Lecture 4

CS 510

NORM of
vectors &
matrices

Consider $\mathbb{R}^n$ & $\mathbb{C}^n$

$x \in \mathbb{R}^n$

$$\|x\|_2 = \left(\sum x_i^2\right)^{1/2}$$

$$\|x\|_p = \left(\sum x_i^p\right)^{1/p}, \quad 1 \le p \le \infty$$

In general a function N from $\mathbb{R}^n \longrightarrow \mathbb{R}_+$ is a norm if

$$N(x) \geq 0, \forall x \in \mathbb{R}^n$$
$$N(x) = 0 \iff x = 0$$
$$N(\alpha x) = |\alpha| N(x), \forall \alpha \in \mathbb{R}, x \in \mathbb{R}^n$$
$$N(x+y) \leq N(x) + N(y)$$

The Euclidean norm $\|x\|_2$ satisfies this.

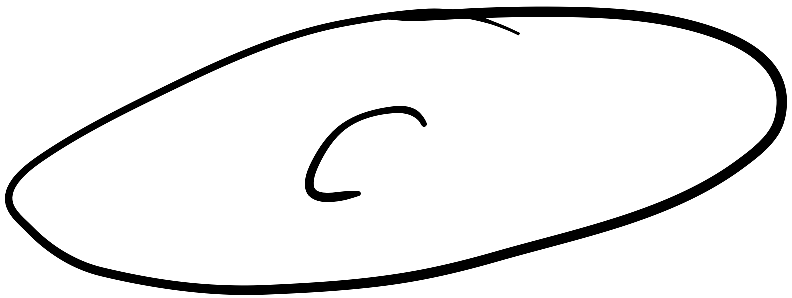
$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

$$\|x\|_\infty = \max\{|x_i| : i = 1, \dots, n\}$$

A subset C of $\mathbb{R}^n$ is convex if whenena $x, y \in C$

$$\alpha x + (1-\alpha) y \in C$$

$$\forall \alpha \in [0,1]$$



The unit ball with respect to a norm $\|\cdot\|$ is

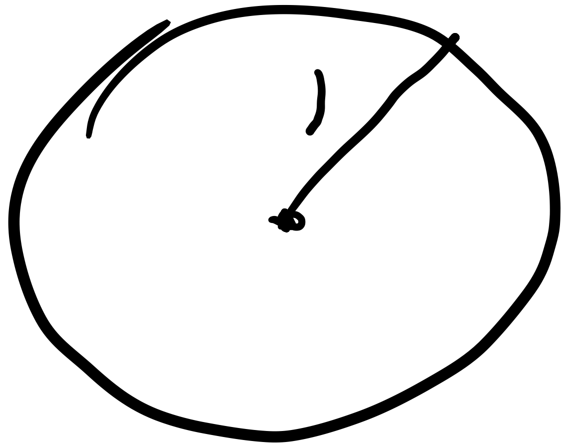
$$B = \{ x : \|x\| \leq 1 \}$$

Claim: B is convex

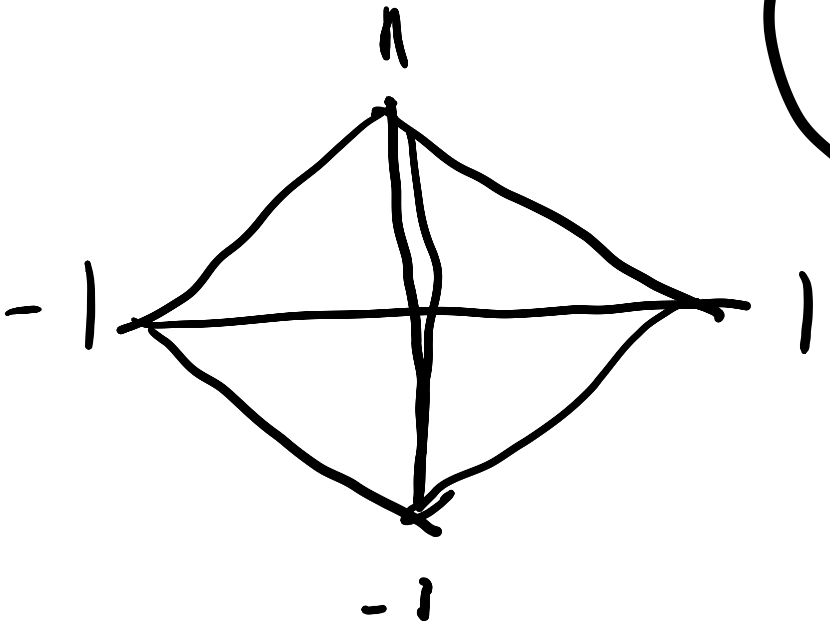
$x, y \in B$, $\|\alpha x + (1-\alpha) y\| \leq \|\alpha x\| + \|(1-\alpha) y\| = \alpha \|x\| + (1-\alpha)\|y\| \leq 1$

$\|\cdot\|_2$

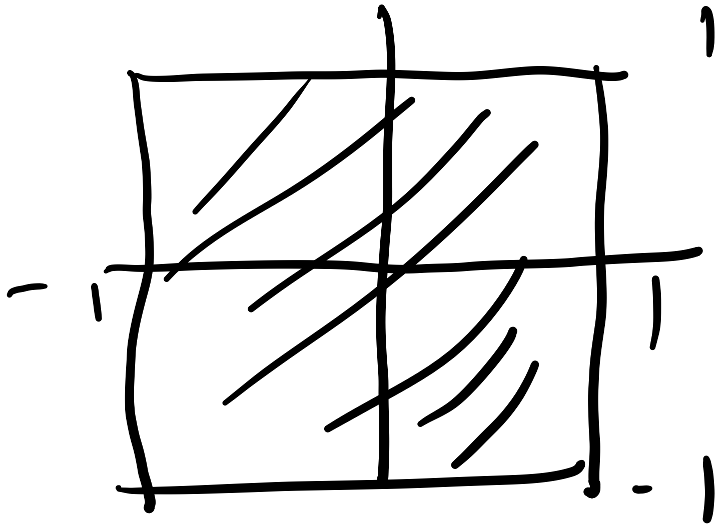
B



$\|\cdot\|_1$



$\|\cdot\|_\infty$



If $x \in \mathbb{C}^n$,

$$\|x\|_2 = \left(\sum_{i=1}^{n} x_i \bar{x}_i\right)^{1/2}$$

$\bar{x}_i$ = conjugate of x

$$\|x\|_1 = \sum |x_i| \;,\quad |x_i| = \sqrt{x_i \bar{x}_i}$$

$$\|x\|_\infty = \max\{\, |x_i| : i = 1, \ldots, n \,\}.$$

Given a sequence of pts

$\{x^0, x^1, \ldots, x^k, \} \subseteq \mathbb{R}^n$

we write $x^k \longrightarrow x_*$

if $\|x^k - x_*\| \longrightarrow 0$.

Consider the set of $n \times n$ real or complex matrices. They can be viewed as vectors in $\mathbb{R}^{n^2}$ or $\mathbb{C}^{n^2}$.

$\therefore$ we can speak of norm for matrices.

Given $n \times n$ matrices we consider a function defined over them as a norm if it satisfies:

1. $N(A) \geq 0$, $\forall A$, $N(A) = 0 \iff A = 0$
2. $N(\alpha A) = |\alpha| N(A)$
3. $N(A+B) \leq N(A) + N(B)$

plus two additional properties

4. $N(AB) \leq N(A)\,N(B)$

5. $N(AX) \leq N(A)\,\|X\|_2$

We write for $N(A) = \|A\|$

But will specify what the norm is..

Given any vector norm on $x \in \mathbb{R}^n$ or $\mathbb{C}^n$, it induces a norm on A.

Example will be given with $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$.

Matrix norm

Let A be $m \times n$ matrix on $\mathbb{R}$ or $\mathbb{C}$.

Given any norm $\|\cdot\|$ on $\mathbb{R}^n$ or $\mathbb{C}^n$, we define

$$\|A\| = \underset{\|x\| \leq 1}{\text{supremum}} \|Ax\|,$$

$\|x\| \leq 1$ can be replaced w. $\|x\| = 1$

What is $\|A\|_1$?

Pick $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$

$$AX = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix}$$

$$\|AX\|_1 = |a_{11}x_1 + a_{12}x_2| + |a_{21}x_1 + a_{22}x_2|$$

when is this sum maximized if $|X|_1 \leq 1$

If we choose $x_1 = \pm 1$, then

$$\|Ax\|_1 = |a_{11}| + |a_{21}|$$

If we choose $x_2 = \pm 1$, then

$$\|Ax\|_1 = |a_{12}| + |a_{22}|$$

So $$\|A\|_1 = \max\{|a_{11}| + |a_{21}|, |a_{12}| + |a_{22}|\}$$

$\therefore$ $\|A\|_1$ = max of 1-norm of columns of A

Works for general A.

What is $\|A\|_\infty$?

Pick $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$

$$Ax = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix}$$

$$\|Ax\|_\infty = \max_{\|x\|_\infty = 1} \left\{ |a_{11}X_{11} + a_{12}x_2|, \; |a_{21}x_1 + a_{22}x_2| \right\}$$

If $\|X\|_\infty = 1$ it means $|X_1| = 1 = |X_2|$
we can have

So
$\|A\|_\infty =$ maximum of $\| \|_1$ norm of rows of A.

What is $\|A\|_2$?

$$\|A\|_2 = \max_{\|x\|_2 \le 1} \|Ax\|_2$$

How to compute this ?

Let A be an $n \times n$ real matrix. eigenvalues of A are defined by $Ax = \lambda x$ where $x \neq 0$

We say λ is an eigen value of A if $AX = \lambda X$ for some $X \neq 0$.

Ex. $A = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$

$$AX = \begin{bmatrix} x_1 - x_2 \\ -x_1 + 2x_2 \end{bmatrix}$$

When is $AX = \lambda X$?

This happens if

$$\det(A - \lambda I) = 0$$

$$A - \lambda I = \begin{bmatrix} 1-\lambda & -1 \\ -1 & 2-\lambda \end{bmatrix}$$

$$\det(A - \lambda I) = (1-\lambda)(2-\lambda) + 1$$

$$= \lambda^2 - 3\lambda + 3 = 0$$

$$\lambda = \frac{+3 \pm \sqrt{9-12}}{2} = \frac{3 \pm \sqrt{3}\, i}{2}$$

So eigenvalues of real matrix could be complex numbers.

And the eigenvectors of A then complex vectors.

Given $A = (a_{ij})$,

$A^T = (a_{ji})$, transpose of A.

$A^* = (\overline{a_{ji}})$, conjugate transpose

Given $x \in \mathbb{R}^n$, $x^T x = \|x\|_2^2$.

Given $x \in \mathbb{C}^n$, $x^* x = \|x\|_2^2$

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

$$x^T y = \sum x_i y_i = \langle x, y \rangle$$

inner product

Given $x, y \in \mathbb{C}^n$

$$x^* y = \sum \bar{x}_i y_i = \langle x, y \rangle$$

inner product

$$\langle x, y \rangle \le \|x\|_2 \|y\|_2$$

Cauchy-Schwartz inequality

Given an $n \times n$ matrix A, eigenvalues of A are solutions to

$$f_A(\lambda) = \det(A - \lambda I) = 0$$

$$\det(A - \lambda I) = \det \begin{bmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & & \\ \vdots & & \ddots & \\ a_{2n} & \cdots & & a_{nn} - \lambda \end{bmatrix}$$

$$f_A(\lambda) = (a_{11}-\lambda)(a_{22}-\lambda)\cdots(a_{nn}-\lambda)$$

+ polynomial of degree $\leq n-2$

$f_A(\lambda)$ has n roots.

$$f_A(\lambda) = (-1)^n \lambda^n +$$

$$(-1)^{n-1}(a_{11}+\cdots+a_{nn})\lambda^{n-1} + \cdots$$

Note that sum of the roots of $f_A(\lambda) = 0$ is

$$a_{11} + \cdots + a_{nn}$$

Called Trace of A

$\mathrm{Tr}(A)$

What is $\|A\|_2$?

Suppose A is real.

$$\|A\|_2 = \max_{\|x\|_2 \le 1} \|Ax\|_2$$

what $(\|Ax\|_2)^2 = (Ax)^T(Ax)$

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

Clearly, maximizing this over $\|x\|_2 \le 1$ →

This is equivalent to

$$\max x^T A^T A x$$

$$\|x\|_2 = 1 \quad , \quad x^T x = 1$$

From Lagrange multiplier optimality condition: (gradient of function is proportional to gradient of constraint)

$$A^T A x = \lambda x$$

This in particular means λ is eigenvalue

From $A^T A x = \lambda x$

and $\|x\|_2 = 1$ we get

$$x^T A^T A x = \lambda x^T x = \lambda$$

This means the maximum occurs when λ is the largest eigen value of $A^T A$. So $\|A\|_2^2 = \sqrt{r(A^T A)}$

$r(A^T A) = |$largest eigenvalue of $A^T A|$

Largest eigenvalue of $A^T A$ is a real number because eigenvalue of real symmetric matrix are real.

So for $n \times n$ real matrix A

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

If A is symmetric,

$$A^T = A$$

So $A^T A = A^2$

Then $\|A\|_2 =$ largest eigenvalue of A in absolute value

What is $\|A\|_2$ if A is complex?

Fact:

If $\lambda_1, \ldots, \lambda_n$ are eigenvalues & $v_1, \ldots, v_n$ corresponding set of orthogonal eigenvectors, then any $x \in \mathbb{C}^n \rightarrow$

$$x = \sum \alpha_i u_i$$

NOW

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

$$Ax = A\left(\sum \alpha_i u_i\right) =$$

$$\sum \alpha_i A u_i = \sum \alpha_i \lambda_i u_i$$

$$\|AX\|_2 = \|\sum \alpha_i \lambda_i u_i\|$$

$$= \sqrt{(\sum \bar{\alpha}_i \bar{\lambda}_i u_i^*)(\sum \alpha_i \lambda_i u_i)}$$

$$= \sqrt{\sum |\alpha_i|^2 |\lambda_i|^2}$$

$$\|X\|_2 = 1 \Rightarrow \sum |\alpha_i|^2 = 1$$

so $\max \|AX\|_2 = \max |\lambda_i|$

$\therefore$ For general A

$$\|A\|_2 = \sqrt{r(A^*A)}$$

Thm. Let $r(A)$ be max $|\lambda|$ such that λ is an eigenvalue of A. Then for any matrix norm $\|A\|$, $r(A) \leq \|A\|$

Pf. $r(A) = |\lambda| = \|\lambda x\|$ for some x, $\|x\|=1$

$= \|Ax\| \leq \|A\| \|x\| = \|A\|$.

Frobenius norm of a matrix

$$F(A) = \left(\sum_{1 \le i,j \le n} |a_{ij}|^2 \right)^{1/2}$$

i.e. we look at A as a vector in $\mathbb{C}^{n^2}$.

We can show it is a matrix norm

The only part we need to prove

$$F(AB) \leq F(A)F(B)$$

Thm. Let A be $n \times n$. Suppose $\|A\| < 1$ for some norm.

Then $I - A$ is invertible

&

$$(I-A)^{-1} = I + A + A^2 + \cdots + A^n + \cdots$$

Pf:

$$(I - A)(I + A \cdots + A^m) =$$

$$I - A^{m+1}$$

so

$$1 + A + A^2 + \cdots + A^m = (I - A)^{-1}(I - A^{m+1})$$

As $m \to \infty$ $A^{m+1} \to 0$

Thm. For any $n \times n$ matrix

$$I + A + \frac{A^2}{2!} + \cdots + \frac{A^n}{n!} + \cdots$$

converges to a matrix we denote by e^A

Linear systems

$Ax = b$, A $n \times n$

If A is invertible there is a unique solution

$$x = A^{-1} b$$

Suppos

$$Ax = b$$

$$A\hat{x} = \hat{b}$$

where $\hat{x}$ is the unique solution to a system with perturbed b. How much does $\hat{x}$ change?

Let $\delta x = x - \hat{x}$, $\delta b = b - \hat{b}$

we ~~get~~

$$A\,\delta x = \delta b$$

or $\delta x = A^{-1}\delta b$

$$\|\delta x\| \leq \|A^{-1}\|\,\|\delta b\| \quad \text{— (1)}$$

Also from $Ax = b$

$$\|b\| \leq \|A\|\,\|x\| \quad \text{— (2)}$$

So

$$\frac{1}{\|x\|} \leq \frac{\|A\|}{\|b\|} \quad — \quad 3$$

From 1 & 3 we get

$$\frac{\|\delta x\|}{\|x\|} \leq \underbrace{\|A\|\,\|A^{-1}\|}_{\text{cond}(A)} \frac{\|\delta b\|}{\|b\|}$$