

Inner Product & Norms

Note Title

LECTURE 05

★ Inner Product

Def. The **inner product** between two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$ is defined as

$$\mathbf{x}^T \mathbf{y} = \sum_{i=1}^m x_i y_i \in \mathbb{R}$$

and is also written as

$$\mathbf{x} \cdot \mathbf{y}, (\mathbf{x}, \mathbf{y}), \text{ or } \langle \mathbf{x}, \mathbf{y} \rangle.$$

The **l^2 -norm** of $\mathbf{x} \in \mathbb{R}^m$ is defined as

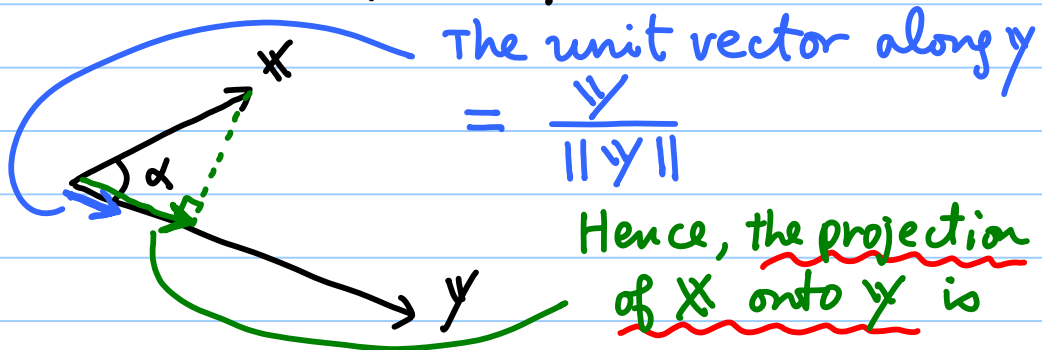
$$\|\mathbf{x}\|_2 := \sqrt{\mathbf{x}^T \mathbf{x}} = \sqrt{\sum_{i=1}^m |x_i|^2},$$

which is the **Euclidean length** of $\mathbf{x}$.

This is often written as $\|\mathbf{x}\|$.

The **angle** α between $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$, can be computed by

$$\cos \alpha = \frac{\mathbf{x}^T \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}$$



$$\begin{aligned} \text{proj}_{\mathbf{y}} \mathbf{x} &= (\|\mathbf{x}\| \cos \alpha) \frac{\mathbf{y}}{\|\mathbf{y}\|} \\ &= \frac{\mathbf{x}^T \mathbf{y}}{\|\mathbf{y}\|^2} \mathbf{y} \end{aligned}$$

LECTURE 05

★ Vector Norms

→ To quantify (or measure) the size (or length) of a vector

Def. A **norm** is a function

$$\|\cdot\| : \mathbb{R}^m \rightarrow \mathbb{R} \quad \text{s.t.}$$

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^m, \forall \alpha \in \mathbb{R}$$

$$(1) \quad \|\mathbf{x}\| \geq 0 \quad \text{and} \quad \|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$$

$$(2) \quad \|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \quad \text{The triangle inequality}$$

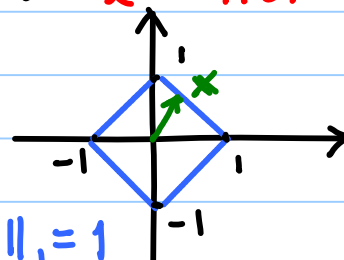
$$(3) \quad \|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$$

Examples

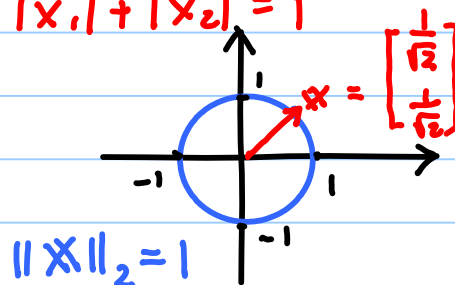
p-norms (or **l^p -norms**)

$$\|\mathbf{x}\|_1 := \sum_{i=1}^m |x_i|$$

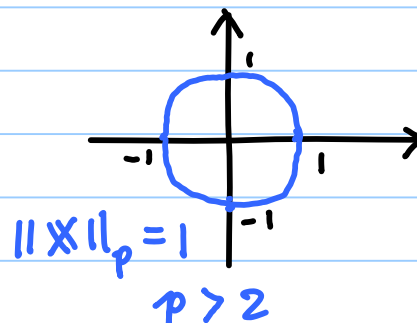
$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \rightarrow \|\mathbf{x}\|_1 = 1$
 $|x_1| + |x_2| = 1$



$$\|\mathbf{x}\|_2 := \left(\sum_{i=1}^m |x_i|^2 \right)^{1/2}$$

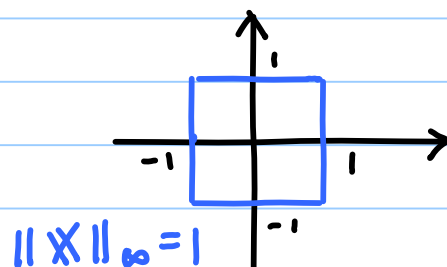


$$\|\mathbf{x}\|_p := \left(\sum_{i=1}^m |x_i|^p \right)^{1/p}$$



LECTURE 05

$$\|X\|_{\infty} := \max_{1 \leq i \leq m} |X_i|$$



Exercise: What is the vector $X \in \mathbb{R}^2$ that achieves $\max \|X\|_1$, subject to $\|X\|_2 = 1$?

★ Matrix Norms

- One can view an $m \times n$ matrix as a vector of length mn , then use one of the vector norms.

Def. The **Frobenius (Hilbert-Schmidt)** norm of $A \in \mathbb{R}^{m \times n}$ is defined as

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

$$= \left(\sum_{j=1}^n \|a_j\|_2^2 \right)^{1/2}$$

$$= \sqrt{\text{tr}(A^T A)}$$

$$= \sqrt{\text{tr}(A A^T)}$$

Def. For $X \in \mathbb{R}^{m \times n}$, $\text{tr}(X) := \sum_{i=1}^{\min(m,n)} X_{ii}$ is called the **trace** of X .

LECTURE 05

- However, $\exists$ different types of matrix norms called **induced matrix norms** (often called **operator norms**), which are defined in terms of the behavior of a matrix as an operator between its normed domain and range space.

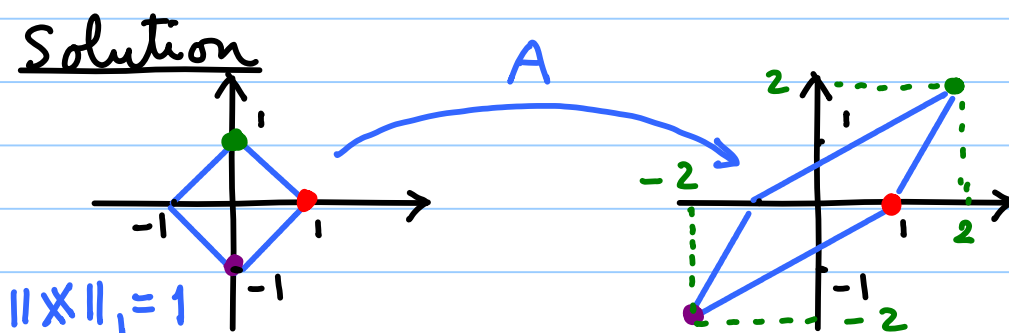
Def. Let $A \in \mathbb{R}^{m \times n}$. Then the **induced matrix** (or **operator**) **norm** is defined as

$$\begin{aligned} \|A\|_p &:= \sup_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \frac{\|Ax\|_p}{\|x\|_p} && \text{"sup" can be replaced by "max" here!} \\ &= \sup_{\substack{x \in \mathbb{R}^n \\ \|x\|_p = 1}} \|Ax\|_p \end{aligned}$$

In other words, $\|A\|_p$ is the smallest constant C satisfying $\|Ax\|_p \leq C \|x\|_p \quad \forall x \in \mathbb{R}^n$.

Example Consider $A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$.

Compute $\|A\|_1$, $\|A\|_2$, $\|A\|_\infty$.



LECTURE 05

Hence, $\sup \|A \mathbb{X}\|_1 = \max \|A \mathbb{X}\|_1$
 $= |2| + |2| = |-2| + |-2| = 4$

achieved for $\mathbb{X} = [0, 1]^T, [0, -1]^T$.

In fact,

$$\begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \rightarrow \left\| \begin{bmatrix} 2 \\ 2 \end{bmatrix} \right\|_1 = 2 + 2 = 4$$

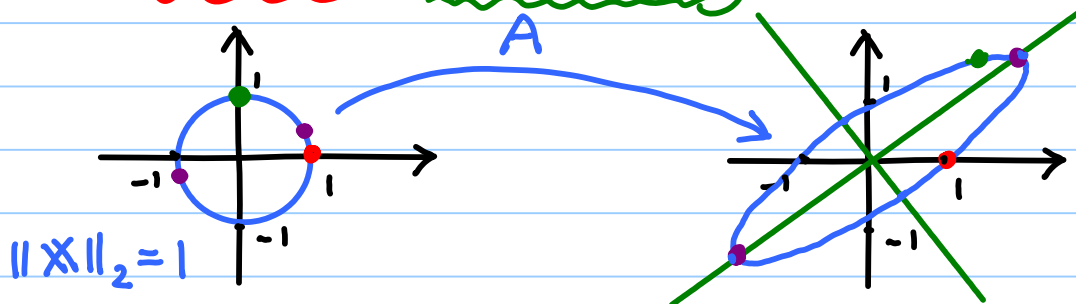
$$\begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \end{bmatrix} \rightarrow \left\| \begin{bmatrix} -2 \\ -2 \end{bmatrix} \right\|_1 = |-2| + |-2| = 4.$$

How about $\|A\|_2$?

$\Rightarrow$ As I'll prove later,

$\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)}$

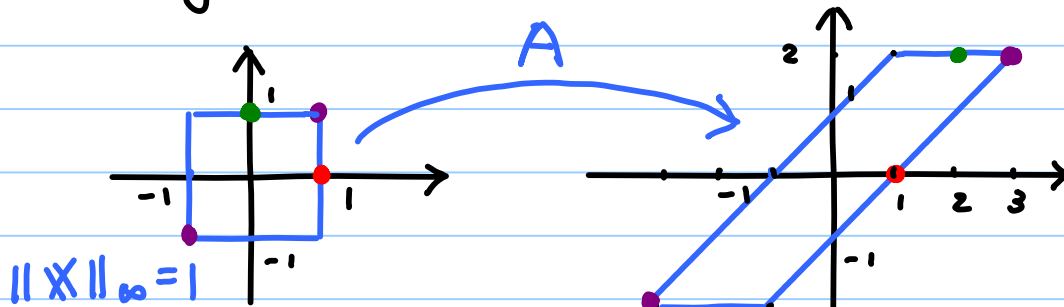
The largest eigenvalue of $A^T A$.



In this case $\|A\|_2 \approx 2.9208$

= the length of the major semi axis of the ellipse.

Finally, $\|A\|_\infty$.



LECTURE 05

From this figure, we can see

$$\|A\|_{\infty} = 3.$$

In fact,
$$\begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+2y \\ 2y \end{bmatrix}$$

$$\begin{aligned} \text{So, } \|A\|_{\infty} &= \max_{\substack{|x| \leq 1 \\ |y| \leq 1}} (|x+2y|, |2y|) \\ &= \max_{\substack{|x| \leq 1 \\ |y| \leq 1}} |x+2y| \\ &= 3 \quad \text{at } \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ or } \begin{pmatrix} -1 \\ -1 \end{pmatrix} \end{aligned}$$

• The p-norm of a diagonal matrix

Say
$$D = \begin{bmatrix} d_1 & & 0 \\ & d_2 & \\ 0 & & \ddots \\ & & & d_m \end{bmatrix}$$

Then, D maps the unit sphere in $\mathbb{R}^m$ (denoted by S^{m-1}) to a hyperellipsoid whose semi-axes are $|d_1|, \dots, |d_m|$.

$$\text{So, } \|D\|_2 = \max_{1 \leq i \leq m} |d_i|$$

$$\begin{aligned} \text{In fact, } \|D\|_p &= \max_{1 \leq i \leq m} |d_i| \\ &\text{for } \forall p \geq 1. \end{aligned}$$

LECTURE 05

- The 1-norm of a matrix
 $A \in \mathbb{R}^{m \times n}$

$$\|A\|_1 = \max_{1 \leq j \leq n} \|a_j\|_1$$

i.e., max. of 1-norms of col. vec's.

(Proof) Suppose $x \in \mathbb{R}^n$

$$\text{Then } \|Ax\|_1 = \left\| \sum_{j=1}^n x_j a_j \right\|_1$$

$$\begin{aligned} &\leq \sum_{j=1}^n |x_j| \|a_j\|_1 \\ &\quad \text{via the triangle ineq.} \quad \nearrow \\ &\leq \max_{1 \leq j \leq n} \|a_j\|_1 \cdot \sum_{j=1}^n |x_j| \\ &= \max_{1 \leq j \leq n} \|a_j\|_1 \cdot \|x\|_1 \end{aligned}$$

$$\text{So } \frac{\|Ax\|_1}{\|x\|_1} \leq \max_{1 \leq j \leq n} \|a_j\|_1$$

Now can this bound be attained at some x ? $\Rightarrow$ Yes!

$$\text{Let } \|a_k\|_1 = \max_{1 \leq j \leq n} \|a_j\|_1$$

Then set $x = e_k$

$$\Rightarrow \frac{\|Ae_k\|_1}{\|e_k\|_1} = \frac{\|a_k\|_1}{1} = \|a_k\|_1$$

//

LECTURE 05

- The 2-norm of a matrix

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

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)}$$

where $\lambda_{\max}(A^T A)$ is the largest (positive) eigenvalue of $A^T A$.

(Proof) Note the def. of $\|A\|_2$, i.e.,

$$\|A\|_2 = \sup_{\|x\|_2=1} \|Ax\|_2$$

Consider functions:

$$f(x) := \|Ax\|_2^2 = (Ax)^T (Ax)$$

$$= x^T A^T A x$$

$$g(x) := \|x\|_2^2 = x^T x$$

Then consider the following problem.

(*) Maximize $f(x)$ subject to $g(x)=1$.

$\Rightarrow$ This can be solved by the method of Lagrange multipliers (MAT 21C)

In other words, define

$$h(x, \lambda) := f(x) - \lambda(g(x) - 1)$$

The solution to (*) $\Leftrightarrow \frac{\partial h}{\partial x_i} = 0, 1 \leq i \leq n$
with $g(x)=1$

LECTURE 05

Can show that $\frac{\partial h}{\partial x_i} = 0 \quad 1 \leq i \leq n$
 leads to $\frac{\partial h}{\partial x} = 0$

i.e., $2 A^T A x - 2 \lambda x = 0$

$A^T A x = \lambda x \Rightarrow$ x : eigenvector
 λ : eigenvalue
 of $A^T A$

Now $g(x) = x^T x = 1$

$\underbrace{x^T A^T A x}_{\geq 0} = \lambda \underbrace{x^T x}_{=1} = \lambda$

So this is also ≥ 0

Finally,

$$\begin{aligned} \|A\|_2 &= \sup_{\|x\|_2=1} \|Ax\|_2 \\ &= \left(\sup_{x^T x=1} x^T A^T A x \right)^{\frac{1}{2}} \\ &= \sqrt{\lambda_{\max}(A^T A)} \quad // \end{aligned}$$

• The ∞ -norm of a matrix

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

$\|A\|_\infty = \max_{1 \leq i \leq m} \|\underbrace{a_{i.}}_{\text{ith row vector of } A}\|_1$

Note: Let $x \in \mathbb{R}^k = \mathbb{R}^{k \times 1}$

Then $x^T \in \mathbb{R}^{1 \times k}$ = a row vector with k entries

$\|x^T\|_1 = \|x\|_1 = \sum_{j=1}^k |x_j|$

also, note $A = \begin{bmatrix} a_{1.} \\ \vdots \\ a_{m.} \end{bmatrix}$

LECTURE 05

(Proof) $\|A \mathbb{X}\|_{\infty} \stackrel{\text{definition}}{=} \max_{1 \leq i \leq m} |a_{i.} \mathbb{X}|$

$$= \max_{1 \leq i \leq m} \left| \sum_{j=1}^n a_{ij} x_j \right|$$

$$\leq \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}| |x_j|$$

$$\leq \|\mathbb{X}\|_{\infty} \cdot \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|$$

So, $\frac{\|A \mathbb{X}\|_{\infty}}{\|\mathbb{X}\|_{\infty}} \leq \max_{1 \leq i \leq m} \|a_{i.}\|_1$

Suppose $\|\mathbb{X}\|_{\infty} = 1$, then for which $\mathbb{X}$, the equality

$$\|A \mathbb{X}\|_{\infty} = \max_{1 \leq i \leq m} \|a_{i.}\|_1$$

is attained?

$\Rightarrow$ Let $\|a_{k.}\|_1 = \max_{1 \leq i \leq m} \|a_{i.}\|_1$

Then define $\mathbb{X}$ as

$$x_j = \begin{cases} 1 & \text{if } a_{kj} \geq 0 \\ -1 & \text{if } a_{kj} < 0. \end{cases}$$

Clearly $\|\mathbb{X}\|_{\infty} = 1$ and

$$|a_{i.} \mathbb{X}| = \left| \sum_{j=1}^n a_{ij} x_j \right|$$

$$\leq \sum_{j=1}^n |a_{ij}| \underbrace{|x_j|}_{=1}$$

LECTURE 05

$$\begin{aligned} &= \sum_{j=1}^n |a_{ij}| \\ &= \|a_{i\cdot}\|_1 \quad 1 \leq i \leq m \end{aligned}$$

But if $i=k$, this becomes an equality,
and the max. is achieved!

$$\begin{aligned} \|A\|_{\infty} &= \max_{1 \leq i \leq m} \|a_{i\cdot}\|_1 \\ &= \|a_{k\cdot}\|_1 \\ &= \max_{1 \leq i \leq m} \|a_{i\cdot}\|_1 \end{aligned}$$

