Math 6610: Analysis of Numerical Methods, I
Linear algebraic preliminaries

Department of Mathematics, University of Utah

Fall 2026

Resources:	Trefethen and Bau 1997, Lectures 1, 2, 3
	Atkinson 1989, Sections 7.1, 7.3
	Salgado and Wise 2022, Sections 1.1, 1.2

Notation

We'll use some standard math notation

- $\mathbb{C}$, $\mathbb{R}$, $\mathbb{N}$
- $\in$, $\forall$, $\exists$, !
- $\left\{x \in \mathbb{C} \;\middle|\; \mathsf{Im}\{x\} \in \mathbb{N}\right\}$
- $z = x + iy$ for $x, y \in \mathbb{R} \Longrightarrow \bar{z} = z^* \coloneqq x - iy$

Notation

We'll use some standard math notation

- $\mathbb{C}$, $\mathbb{R}$, $\mathbb{N}$
- $\in$, $\forall$, $\exists$, !
- $\{x \in \mathbb{C} \;\big|\; \mathsf{Im}\{x\} \in \mathbb{N}\}$
- $z = x + iy$ for $x, y \in \mathbb{R} \Longrightarrow \bar{z} = z^* \coloneqq x - iy$

Vectors, matrices, etc:

- $\mathbf{u} \in \mathbb{C}^n$
- $\mathbf{A} \in \mathbb{C}^{m \times n}$
- linear independence
- rank
- (conjugate) transpose
- determinant
- matrix inverse
- subspaces defined by $\mathbf{A}$: range, kernel, cokernel, corange

Hilbertian structure

$\mathbb{C}^n$ endowed with the standard inner product is a Hilbert space. If $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$,

- $\langle \mathbf{u}, \mathbf{v} \rangle$, $\|\mathbf{u}\|$
- $\angle(\mathbf{u}, \mathbf{v})$
- $\mathbf{u} \perp \mathbf{v}$
- $\mathsf{Proj}_{\mathbf{v}} \mathbf{u}$
- orthogonal and orthonormal sets

All the above is also well-defined in $\mathbb{R}^n$.

Matrices and matrix algebra

An $m \times n$ matrix $\mathbf{A}$ is a tableau of elements (from $\mathbb{R}$ or $\mathbb{C}$):

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in \mathbb{C}^{m\times n}, \qquad (A)_{j,k} = a_{j,k}, \qquad j \in [m], k \in [n]$$

A matrix is "just" a vector with "2D" indices.

Matrices come with a natural algebra, i.e., sum and product operations involving matrices:

- Product of a scalar and a matrix
- Sum of two matrices (of the same size)
- Product of two matrices (of conforming sizes)

$$\mathbf{A} \in \mathbb{C}^{m\times n}, \mathbf{B} \in \mathbb{C}^{n\times k} \quad \Longrightarrow \quad \mathbf{AB} \in \mathbb{C}^{m\times k}, \;\; (AB)_{j,k} = \sum_{\ell=1}^{n} a_{j\ell} b_{\ell k}.$$

The "four fundamental subspaces"

Let $\mathbf{A} \in \mathbb{C}^{m \times n}$ be given. The four fundamental subspaces are uniquely defined:

- $\mathbb{C}^m \supset \mathscr{R}(\mathbf{A}) = \mathsf{range}(\mathbf{A}) = \mathsf{Im}(\mathbf{A})$, the "column space" of $\mathbf{A}$
- $\mathbb{C}^n \supset \mathscr{R}(\mathbf{A}^*) = \mathsf{corange}(\mathbf{A})$, the "row space" or "corange" of $\mathbf{A}$.
- $\mathbb{C}^n \supset \mathscr{K}(\mathbf{A}) = \mathsf{ker}(\mathbf{A})$, the "nullspace" or "kernel" of $\mathbf{A}$
- $\mathbb{C}^m \supset \mathscr{K}(\mathbf{A}^*) = \mathsf{coker}(\mathbf{A})$, the "left nullspace" or "cokernel" of $\mathbf{A}$.

The "four fundamental subspaces"

Let $\mathbf{A} \in \mathbb{C}^{m \times n}$ be given. The four fundamental subspaces are uniquely defined:

- $\mathbb{C}^m \supset \mathscr{R}(\mathbf{A}) = \mathsf{range}(\mathbf{A}) = \mathsf{Im}(\mathbf{A})$, the "column space" of $\mathbf{A}$
- $\mathbb{C}^n \supset \mathscr{R}(\mathbf{A}^*) = \mathsf{corange}(\mathbf{A})$, the "row space" or "corange" of $\mathbf{A}$.
- $\mathbb{C}^n \supset \mathscr{K}(\mathbf{A}) = \mathsf{ker}(\mathbf{A})$, the "nullspace" or "kernel" of $\mathbf{A}$
- $\mathbb{C}^m \supset \mathscr{K}(\mathbf{A}^*) = \mathsf{coker}(\mathbf{A})$, the "left nullspace" or "cokernel" of $\mathbf{A}$.

Essentially by definition: $\mathscr{K}(\mathbf{A})$ contains all vectors $\mathbf{v}$ satisfying $\mathbf{A}\mathbf{v} = \mathbf{0}$. I.e., if $\mathbf{w}_1, \ldots, \mathbf{w}_m \in \mathbb{C}^n$ are conjugate-transposed rows of $\mathbf{A}$, then $\mathbf{v}$ is orthogonal to $\mathsf{span}\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$.

The "four fundamental subspaces"

Let $\mathbf{A} \in \mathbb{C}^{m \times n}$ be given. The four fundamental subspaces are uniquely defined:

- $\mathbb{C}^m \supset \mathscr{R}(\mathbf{A}) = \mathsf{range}(\mathbf{A}) = \mathsf{Im}(\mathbf{A})$, the "column space" of $\mathbf{A}$
- $\mathbb{C}^n \supset \mathscr{R}(\mathbf{A}^*) = \mathsf{corange}(\mathbf{A})$, the "row space" or "corange" of $\mathbf{A}$.
- $\mathbb{C}^n \supset \mathscr{K}(\mathbf{A}) = \mathsf{ker}(\mathbf{A})$, the "nullspace" or "kernel" of $\mathbf{A}$
- $\mathbb{C}^m \supset \mathscr{K}(\mathbf{A}^*) = \mathsf{coker}(\mathbf{A})$, the "left nullspace" or "cokernel" of $\mathbf{A}$.

Essentially by definition: $\mathscr{K}(\mathbf{A})$ contains all vectors $\mathbf{v}$ satisfying $\mathbf{A}\mathbf{v} = \mathbf{0}$. I.e., if $\mathbf{w}_1, \ldots, \mathbf{w}_m \in \mathbb{C}^n$ are conjugate-transposed rows of $\mathbf{A}$, then $\mathbf{v}$ is orthogonal to $\mathsf{span}\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$.

Theorem (Fundamental Theorem of Linear Algebra)

For any $\mathbf{A} \in \mathbb{C}^{m \times n}$,

$$\begin{aligned}
n &= \dim \mathsf{corange}(\mathbf{A}) + \dim \mathsf{ker}(\mathbf{A}), & \mathsf{corange}(\mathbf{A}) &\perp \mathsf{ker}(\mathbf{A}), & \mathbb{C}^n &= \mathsf{corange}(\mathbf{A}) \oplus \mathsf{ker}(\mathbf{A}) \\
m &= \dim \mathsf{range}(\mathbf{A}) + \dim \mathsf{coker}(\mathbf{A}), & \mathsf{range}(\mathbf{A}) &\perp \mathsf{coker}(\mathbf{A}), & \mathbb{C}^m &= \mathsf{range}(\mathbf{A}) \oplus \mathsf{coker}(\mathbf{A})
\end{aligned}$$

Norms

recall : Δ ineq $\Rightarrow$ $\|\theta x + (1-\theta) y\| \leq \theta \|x\| + (1-\theta)\|y\|$

$\Rightarrow$ $\|\cdot\|$ convex.

Metrizing linear spaces is a big business in mathematics.

Given a vector space V, a map $\|\cdot\| : V \to \mathbb{R}$ is a *norm* if it satisfies all the following properties:

- $\|\mathbf{x}\| \geq 0 \;\; \forall \mathbf{x} \in V$
- $\|\mathbf{x}\| = 0$ iff $\mathbf{x} = \mathbf{0}$
- $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \;\; \forall \mathbf{x}, \mathbf{y} \in V$
- $\|c\mathbf{x}\| = |c|\|\mathbf{x}\| \;\; \forall \mathbf{x} \in V,\; c \in \mathbb{C}.$

} Measures 'size' of vectors in V. : get a distance by $d(x,y) \triangleq \|x-y\|$

We are mostly concerned with standard examples $V = \mathbb{R}^n, \mathbb{C}^n, \mathbb{C}^{m \times n}$, etc.

recall : $\|x\| = \|x + y - y\| \leq \|x - y\| + \|y\| \Rightarrow \|x\| - \|y\| \leq \|x-y\|$

also by sym : $\|y\| - \|x\| \leq \|x-y\|$

$\Rightarrow$ $\big|\, \|x\| - \|y\| \,\big| \leq \|x-y\|$ $\Rightarrow$ $\|\cdot\|$ cts

Vector norms

The "standard" examples of vector norms are the ℓ^p norms.

With $\mathbf{x} \in \mathbb{C}^n$:

$$\|\mathbf{x}\|_p^p := \sum_{j\in[n]} |x_j|^p, \qquad p \in [1,\infty)$$

$$\|\mathbf{x}\|_\infty := \max_{j\in[n]} |x_j|, \qquad p = \infty.$$

Take $x = e_1 = (1,0,0,\cdots)$
$y = e_2 = (0,1,0,\cdots)$

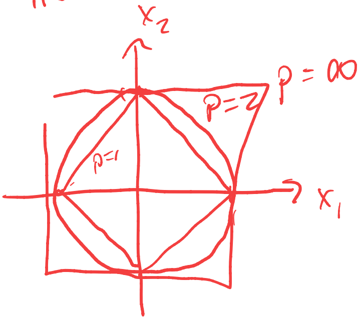
Then
$\|x+y\|_p = (1^p + 1^p)^{\frac{1}{p}} = 2^{\frac{1}{p}}$

But by Δ ineq:

$\|x+y\|_p \le \|x\|_p + \|y\|_p$
$= 2.$

and for $p<1$, $2^{\frac{1}{p}} > 2$

$\mathbb{R}^2$



$p=1$

$\|x\|_1 = \sum |x_i| = |x_1| + |x_2| \overset{set}{=} 1$

$\Rightarrow |x|_1 = 1 - |x|_2.$

$\|x_2\|_2 = \sqrt{x_1^2 + x_2^2} \overset{set}{=} 1$

$\|x\|_\infty \overset{set}{=} 1 \Rightarrow x_1, x_2 \le 1.$

Vector norms

The "standard" examples of vector norms are the ℓ^p norms.

With $\mathbf{x} \in \mathbb{C}^n$:

$$\|\mathbf{x}\|_p^p := \sum_{j \in [n]} |x_j|^p, \qquad p \in [1, \infty)$$

$$\|\mathbf{x}\|_\infty := \max_{j \in [n]} |x_j|, \qquad p = \infty.$$

Example

Show that $\|\cdot\|_2$ on $\mathbb{C}^n$ is a norm.

First, observe that for $\langle x, y \rangle = y^* x = \sum_i x_i y_i^*$, we have $\|x\|_2^2 = \langle x, x \rangle$.

Recall that for this Euclidean inner product on $\mathbb{R}^n$ we define

$$\langle x, y \rangle = \|x\| \, \|y\| \cos(\angle(x,y))$$

$$\Rightarrow \quad |\langle x, y \rangle| \leq \|x\| \, \|y\|$$

This carries to $\mathbb{C}^n$ as well (and to general Hilbert spaces)

we call this the Cauchy-Schwartz inequality

$$|\langle x, y \rangle| \leq \|x\| \, \|y\|$$

WTS $\|\cdot\|_2$ a norm:

recall: $\|x\|_2 = \sqrt{\sum_i |x_i|^2}$

$\|x\|_2 \geq 0$ ✓

$\|x\|_2 = 0 \Rightarrow \|x\|_2 = 0$ ✓

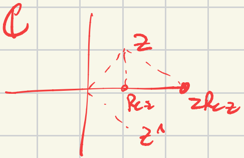
$\|cx\|_2 = |c|\|x\|_2$ ✓

for Δ-ineq:

$$
\begin{aligned}
\|x+y\|_2^2 &= \langle x+y, x+y\rangle \\
&= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\
&= \|x\|_2^2 + \|y\|_2^2 + \langle x,y\rangle + \langle x,y\rangle^* \\
&= \|x\|_2^2 + \|y\|_2^2 + 2\,\mathrm{Re}\langle x,y\rangle \\
&\leq \|x\|_2^2 + \|y\|_2^2 + 2|\langle x,y\rangle| \\
&\overset{\text{Cauchy-Schwarz}}{\leq} \|x\|_2^2 + \|y\|_2^2 + 2\|x\|\|y\| = (\|x\|+\|y\|)^2.
\end{aligned}
$$

$\square$

since $(\langle x,y\rangle)^* = \left(\sum_i x_i y_i^*\right)^* = \sum_i x_i^* y_i$
$= \langle y,x\rangle$



$\mathrm{Re}\, z \leq |z| = \sqrt{\mathrm{Re}\,z^2 + \mathrm{Im}\,z^2}$

Matrix norms

convert matrix to vector by "flattening"; apply vector norm

One straightforward class of matrix norms consists of "entrywise" norms:

$$\|\mathbf{A}\|_{p,p} := \|\mathsf{vec}(\mathbf{A})\|_p, \qquad p \in [1, \infty],$$

where $\mathsf{vec}(\cdot)$ is the *vectorization* function.

$$\mathbb{C}^{m\times n} \ni \underline{A} = \begin{pmatrix} | & & | \\ \underline{a}_1 & \cdots & \underline{a}_n \\ | & & | \end{pmatrix} \Rightarrow \mathsf{vec}(\underline{A}) = \begin{pmatrix} | \\ \underline{a}_1 \\ | \\ | \\ \underline{a}_2 \\ | \\ \vdots \\ | \\ \underline{a}_n \\ | \end{pmatrix} \in \mathbb{C}^{mn}$$

could also stack across rows; doesn't matter since vector p norms are permutation invariant.

Matrix norms

One straightforward class of matrix norms consists of "entrywise" norms:

$$\|\mathbf{A}\|_{p,p} \coloneqq \|\mathsf{vec}(\mathbf{A})\|_p, \qquad\qquad p \in [1, \infty],$$

where $\mathsf{vec}(\cdot)$ is the *vectorization* function.

There are "mixed" entrywise norm definitions, corresponding to taking ℓ^p vector norms of each column, and then a vector ℓ^q norm of the resulting vector of norms,

$$\|\mathbf{A}\|_{p,q} \coloneqq \left(\sum_{j \in [n]} \left(\sum_{i \in [m]} |a_{i,j}|^p \right)^{q/p} \right)^{1/q}.$$

↳ more useful for applications where matrices are data arrays rather than linear maps

$$\begin{pmatrix} | & & | \\ a_1 & \cdots & a_n \\ | & & | \end{pmatrix} \to \left\| \begin{pmatrix} \|a_1\|_p \\ \vdots \\ \|a_n\|_p \end{pmatrix} \right\|_q$$

Matrix norms

One straightforward class of matrix norms consists of "entrywise" norms:

$$\|\mathbf{A}\|_{p,p} \coloneqq \|\mathsf{vec}(\mathbf{A})\|_p, \qquad\qquad p \in [1, \infty],$$

where $\mathsf{vec}(\cdot)$ is the *vectorization* function.

There are "mixed" entrywise norm definitions, corresponding to taking ℓ^p vector norms of each column, and then a vector ℓ^q norm of the resulting vector of norms,

$$\|\mathbf{A}\|_{p,q} \coloneqq \left(\sum_{j \in [n]} \left(\sum_{i \in [m]} |a_{i,j}|^p \right)^{q/p} \right)^{1/q}.$$

will show up loads later...

keep an eye on it!

A particularly useful entrywise norm is the *Frobenius norm*,

$$\|\mathbf{A}\|_F \coloneqq \|\mathbf{A}\|_{2,2}. = \sqrt{Tr(A^T A)}$$

$$= \sqrt{\sum_i \sigma_i^2}$$

$$= \sqrt{\langle A, A \rangle}$$

Induced matrix norms

↳ 'natural' norm on linear maps between normed spaces.

A more conceptual collection of matrix norms are *induced* by vector norms.

By viewing $\mathbf{A} \in \mathbb{C}^{m \times n}$ as the mapping $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$, norms can be defined as the maximum relative "size" of this mapping:

$$\|\mathbf{A}\|_p \coloneqq \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_p}{\|\mathbf{x}\|_p}, \quad p \in [1, \infty].$$

linearity $= \sup_{\|x\|_p = 1} \|Ax\|$

note for all $x \in V$

$\|A\|_p \leq \|Ax\|_p / \|x\|_p$

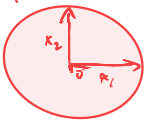
$\Rightarrow \|Ax\| \leq \|A\|_p \|x\|_p$

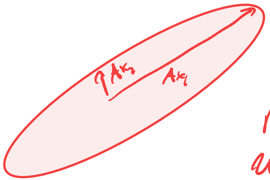
max stretch factor.

(Note that $\mathbf{A}$ can be rectangular here.)

That these are proper norms is direct from the fact that $\| \cdot \|_p$ is a norm on $\mathbb{C}^m$.

$\| \cdot \|_p$ - unit ball





$\|A\|_p$ measures maximal norm amplification under A.

Norm equivalence

$\| \ \|_{L^2}$, $\| \ \|_{L^\infty}$ not equivalent on say $\mathcal{C}([a,b])$. e.g.

Cannot emphasize enough that this only holds for $\dim V < \infty$

A rather important and useful fact is that any two norms on the same finite-dimensional vector space are *equivalent*.

Theorem (All norms on a finite-dimensional space are equivalent)

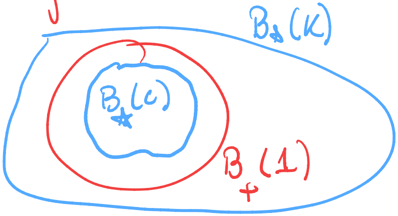
Let V be an n-dimensional vector space, and let $\|\cdot\|_*$ and $\|\cdot\|_+$ be any two norms on this space. Then there are strictly positive constants c and k such that for all $\mathbf{x} \in V$,

$$c\|\mathbf{x}\|_* \leq \|\mathbf{x}\|_+ \leq k\|\mathbf{x}\|_*. \quad \iff \frac{1}{k}\|x\|_+ \leq \|x\|_* \leq \frac{1}{c}\|x\|_+$$

The constants c and k can depend on V (in particular n) and the choice of $\|\cdot\|_*$ and $\|\cdot\|_+$, but not on $\mathbf{x}$.

Note that the above applies equally to spaces containing vectors or matrices.

pictorally:



We say $\| \ \|_+$, $\| \ \|_*$ induce same topology on V.

Convergence in one norm implies convergence in another.

Norm equivalence

A rather important and useful fact is that any two norms on the same finite-dimensional vector space are *equivalent*.

Theorem (All norms on a finite-dimensional space are equivalent)

Let V be an n-dimensional vector space, and let $\|\cdot\|_*$ and $\|\cdot\|_+$ be any two norms on this space. Then there are strictly positive constants c and k such that for all $\mathbf{x} \in V$,

$$c\|\mathbf{x}\|_* \leq \|\mathbf{x}\|_+ \leq k\|\mathbf{x}\|_*.$$

The constants c and k can depend on V (in particular n) and the choice of $\|\cdot\|_*$ and $\|\cdot\|_+$, but not on $\mathbf{x}$.

Note that the above applies equally to spaces containing vectors or matrices.

The good news: Norm equivalence suggests it doesn't matter which norm you pick. → theoretically

The bad news: To prove something, it typically matters which norm you pick. → practically,

Typically, we like $p = 1, 2, \infty$ for vectors, and the Frobenius norm or induced $p = 1, 2, \infty$ for matrices.

Example

Compute c and k such that,

$$c\|\mathbf{x}\|_1 \leq \|\mathbf{x}\|_2 \leq k\|\mathbf{x}\|_1, \qquad \forall\, \mathbf{x} \in \mathbb{C}^n.$$

Also, identify examples of vectors $\mathbf{x}$ that achieve the upper and lower bounds above.

for K:

this one picks up non-neg. cross terms

$$\|x\|_2^2 = \sum_i |x_i|^2 \leq \left(\sum_i |x_i|\right)^2 = \|x\|_1^2$$

$\Rightarrow K = 1$ works.

For $x = e_1 = (1, 0, \ldots, 0)$, $\|x\|_2^2 = 1 = \|x\|_1$, so $K = 1$ is sharp.

for c:

vec of $|x_i|$ vec of 1 C.S.

$$\|x\|_1 = \sum_i |x_i| = \langle |\underline{x}|, \underline{1} \rangle \leq \| |x| \|_2 \, \|\underline{1}\|_2 = \|x\|_2 \sqrt{n}$$

Thus $c = \frac{1}{\sqrt{n}}$ works. Sharp for $\underline{x} = \underline{1}$.

A zoo of norms, formulas, and equivalences: Matrix norm bounds

Example Recall: $\frac{1}{\sqrt{n}} \|x\|_1 \leq \|x\|_2 \leq \|x\|_1$

Compute c and k such that,

$$c\|\mathbf{A}\|_1 \leq \|\mathbf{A}\|_2 \leq k\|\mathbf{A}\|_1, \qquad \forall\ \mathbf{A} \in \mathbb{C}^{m\times n}$$

Also, identify examples of matrices $\mathbf{A}$ that achieve the upper and lower bounds above.

Start with c:

$$\|A\|_1 = \sup_{x\neq 0} \frac{\|\underbrace{Ax}_{\in \mathbb{C}^m}\|_1}{\|x\|_1} \leq \sup_{x\neq 0} \frac{\sqrt{m}\|Ax\|_2}{1 \cdot \|x\|_2} = \sqrt{m}\|A\|_2$$

$\therefore c = \frac{1}{\sqrt{m}}$ works...

Check: sharp w/ $A = e_1 1^*$

For k:

$$\|A\|_2 = \sup_{x\neq 0} \frac{\|Ax\|_2}{\|\underbrace{x}_{\in \mathbb{C}^n}\|_2} \leq \sup_{x\neq 0} \frac{\|Ax\|_1}{\frac{1}{\sqrt{n}}\|x\|_1} \Rightarrow \|A\|_2 \leq \sqrt{n}\|A\|_1$$

Check: sharp for $A = 1 e_1^*$

Thus

$$\boxed{\frac{1}{\sqrt{m}}\|A\|_1 \leq \|A\|_2 \leq \sqrt{n}\|A\|_1}$$

A zoo of norms, formulas, and equivalences: A matrix 2-norm

Example

Compute $\|\mathbf{A}\|_2$, where,

$$\mathbf{A} = \begin{pmatrix} -2 & -1 \\ 1 & 2 \end{pmatrix}$$

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

$\begin{cases} \text{maximize } \|Ax\| \text{ with constraint} \\ \|x\|_2 = 1. \end{cases}$

$$\|x\|_2^2 = 1 \Rightarrow x_1^2 + x_2^2 = 1 \Rightarrow x_2 = \pm\sqrt{1-x_1^2}.$$

Then

$$Ax = \begin{pmatrix} -2 & -1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -2x_1 - x_2 \\ x_1 + 2x_2 \end{pmatrix}$$

$$\|Ax\|_2^2 = (-2x_1 - x_2)^2 + (x_1 + 2x_2)^2 = 4x_1^2 + 4x_1x_2 + x_2^2 + x_1^2 + 4x_1x_2 + 4x_2^2$$

$$= 5x_1^2 + 8x_1x_2 + 5x_2^2$$

Sub $x_2 = \sqrt{1-x_1^2}$:

$$\|Ax\|_2^2 = 5x_1^2 + 8x_1\sqrt{1-x_1^2} + 5(1-x_1^2)$$

$$= 5 + 8x_1\sqrt{1-x_1^2} =: f(x_1)$$

Then,

since $x_1 \in [0,1]$ we check: $f(0) = 5$, $f(1) = 5$

$$\frac{d}{dx_1} f(x_1) = 8\sqrt{1-x_1^2} + 8x_1\left(\frac{1}{2}(1-x_1^2)^{-\frac{1}{2}}(-2x_1)\right)$$

solve $\frac{d}{dx_1} f(x_1) = 0 \Rightarrow x_1 = \pm 1/\sqrt{2}$. , take positive for max f

Then $f(1/\sqrt{2}) = 9 \Rightarrow \|A\|_2^2 = 9 \Rightarrow \|A\|_2 = 3$.

Hilbertian structure

Of special interest are the norms arising from inner products: these norms induce Euclidean-like geometry (Hilbert spaces).

The prototypical example on $\mathbb{C}^n$ is the ℓ^2 norm: for $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, we have,

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^* \mathbf{x} = \sum_{i \in [n]} x_i y_i^*, \qquad \|\mathbf{x}\|_2^2 = \langle \mathbf{x}, \mathbf{x} \rangle .$$

Inner products are *bilinear* forms over $\mathbb{R}$ and *sesquilinear* forms over $\mathbb{C}$.

Other examples: Let $Q \in \mathbb{C}^{n \times n}$ be Hermitian positive definite,

then $\langle x, y \rangle_Q := \langle Qx, y \rangle = y^* Q x$; $\|x\|_Q^2 = \langle Qx, x \rangle$

can show

$$\sqrt{\lambda_{min} Q}\, \|x\|_2 \leq \|x\|_Q \leq \sqrt{\lambda_{max} Q}\, \|x\|_2 .$$

; standard is $Q = I$.

Hilbertian structure

Of special interest are the norms arising from inner products: these norms induce Euclidean-like geometry (Hilbert spaces).

The prototypical example on $\mathbb{C}^n$ is the ℓ^2 norm: for $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, we have,

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^* \mathbf{x} = \sum_{i \in [n]} x_i y_i^*, \qquad \|\mathbf{x}\|_2^2 = \langle \mathbf{x}, \mathbf{x} \rangle .$$

$\langle a\underline{x} + y, \underline{z} \rangle = a \langle \underline{x}, \underline{z} \rangle + \langle y, z \rangle$

$\langle \underline{x}, a y + \underline{z} \rangle = a^* \langle x, y \rangle + \langle x, z \rangle$

Inner products are *bilinear* forms over $\mathbb{R}$ and *sesquilinear* forms over $\mathbb{C}$.

One of the most useful algebraic properties of inner products that give rise to a norm $\|\cdot\|$ is the *Cauchy-Schwarz* inequality:

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\|$$

From this property one can observe that the following geometric structure of elements $\mathbf{x}, \mathbf{y}$ is reasonable:

$$\cos\left(\angle(\mathbf{x}, \mathbf{y})\right) := \frac{|\langle \mathbf{x}, \mathbf{y} \rangle|}{\|\mathbf{x}\| \, \|\mathbf{y}\|}, \qquad \mathbf{x} \perp \mathbf{y} \text{ iff } \langle \mathbf{x}, \mathbf{y} \rangle = 0.$$

might be complex valued

we used this for only $\|\cdot\|_2$ on $\mathbb{R}^n$ above, but it holds generally.

Hilbertian structure

Of special interest are the norms arising from inner products: these norms induce Euclidean-like geometry (Hilbert spaces).

The prototypical example on $\mathbb{C}^n$ is the ℓ^2 norm: for $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, we have,

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^* \mathbf{x} = \sum_{i \in [n]} x_i y_i^*, \qquad\qquad \|\mathbf{x}\|_2^2 = \langle \mathbf{x}, \mathbf{x} \rangle .$$

Inner products are *bilinear* forms over $\mathbb{R}$ and *sesquilinear* forms over $\mathbb{C}$.

One of the most useful algebraic properties of inner products that give rise to a norm $\|\cdot\|$ is the *Cauchy-Schwarz* inequality:

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\|$$

From this property one can observe that the following geometric structure of elements $\mathbf{x}, \mathbf{y}$ is reasonable:

$$\cos\left(\angle(\mathbf{x}, \mathbf{y})\right) \coloneqq \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \, \|\mathbf{y}\|}, \qquad\qquad \mathbf{x} \perp \mathbf{y} \text{ iff } \langle \mathbf{x}, \mathbf{y} \rangle = 0.$$

There are Hilbertian norms even for matrices, with a common example being the Frobenius norm:

$$\langle \mathbf{A}, \mathbf{B} \rangle_F \coloneqq \mathrm{Tr}\left(\mathbf{B}^* \mathbf{A}\right), \qquad\qquad \|\mathbf{A}\|_F^2 = \langle \mathbf{A}, \mathbf{A} \rangle_F$$

yields geometry for spaces of matrices

The Pythagorean Theorem

The sledgehammer killing an ant way to prove the Pythagorean Theorem: Let $\mathbf{x}_1$, $\mathbf{x}_2$ be two orthogonal vectors (say in $\mathbb{C}^n$).

Since $\langle \mathbf{x}_1, \mathbf{x}_2 \rangle = 0$, then

$$\begin{aligned} \|\mathbf{x}_1 + \mathbf{x}_2\|_2^2 &= \langle \mathbf{x}_1 + \mathbf{x}_2, \mathbf{x}_1 + \mathbf{x}_2 \rangle \\ &= \langle \mathbf{x}_1, \mathbf{x}_1 \rangle + \langle \mathbf{x}_2, \mathbf{x}_2 \rangle + \underbrace{\langle \mathbf{x}_1, \mathbf{x}_2 \rangle}_{0} + \underbrace{\langle \mathbf{x}_2, \mathbf{x}_1 \rangle}_{0} \\ &= \|\mathbf{x}_1\|_2^2 + \|\mathbf{x}_2\|_2^2 . \end{aligned}$$

The Pythagorean Theorem

The sledgehammer killing an ant way to prove the Pythagorean Theorem: Let $\mathbf{x}_1$, $\mathbf{x}_2$ be two orthogonal vectors (say in $\mathbb{C}^n$).

Since $\langle \mathbf{x}_1, \mathbf{x}_2 \rangle = 0$, then

$$\begin{aligned}
\|\mathbf{x}_1 + \mathbf{x}_2\|_2^2 &= \langle \mathbf{x}_1 + \mathbf{x}_2, \mathbf{x}_1 + \mathbf{x}_2 \rangle \\
&= \langle \mathbf{x}_1, \mathbf{x}_1 \rangle + \langle \mathbf{x}_2, \mathbf{x}_2 \rangle + \underbrace{\langle \mathbf{x}_1, \mathbf{x}_2 \rangle}_{0} + \underbrace{\langle \mathbf{x}_2, \mathbf{x}_1 \rangle}_{0} \\
&= \|\mathbf{x}_1\|_2^2 + \|\mathbf{x}_2\|_2^2 .
\end{aligned}$$

One of the more useful extensions of this (not apparent from $n = 2$) is: If $\mathbf{x}_1, \ldots \mathbf{x}_k$ are k mutually orthogonal vectors in $\mathbb{C}^n$, then,

$$\left\| \sum_{j \in [k]} \mathbf{x}_j \right\|_2^2 = \sum_{j \in [k]} \|\mathbf{x}_j\|_2^2$$

Indeed:

$$\left\| \sum x^i \right\|_2^2 = \left\langle \sum_i x^i , \sum_j x^j \right\rangle = \sum_{ij} \langle x^i, x^j \rangle = \sum_{ij} \delta_{ij} \langle x^i, x^j \rangle = \sum_k \langle x^k, x^k \rangle = \sum \|x^k\|_2^2$$

References

Atkinson, Kendall (1989). *An Introduction to Numerical Analysis*. New York: Wiley. ISBN: 978-0-471-62489-9.
Salgado, Abner J. and Steven M. Wise (2022). *Classical Numerical Analysis: A Comprehensive Course*. Cambridge: Cambridge University Press. ISBN: 978-1-108-83770-5. DOI: 10.1017/9781108942607.
Trefethen, Lloyd N. and David Bau (1997). *Numerical Linear Algebra*. SIAM: Society for Industrial and Applied Mathematics. ISBN: 0-89871-361-7.