z_n x^{(n)}$ for all $z = [z_1 \ \ldots \ z_n]^T \in \mathbf{F}^n$. Assume that f_1 and f_2 are*

准范数

(a) *Positive: $f_i(x) \geq 0$ for all $x \in V$, and $f_i(x) = 0$ if and only if $x = 0$*
(b) *Homogeneous: $f_i(\alpha x) = |\alpha| f_i(x)$ for all $\alpha \in \mathbf{F}$ and all $x \in V$*
(c) *Continuous: $f_i(x(z))$ is continuous on $\mathbf{F}^n$ with respect to the Euclidean norm*

Then there exist finite positive constants C_m and C_M such that

$$C_m f_1(x) \leq f_2(x) \leq C_M f_1(x) \text{ for all } x \in V$$

Define $h(z) = f_2(x(z))/f_1(x(z))$
$\|z\|_2 = 1$

(a) (c)
→
The Weierstrass theorem

$C_m f_1(x(z)) \leq f_2(x(z)) \leq C_M f_1(x(z))$

有限维向量范数的收敛性

Corollary 5.4.5. *Let* $\|\cdot\|_\alpha$ *and* $\|\cdot\|_\beta$ *be given norms on a finite-dimensional real or complex vector space* V. *Then there exist finite positive constants* C_m *and* C_M *such that* $C_m\|x\|_\alpha \le \|x\|_\beta \le C_M\|x\|_\alpha$ *for all* $x \in V$.

有限维实或复向量空间的序列关于不同范数收敛到同一点

Corollary 5.4.6. *If* $\|\cdot\|_\alpha$ *and* $\|\cdot\|_\beta$ *are norms on a finite-dimensional real or complex vector space* V, *and if* $\{x^{(k)}\}$ *is a given sequence of vectors in* V, *then* $\lim_{k\to\infty} x^{(k)} = x$ *with respect to* $\|\cdot\|_\alpha$ *if and only if* $\lim_{k\to\infty} x^{(k)} = x$ *with respect to* $\|\cdot\|_\beta$.

有限维向量空间所有范数都是等价的

Definition 5.4.7. *Two given norms on a real or complex vector space are* equivalent *if whenever a sequence* $\{x^{(k)}\}$ *of vectors converges to a vector* x *with respect to the one of the norms, then it converges to* x *with respect to the other norm.*

有限维向量范数的收敛性

- （关于任一个基）依分量收敛等价于对任何向量范数收敛

Since all norms on $\mathbf{R}^n$ or $\mathbf{C}^n$ are equivalent to $\|\cdot\|_\infty$, for a given sequence of vectors $x^{(k)} = [x_i^{(k)}]_{i=1}^n$ we have $\lim_{k\to\infty} x^{(k)} = x$ with respect to any norm if and only if $\lim_{k\to\infty} x_i^{(k)} = x_i$ for each $i = 1, \ldots, n$.

- 收敛的充要条件：Cauchy序列

Definition 5.4.9. *A sequence $\{x^{(k)}\}$ in a vector space V is a* Cauchy sequence *with respect to a norm $\|\cdot\|$ if for each $\epsilon > 0$ there is a positive integer $N(\epsilon)$ such that $\|x^{(k_1)} - x^{(k_2)}\| \le \epsilon$ whenever $k_1, k_2 \ge N(\epsilon)$.*

Theorem 5.4.10. *Let $\|\cdot\|$ be a given norm on a finite-dimensional real or complex vector space V, and let $\{x^{(k)}\}$ be a given sequence of vectors in V. The sequence $\{x^{(k)}\}$ converges to a vector in V if and only if it is a Cauchy sequence with respect to the norm $\|\cdot\|$.*

向量范数的几何性质

- 向量范数的单位球
  - 紧集
  - 均衡集
  - 凸集
  - 以0为内点

$$\|\alpha x + (1-\alpha)y\| \le \|\alpha x\| + \|(1-\alpha)y\|$$
$$= \alpha\|x\| + (1-\alpha)\|y\| \le \alpha + (1-\alpha) \le 1$$

Theorem 5.5.8. *A set B in a finite-dimensional real or complex vector space V with positive dimension is the unit ball of a norm if and only if B (i) is compact, (ii) is convex, (iii) is equilibrated, and (iv) has 0 as an interior point.*

$$\|x\| = \begin{cases} 0 & \text{if } x = 0 \\ \min\left\{\frac{1}{t} : t > 0 \text{ and } tx \in B\right\} & \text{if } x \neq 0 \end{cases}$$

向量范数的单调性

- 向量范数是单调的当且仅当它是绝对的

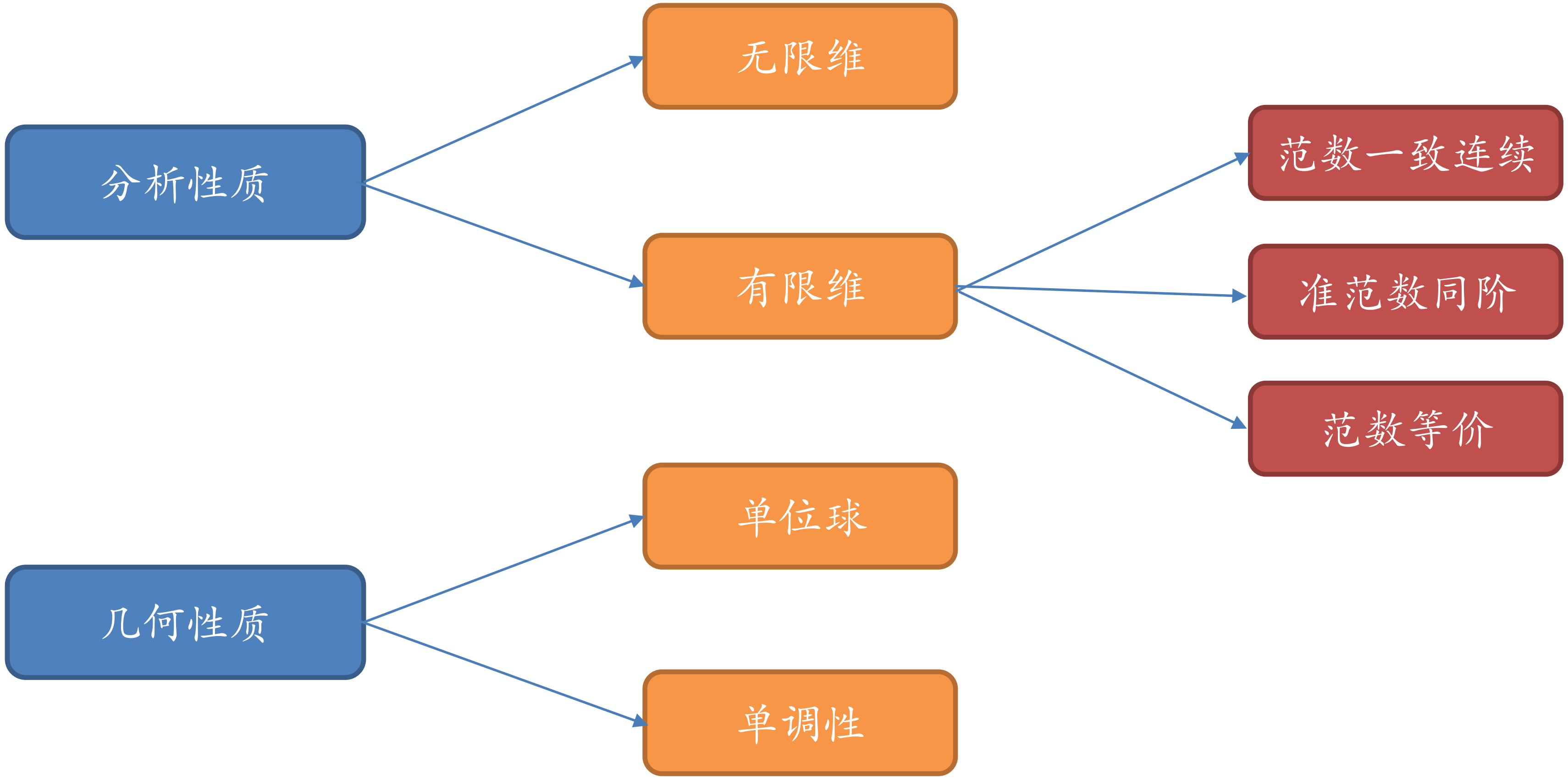
单调的，指的是对所有 $x, y \in \mathbf{F}^n$，$|x| \leqslant |y|$ 蕴涵 $\|x\| \leqslant \|y\|$；

绝对的，指的是对所有 $x \in \mathbf{F}^n$ 有 $\|x\| = \||x|\|$.

$$
\begin{aligned}
&\|[x_1,\cdots,x_{k-1},\alpha x_k,x_{k+1},\cdots,x_n]^T\| \\
&= \left\|\frac{1}{2}(1-\alpha)[x_1,\cdots,x_{k-1},-x_k,x_{k+1},\cdots,x_n]^T+\frac{1}{2}(1-\alpha)x+\alpha x\right\| \\
&\leqslant \frac{1}{2}(1-\alpha)\|[x_1,\cdots,x_{k-1},-x_k,x_{k+1},\cdots,x_n]^T\|+\frac{1}{2}(1-\alpha)\|x\|+\alpha\|x\| \\
&= \frac{1}{2}(1-\alpha)\|x\|+\frac{1}{2}(1-\alpha)\|x\|+\alpha\|x\| = \|x\|. \qquad (5.5.11)
\end{aligned}
$$

LAMDA Learning And Mining from DatA

向量范数的性质



Learning And Mining from DatA

矩阵范数

- M_n本身是n^2维向量空间
- $M_{m,n}$上定义的范数只满足前3条

A function $|||\cdot|||: M_n \to \mathbf{R}$ is a *matrix norm* if, for all $A, B \in M_n$, it satisfies the following five axioms:

(1)	$\|\|\|A\|\|\| \geq 0$	Nonnegative
(1a)	$\|\|\|A\|\|\| = 0$ if and only if $A = 0$	Positive
(2)	$\|\|\|cA\|\|\| = \|c\| \, \|\|\|A\|\|\|$ for all $c \in \mathbf{C}$	Homogeneous
(3)	$\|\|\|A + B\|\|\| \leq \|\|\|A\|\|\| + \|\|\|B\|\|\|$	Triangle Inequality
(4)	$\|\|\|AB\|\|\| \leq \|\|\|A\|\|\| \, \|\|\|B\|\|\|$	Submultiplicativity

矩阵的向量范数——矩阵元范数

Example. The l_1-*norm* defined for $A \in M_n$ by

$$\|A\|_1 = \sum_{i,j=1}^{n} |a_{ij}| \tag{5.6.0.1}$$

is a matrix norm because

$$\begin{aligned}
\|AB\|_1 &= \sum_{i,j=1}^{n} \left| \sum_{k=1}^{n} a_{ik} b_{kj} \right| \le \sum_{i,j,k=1}^{n} |a_{ik} b_{kj}| \\
&\le \sum_{i,j,k,m=1}^{n} |a_{ik} b_{mj}| = \left(\sum_{i,k=1}^{n} |a_{ik}| \right) \left(\sum_{j,m=1}^{n} |b_{mj}| \right) \\
&= \|A\|_1 \|B\|_1
\end{aligned}$$

矩阵的向量范数——矩阵元范数

Example. The l_2-norm (*Frobenius norm*, *Schur norm*, or *Hilbert–Schmidt norm*) defined for $A \in M_n$ by

$$\|A\|_2 = |\operatorname{tr} AA^*|^{1/2} = \left(\sum_{i,j=1}^{n} |a_{ij}|^2\right)^{1/2} \qquad (5.6.0.2)$$

$$\|AB\|_2 = \left(\sum_{i,j=1}^{n} \left|\sum_{k=1}^{n} a_{ik} b_{kj}\right|^2\right)^{1/2} \leq \left(\sum_{i,j=1}^{n} \left(\sum_{k=1}^{n} |a_{ik}|^2\right)\left(\sum_{m=1}^{n} |b_{mj}|^2\right)\right)^{1/2}$$

$$= \left(\sum_{i,k=1}^{n} |a_{ik}|^2\right)^{1/2} \left(\sum_{m,j=1}^{n} |b_{mj}|^2\right)^{1/2} = \|A\|_2 \|B\|_2$$

$$\|A\|_2 = \|A^*\|_2$$

$$\|A\|_2 = \|UAV\|_2$$

酉不变

$$\|A\|_2 = \sqrt{\sigma_1(A)^2 + \cdots + \sigma_n(A)^2}$$

矩阵的向量范数——矩阵元范数

- 不是所有的L_p向量范数都是矩阵范数

The l_∞-*norm* defined for $A \in M_n$ by

$$\|A\|_\infty = \max_{1 \le i, j \le n} |a_{ij}|$$

$$J = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \in M_2$$

$$J^2 = 2J, \ \|J\|_\infty = 1$$

$$2\|J\|_\infty = 2$$

向量范数诱导的矩阵范数

Definition 5.6.1. *Let* $\|\cdot\|$ *be a norm on* $\mathbf{C}^n$. *Define* $|||\cdot|||$ *on* M_n *by*

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

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

(a) $|||I||| = 1$

$$|||A||| = \max_{\|x\|=1} \|Ax\| \geq \|A\frac{y}{\|y\|}\| = \|Ay\|/\|y\|$$

(b) $\|Ay\| \le |||A|||\,\|y\|$ *for any* $A \in M_n$ *and any* $y \in \mathbf{C}^n$

(c) $|||\cdot|||$ *is a matrix norm on* M_n

$$\|ABx\| = \|A(Bx)\| \le |||A|||\,\|Bx\| \le |||A|||\,|||B|||$$

l_1 范数诱导的矩阵范数

- 极大列和矩阵范数

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

$$\|Ax\|_1 = \|x_1 a_1 + \cdots + x_n a_n\|_1 \le \sum_{i=1}^{n} \|x_i a_i\|_1 = \sum_{i=1}^{n} |x_i| \, \|a_i\|_1$$

Learning And Mining from DatA

l_∞范数诱导的矩阵范数

- 极大行和矩阵范数

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

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

Learning And Mining from DatA

l_2范数诱导的矩阵范数

- 谱范数

$$|||A|||_2 = \sigma_1(A), \text{ the largest singular value of } A$$

$$\begin{aligned}
\max_{\|x\|_2=1} \|Ax\|_2 &= \max_{\|x\|_2=1} \|V\Sigma W^* x\|_2 = \max_{\|x\|_2=1} \|\Sigma W^* x\|_2 \\
&= \max_{\|Wy\|_2=1} \|\Sigma y\|_2 = \max_{\|y\|_2=1} \|\Sigma y\|_2 \\
&\le \max_{\|y\|_2=1} \|\sigma_1 y\|_2 = \sigma_1 \max_{\|y\|_2=1} \|y\|_2 = \sigma_1
\end{aligned}$$

酉不变

$$|||UAV|||_2 = |||A|||_2$$

Schatten范数

- 将l_p范数应用到矩阵的奇异值组成的向量上

$$\|A\|_p = \left(\sum_{i=1}^{\min\{m,n\}} \sigma_i^p(A) \right)^{1/p}$$

- p=1, Nuclear norm: surrogate for minimizing the rank
- P=2, Frobenius norm

Learning And Mining from DatA

构造新的矩阵范数

Theorem 5.6.7. *Suppose that* $|||\cdot|||$ *is a matrix norm on* M_n *and* $S \in M_n$ *is nonsingular. Then the function*

$$|||A|||_S = |||SAS^{-1}||| \quad \textit{for all } A \in M_n$$

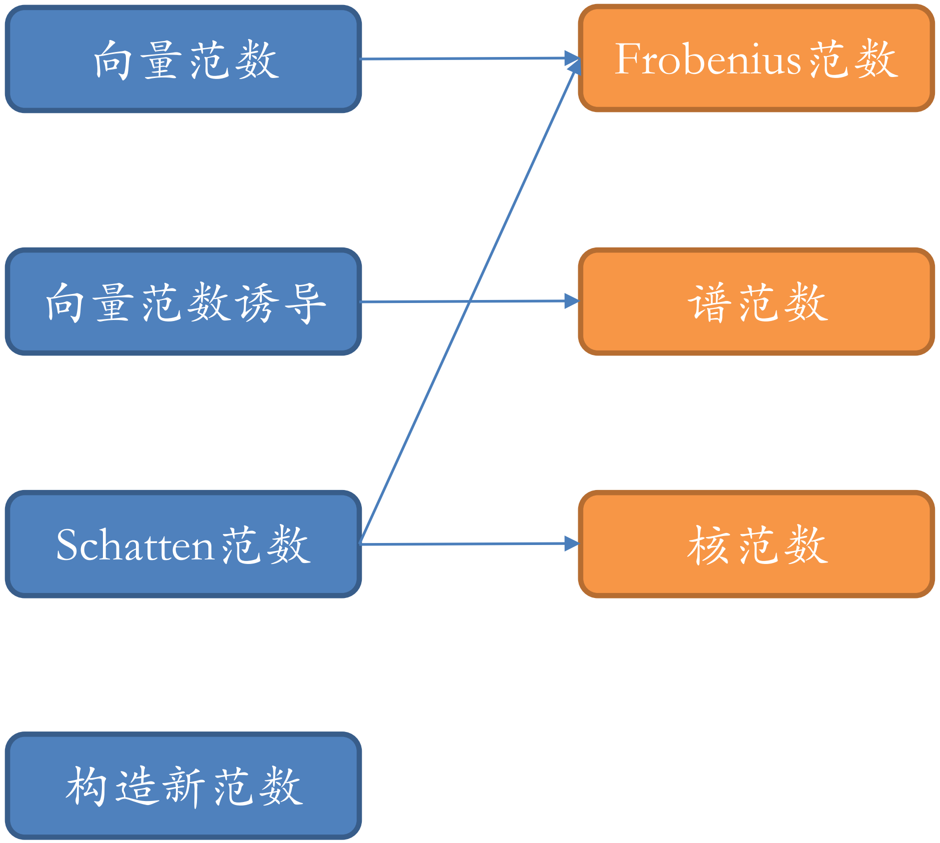
is a matrix norm. Moreover, if $|||\cdot|||$ *is induced by the norm* $\|\cdot\|$ *on* $\mathbf{C}^n$*, then the matrix norm* $|||\cdot|||_S$ *is induced by the norm* $\|\cdot\|_S$ *on* $\mathbf{C}^n$ *defined in (5.2.6).*

$$\begin{aligned} |||AB|||_S &= |||SABS^{-1}||| = |||(SAS^{-1})(SBS^{-1})||| \\ &\leq |||SAS^{-1}|||\ |||SBS^{-1}||| = |||A|||_S\ |||B|||_S \end{aligned}$$

$$\max_{\|x\|_S=1} \|Ax\|_S = \max_{\|Sx\|=1} \|SAx\| = \max_{\|y\|=1} \left\|SAS^{-1}y\right\| = |||SAS^{-1}|||$$

Learning And Mining from DatA

矩阵范数的种类



谱半径

Definition 1.2.9. *Let* $A \in M_n$. *The* spectral radius *of* A *is* $\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}$.

Theorem 5.6.9. *Let* $|||\cdot|||$ *be a matrix norm on* M_n, *let* $A \in M_n$, *and let* λ *be an eigenvalue of* A. *Then*

(a) $|\lambda| \le \rho(A) \le |||A|||$

$$|\lambda|\,|||X||| = |||\lambda X||| = |||AX||| \le |||A|||\,|||X|||$$

If A *is nonsingular, then*

(b) $\rho(A) \ge |\lambda| \ge 1/|||A^{-1}|||$

矩阵范数的下确界

- 谱范数是所有矩阵范数的下确界

$$\inf\{|||A||| : |||\cdot||| \text{ is an induced matrix norm}\}$$

Lemma 5.6.10. *Let $A \in M_n$ and $\epsilon > 0$ be given. There is a matrix norm $|||\cdot|||$ such that $\rho(A) \le |||A||| \le \rho(A) + \epsilon$.*

$$A = U\Delta U^*$$

$$D_t = \text{diag}(t, t^2, t^3, \ldots, t^n)$$

$$D_t \Delta D_t^{-1} = \begin{bmatrix} \lambda_1 & t^{-1}d_{12} & t^{-2}d_{13} & \ldots & t^{-n+1}d_{1n} \\ 0 & \lambda_2 & t^{-1}d_{23} & \ldots & t^{-n+2}d_{2n} \\ 0 & 0 & \lambda_3 & \ldots & t^{-n+3}d_{3n} \\ \cdot & \cdot & \cdot & \ldots & \cdot \\ 0 & 0 & 0 & \ldots & t^{-1}d_{n-1,n} \\ 0 & 0 & 0 & 0 & \lambda_n \end{bmatrix}$$

$$|||D_t \Delta D_t^{-1}|||_1 \le \rho(A) + \epsilon$$

$$|||B||| = |||D_t U^* B U D_t^{-1}|||_1 = |||(D_t U^*) B (D_t U^*)^{-1}|||_1$$

收敛性

- 矩阵幂级数的收敛性

Lemma 5.6.11. *Let* $A \in M_n$ *be given. If there is a matrix norm* $|||\cdot|||$ *such that* $|||A|||$ < 1, *then* $\lim_{k\to\infty} A^k = 0$, *that is, each entry of* A^k *tends to zero as* $k \to \infty$.

$$|||A^k||| \leq |||A|||^k \to 0 \text{ as } k \to \infty$$

向量范数等价

Theorem 5.6.12. *Let* $A \in M_n$. *Then* $\lim_{k\to\infty} A^k = 0$ *if and only if* $\rho(A) < 1$.

收敛性

- 任意矩阵k次幂级数的entry可以被谱半径bound

Corollary 5.6.13. *Let* $A \in M_n$ *and* $\epsilon > 0$ *be given. There is a constant* $C = C(A, \epsilon)$ *such that* $|(A^k)_{ij}| \leq C(\rho(A) + \epsilon)^k$ *for all* $k = 1, 2, \ldots$ *and all* $i, j = 1, \ldots, n$.

$\tilde{A} = [\rho(A) + \epsilon]^{-1} A$ $\qquad \tilde{A}^k \to 0 \text{ as } k \to \infty$ $\qquad |(\tilde{A}^k)_{ij}| \leq C$

- 谱半径的极限定义

Corollary 5.6.14 (Gelfand formula). *Let* $|||\cdot|||$ *be a matrix norm on* M_n *and let* $A \in M_n$. *Then* $\rho(A) = \lim_{k\to\infty} |||A^k|||^{1/k}$.

$\rho(A)^k = \rho(A^k) \leq |||A^k||| \longrightarrow \rho(A) \leq |||A^k|||^{1/k}$

$\tilde{A} = [\rho(A) + \epsilon]^{-1} A \longrightarrow |||\tilde{A}^k||| \leq 1 \text{ for all } k \geq N \longrightarrow |||A^k|||^{1/k} \leq \rho(A) + \epsilon$

Learning And Mining from DatA

总结

- 向量范数
  - 种类
  - 分析性质
  - 几何性质
- 矩阵范数
  - 种类
  - 谱半径及收敛性