Matrix norms

1. The Frobenius norm of a matrix $X \in \mathbb{R}^{m\times n}$

is $\|X\|_F = \left(Tr\,(X^T X)\right)^{1/2}$

$= \left(\sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}^2\right)^{1/2}$

X_{ij}

$i=1,\dots,m$

$j=1,\dots,n$

where Tr is the trace function:

For $A = [A_{ij}] \in \mathbb{R}^{n\times n}$

$Tr(A) = \sum_{i=1}^{n} A_{ii}$

$Tr(AB) = Tr(BA)$

$Tr(A+B) = TrA + TrB$

$Tr(aA) = a\,TrA$ for $a \in \mathbb{R}$.

(Exercise.)

2. The sum-absolute-value norm

$\|X\|_{sav} = \sum_{i=1}^{m}\sum_{j=1}^{n} |X_{ij}|$

(ℓ_1-norm for vectors)

3. The maximum-absolute-value norm

$\|X\|_{mav} = \max\left\{ |X_{ij}| : \begin{array}{l} i=1,\dots m \\ j=1,\dots n \end{array}\right\}$

(ℓ_∞-norm for vectors)

Operator norm: (induced norm)

Suppose $\|\cdot\|_a$, $\|\cdot\|_b$ are norms on $\mathbb{R}^m$ & $\mathbb{R}^n$, respectively. For $X \in \mathbb{R}^{m\times n}$,

$$\|X\|_{a,b} \triangleq \sup\{\|Xu\|_a : \|u\|_b \le 1\}$$

When $\|\ \|_a$, $\|\ \|_b$ are both Euclidean norms, the operator norm of X is its maximum singular value.

$$\|X\|_{2,2} \triangleq \sup\{\|Xu\|_2 : \|u\|_2 \le 1\}$$
$$= \sup\{(u^TX^TXu)^{1/2} : \|u\|_2 \le 1\}$$
$$= \sigma_{max}(X)$$
$$= (\lambda_{max}(X^TX))^{1/2};$$

also called the spectral norm or l_2-norm $(a=b=2)$.

$A \in \mathbb{R}^{n\times n}$.

$A = UDU^T$

U: orthogonal matrix.

D: diagonal matrix with eigenvalues of A.

$A \in \mathbb{R}^{m\times n}$ $A = \Sigma D \Lambda$

See Appendix A5

Suppose that S is a nonempty set of real numbers. A real number b is called a least upper bound of S if
(lower)

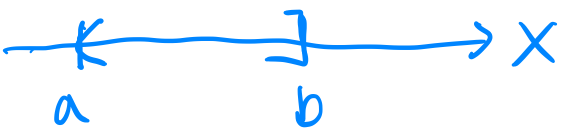
① $\forall a \in S$, $a \le b$.
($\geq$)

② If c is such that $a \le c$ $\forall a \in S$
($\geq$)

then $b \le c$.
($\geq$)

Denote this $b = \sup S$

$\inf(S)$ least lower bound.



$S = \{x : a < x \le b\}$
$= (a, b]$.

$a = \inf(S)$.

$b = \sup(S) := \max(S)$.

ℓ_∞-norm on $\mathbb{R}^m \cdot \mathbb{R}^n$.

$$\|X\|_\infty = \sup\{\|Xu\|_\infty : \|u\|_\infty \le 1\}$$

$$= \max_{i=1,\dots,m} \sum_{j=1}^{n} |X_{ij}|.$$

Proof.

$$\sup\{\|Xu\|_\infty : \|u\|_\infty \le \}$$

$$= \sup\left\{\max_i \left|\sum_j X_{ij} u_j\right| : \|u\|_\infty \le 1\right\}$$

$$\le \sup\Big\{\max_i \sum_j |X_{ij}| \cdot |u_j| : \underbrace{\|u\|_\infty \le 1}_{\Downarrow\ |u_j| \le 1}\Big\}$$

$$\le \sup\{\max_i \sum_j |X_{ij}|\}$$

$$= \max_i \sum_j |X_{ij}|.$$

We can achieve this term by choosing

$u_j = \mathrm{sgn}(X_{ij})$ for the i that maximizes $\sum_j |X_{ij}|$.

$$= \begin{cases} +1 & \text{if } X_{ij} > 0. \\ 0 & \text{if } X_{ij} = 0 \\ -1 & \text{if } X_{ij} < 0 \end{cases}$$

ℓ_1-norm on $\mathbb{R}^m \cdot \mathbb{R}^n$.

max-column-sum norm

$$\|X\|_1 = \max_{j=1,\dots,n} \sum_{i=1}^{m} |X_{ij}|$$

(Exercise.)

Definition

Let V be a vector space over a field $\mathbb{F}$.
$\mathbb{R}^n$, $\mathbb{R}^{m\times n}$ $\qquad$ $\mathbb{R}$

An inner product on V is a function

$$\langle \,,\, \rangle : V \times V \to F$$

such that $\forall\, x, y, z \in V$ & $c \in F$

(a) $\langle x+z, y\rangle = \langle x, y\rangle + \langle z, y\rangle$.

(b) $\langle cx, y\rangle = c\,\langle x, y\rangle$

(c) $\overline{\langle x, y\rangle} = \langle y, x\rangle$.

(d) $\langle x, x\rangle > 0$ if $x \neq 0$.

$F = \mathbb{R}$. $V : \mathbb{R}^n$

$\underline{\langle \,,\, \rangle}$ has two arguments from V, V.
it maps $(x, y) \in V\times V$ to a number in F

For $x, y \in \mathbb{R}^n$

$$\langle x, y\rangle = x^T y = \sum_{i=1}^{n} x_i y_i$$

$\|x\|_2 \triangleq \sqrt{\langle x, x\rangle}$ is a norm (prove it).

Cauchy-Schwartz inequality says that

$$|x^T y| \leq \|x\|_2 \cdot \|y\|_2.$$

Two vectors $x, y \in V$ are orthogonal if

$\langle x, y \rangle = 0$.

A vector v is a unit vector if

$\| v \| = 1$.

The standard inner product on $\mathbb{R}^{m \times n}$

is given by

$$\langle X, Y \rangle = \mathrm{Tr}(X^T Y)$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} X_{ij} Y_{ij}$$

for $X, Y \in \mathbb{R}^{m \times n}$.

Exercise: check: this is an inner product

Dual norm

Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$. The associated dual norm, denoted $\|\cdot\|_*$, is defined as

$$\|z\|_* = \sup\{z^T x : \underline{\|x\| \le 1}\}$$

(Exercise: show that $\|\cdot\|_*$ is a norm).

By definition, $z^T x \le \|x\| \cdot \|z\|_*$ $\forall x, z$

$$z^T \frac{x}{\|x\|} \le \|z\|_*$$

$\|x\|_{**} = \|x\|$ $\forall x \in \mathbb{R}^n$, n finite.

The dual of the Euclidean norm is the Euclidean norm, since

$$\sup\{z^T x : \|x\|_2 \le 1\}$$

$$\le \sup\{\|x\|_2 \|z\|_2 : \|x\|_2 \le 1\}$$ by Cauchy-Schwartz inequality.

$$= \|z\|_2$$

The dual norm of ℓ_∞-norm is the ℓ_1-norm

proof: $\sup\{z^T x : \|x\|_\infty \le 1\}$

$\le \sup\{\sum_i |z_i||x_i| : \|x\|_\infty \le 1\}$

$\le \sup\{\max_i |x_i| \cdot \sum_j |z_j| : \|x\|_\infty \le 1\}$

$\le \sum_j |z_j| = \|z\|_1$ #

The dual norm of ℓ_1-norm is the ℓ_∞ norm

(Exercise).

Norm ball: for $x \in \mathbb{R}^n$, $r > 0$

$$B(x, r) = \{ y \in \mathbb{R}^n : \| y - x \| \le r \}$$

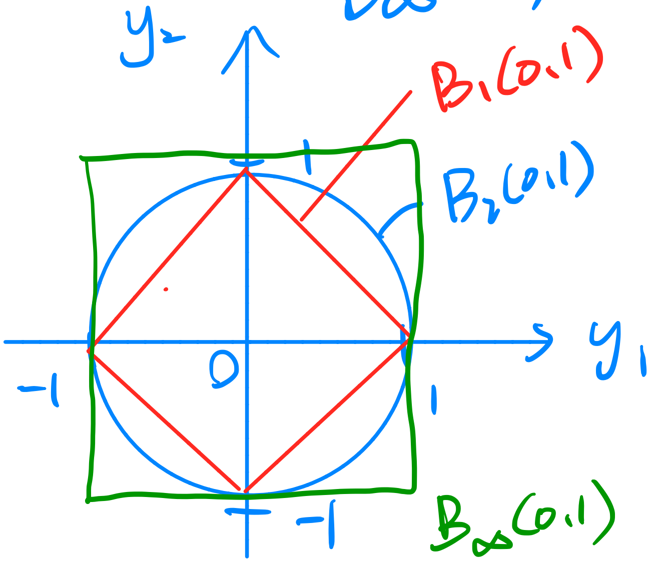
center radius

For $n=2$, $x=0$, $r=1$

2-norm Ball $B_2(0,1) = \{ y \in \mathbb{R}^2 : |y_1|^2 + |y_2|^2 \le 1 \}$

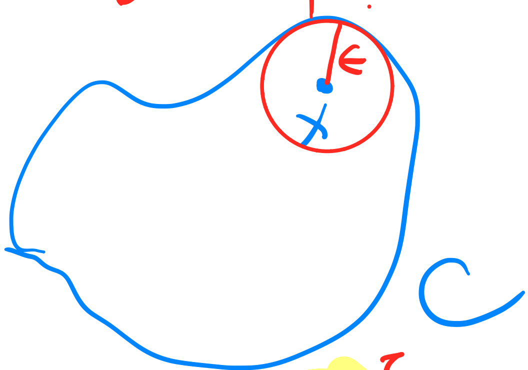
1 $B_1(0,1)$ $|y_1| + |y_2| \le 1$

∞ $B_\infty(0,1)$ $|y_1| \le 1, |y_2| \le 1$



Analysis

An element $x \in C \subseteq \mathbb{R}^n$ called an **interior point** of C if there exists an $\epsilon > 0$ for which $\underbrace{\{y : \|y-x\|_2 < \epsilon\}}_{= B(x,\epsilon)} \subset C$



The set of all points interior ~~to~~ C is called the interior of C, and is denoted $\text{int } C$

$C_1 = \{x : \|x\|_1 < 1\}$. $n=2$.

$C = (a, b] = \{x : a < x \le b\}$.

b is not an interior point of C.

$\text{int } C_1 = C_1$

$C_2 = \{x : \|x\|_1 \le 1\}$

If x has $\|x\|_1 = 1$.

$x \notin \text{int } C_2 = C_1$.

A set C is **open** if $\text{int } C = C$.

$C = \{x : \|x\|_2 < 1\}$

open ball $\Rightarrow$ $\mathbb{R}^n \setminus C$ is closed

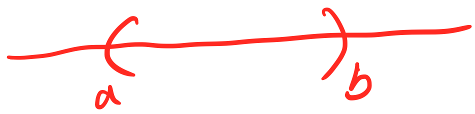
A set $C \subset \mathbb{R}^n$ is **closed** if its complement

$\mathbb{R}^n \setminus C = \{y \in \mathbb{R}^n : y \notin C\}$ is open.

set difference

The closure of C is the set of all limit points of convergent sequences.

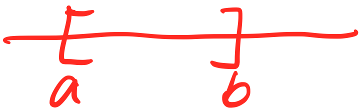
$$\mathrm{cl}\, C \triangleq \mathbb{R}^n \setminus \mathrm{int}(\mathbb{R}^n \setminus C)$$

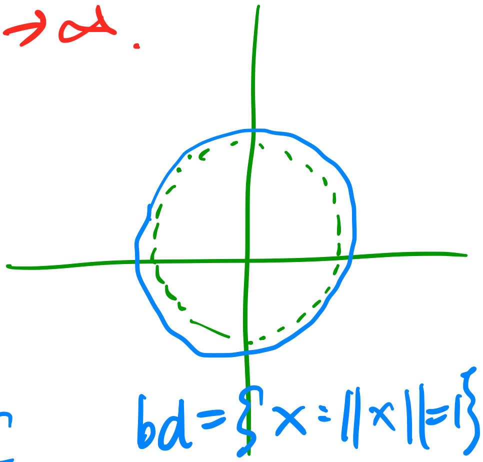
C: 

$\{a+\frac{1}{n} : n \in \mathbb{Z}^+\}$

$a+1, a+\frac{1}{2}, a+\frac{1}{3}, \cdots,$

$a \in \mathrm{cl}\, C$. $\to a$ as $n \to \infty$.

$\mathrm{cl}\, C =$ 



The boundary of C is

$$\mathrm{bd}\, C \triangleq \mathrm{cl}\, C \setminus \mathrm{int}\, C.$$

$\mathrm{bd} = \{x : \|x\| = 1\}$

A boundary point $x \in \mathrm{bd}\, C$ satisfies that

$\forall \epsilon > 0$, $\exists y \in C$, and $z \notin C$ with

$$\|y - x\|_2 < \epsilon, \quad \|z - x\|_2 < \epsilon$$